Le Frido 2022,
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 2023. Téléchargeable sur

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

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

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

3. Et bien entendu les sources LaTeX

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

Copyright 2011-2022 Laurent Claessens, and many contributors. A complete list could be retrieved from the git log.
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the chapter entitled "GNU Free Documentation License".

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

ISBN : 979-10-97085-30-8

Thèmes

Ceci est une sorte d'index thématique.

1 : cardinalité
2 : morphismes et compagnie
3 : arithmétique modulo, théorème de Bézout
4 : équations diophantiennes
5 : types d'anneaux
6 : sous-groupes
7 : groupe symétrique
8 : action de groupe
9 : classification de groupes
10 : produit semi-direct de groupes
11 : théorie des représentations
12 : espaces vectoriels
13 : rang
14 : formes bilinéaires et quadratiques
15 : endomorphismes cycliques
16 : extension de corps et polynômes
17 : polynômes
18 : polynôme d'endomorphismes
19 : dualité
20 : injections
21 : zoologie de l'algèbre
22 : tribu, algèbre de parties, λ-systèmes et co.
23 : théorie de la mesure
24 : normes
25 : inégalités
26 : constructions topologiques
27 : espaces métriques, normés
28 : limite et continuité
29 : intégration
30 : connexité
31 : compacts
32 : densité
33 : application réciproque
34 : applications continues et bornées
35 : suite de Cauchy, espace complet
36 : caractérisations séquentielles
37 : valeurs propres, définie positive
38 : norme matricielle, norme opérateur et rayon spectral
39 : série de matrices
40 : décomposition de matrices
41 : systèmes d'équations linéaires
42 : réduction, diagonalisation
43 : déterminant
44 : espaces de fonctions
45 : fonctions Lipschitz
46 : suites et séries
47 : suite de fonctions
48 : exponentielle
49 : logarithme
50 : fonction puissance
51 : sommation finie et infinie
52 : polynôme de Taylor
53 : formule des accroissements finis
54 : dérivation
55 : différentiabilité
56 : équations différentielles
57 : convexité
58 : espaces de Hilbert, base hilbertienne
59 : analyse complexe, fonctions holomorphes
60 : permuter des limites
61 : déduire la nullité d'une fonction depuis son intégrale
62 : inversion locale, fonction implicite
63 : points fixes
64 : changement de variables
65 : techniques de calcul
66 : méthodes de calcul
67 : méthode de Newton
68 : opérations sur les distributions
69 : convolution
70 : séries de Fourier
71 : transformée de Fourier
72 : gaussienne
73 : lemme de transfert
74 : invariants de similitude
75 : isométries
76 : enveloppes
77 : intégration sur des variétés
78 : dénombrements
79 : caractérisation de distributions en probabilités
80 : théorème central limite
81 : probabilités et espérances conditionnelles

Thème 1 : cardinalité Le Frido[1] ne définit pas la notion de nombre cardinal ; ça nous mènerait trop loin. Au lieu de cela, nous allons nous contenter des notions d'équipotence, surpotence et subpotence, et démontrer un certain nombre de résultats en utilisant sans retenue le lemme de Zorn 1.22.

1. Ici je mets la référence [1] ; pas parce qu'elle est utile ici, mais parce que je veux être sûr qu'elle soit numéro 1 de façon à être facilement reconnaissable.

(1) Cardinal d'un ensemble fini, définition 1.120.

(2) Définition d'équipotence, surpotence et subpotence, notations $A \succ B$ et $A \approx B$, définition 1.109.

(3) Toute partie d'un ensemble fini est finie, lemme 1.113.

(4) Si A est un ensemble fini ou dénombrable, alors il existe une surjection $\mathbb{N} \to A$, lemme 1.125.

(5) Si A est un ensemble infini et si $f\colon A \to B$ est une application injective, alors $f(A)$ est infini, proposition 1.118.

(6) Toute partie infinie de $\mathbb{N}$ est dénombrable, proposition 1.126

(7) Une bijection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$, proposition 1.129.

(8) Une décomposition de $\mathbb{N}$ en une infinité de parties équipotentes à $\mathbb{N}$, corolaire 1.130.

(9) Si il existe une surjection $\mathbb{N} \to A$, alors A est fini ou dénombrable, lemme 1.131.

(10) Une union dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable, proposition 1.132

(11) Tout ensemble infini contient une partie en bijection avec $\mathbb{N}$, proposition 1.135.

(12) Toute partie d'un ensemble fini est finie, et toute partie d'un ensemble dénombrable est finie ou dénombrable, proposition 1.136.

(13) Si $A \succeq B$ et $B \succeq A$, alors $A \approx B$, théorème de Cantor-Schröder-Bernstein 1.139

(14) Le théorème de Cantor 1.142 dit qu'il n'existe pas de surjection d'un ensemble vers son ensemble des parties. On en déduit qu'il n'existe pas d'ensemble contenant tous les ensembles (corolaire 1.144).

(15) Si A est infini et si $A \succeq B$, alors $A \approx A \cup B$ par le lemme 1.146.

(16) Si S est un ensemble infini alors il existe une bijection $\varphi\colon \{0,1\} \times S \to S$, proposition 1.147.

(17) Si A est infini, alors $A \times \mathbb{N} \approx A$, proposition 1.149.

(18) Si A est infini et si $B \prec A$, alors $A \backslash B \approx A$, lemme 1.151.

(19) Si A est infini, alors $A \approx A \times A$, théorème 1.155.

Il y a aussi des résultats de cardinalité autour des extensions de corps.

(1) Si $\mathbb{K}$ est un corps infini, alors $\mathbb{K}[X] \approx \mathbb{K}$.

(2) Le théorème de Steinitz 6.128 affirme que tout corps admet une unique clôture algébrique. La preuve utilise pas mal de cardinalité ainsi que le lemme de Zorn 1.22.

Thème 2 : morphismes et compagnie

(1) Un morphisme est un concept algébrique. Il s'agit d'une application (pas spécialement inversible) qui préserve la structure. Quand on parle de morphisme, il faut donc préciser la structure. On dit « morphisme de groupe », « morphisme d'espace vectoriel », « morphisme d'anneaux », etc.

(2) Morphisme de module, définition 1.274.

(3) Morphisme de groupes, définition 1.35.

(4) Un isomorphisme d'espaces topologiques est une application continue, inversible, dont l'inverse est continue, 7.44. On dit aussi un homéomorphisme.

(5) Un difféomorphisme est différentiable d'inverse différentiable, définition 11.172.

(6) Un C^k-difféomorphisme est une application C^k d'inverse C^k. Définition 11.172.

Le mot « homomorphisme » signifie exactement « morphisme », et, sauf incohérence de ma part, il n'est pas utilisé dans le Frido.

Thème 3 : arithmétique modulo, théorème de Bézout

(1) Pour $\mathbb{Z}^*$ c'est le théorème 1.228.

(2) Théorème de Bézout dans un anneau principal : corolaire 3.92.

(3) Théorème de Bézout dans un anneau de polynômes : théorème 6.44.

(4) En parlant des racines de l'unité et des générateurs du groupe unitaire dans le lemme 19.6. Au passage nous y parlerons de solfège.

(5) La proposition 1.231 qui donne tout entier assez grand comme combinaisons de a et b à coefficients positifs est utilisée en chaines de Markov, voir la définition 38.47 et ce qui suit.

(6) PGCD et PPCM sont dans la définition 1.215.

(7) Calcul effectif du PGCD puis des coefficients de Bézout : sous-sections 3.2.1.1 et 3.2.1.2.

Thème 4 : équations diophantiennes

(1) Équation $ax + by = c$ dans $\mathbb{N}$, équation (3.50).

(2) Dans 3.2.7, nous résolvons $ax + by = c$ en utilisant Bézout (théorème 1.228).

(3) L'exemple 3.113 donne une application de la pure notion de modulo pour $x^2 = 3y^2 + 8$. Pas de solutions.

(4) L'exemple 3.114 résout l'équation $x^2 + 2 = y^3$ en parlant de l'extension $\mathbb{Z}[i\sqrt{2}]$ et de stathme.

(5) Les propositions 3.118 et 3.120 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 décompositions de fractions en éléments simples, théorème 26.53.

(7) La proposition 1.129 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.

Thème 5 : types d'anneaux

(1) Définition d'anneau : définition 1.38.

(2) La définition d'anneau principal est 3.75 ; pour un idéal principal, c'est 3.74.

(3) $\mathbb{Z}$ est intègre, exemple 1.246, principal et euclidien (proposition 3.108).

(4) $\mathbb{Z}[X]$ n'est pas principal (voir (5)).

(5) Si A est un anneau intègre [2] qui n'est pas un corps, alors $A[X]$ n'est pas principal, lemme 3.109.

(6) L'anneau des fonctions holomorphes sur un compact donné est principal, proposition 26.23.

(7) L'anneau $\mathbb{Z}[i\sqrt{3}]$ n'est pas factoriel, exemple 3.73.

(8) L'anneau $\mathbb{Z}[i\sqrt{5}]$ n'est ni factoriel ni principal, exemple 3.99.

(9) Tous les idéaux de $\mathbb{Z}/6\mathbb{Z}$ sont principaux, mais $\mathbb{Z}/6\mathbb{Z}$ n'est pas principal. Exemple 3.85.

Thème 6 : sous-groupes

(1) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.293.

(2) Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 13.40 ou proposition 13.41.

Thème 7 : groupe symétrique

(1) Définition 1.179.

(2) La signature $\epsilon\colon S_n \to \{-1, 1\}$ est l'unique morphisme surjectif de S_n sur $\{-1, 1\}$, proposition 1.205(1).

(3) La table des caractères du groupe symétrique S_4 est donné dans la section 16.5.

2. Définition 1.221.

(4) Le groupe symétrique S_4 est le groupe des symétries affines du tétraèdre régulier [3], proposition 18.188.

(5) Le groupe alterné A_5 est l'unique groupe simple d'ordre [4] 60, proposition 5.40.

(6) La proposition 5.30 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 8 : action de groupe

(1) Définition d'une action de groupe sur un ensemble : 1.311.

(2) Action du groupe modulaire sur le demi-plan de Poincaré, théorème 23.95.

(3) La formule de Burnside (théorème 2.40) 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, 18.9.15.1 et 18.9.15.2.

(5) Le groupe symétrique S_n agit sur l'anneau $\mathbb{K}[T_1, \ldots, T_n]$, lemme 1.312.

Thème 9 : classification de groupes

(1) Ordre d'un groupe, définition 1.172.

(2) Structure des groupes d'ordre pq, théorème 5.25.

(3) Le groupe alterné est simple, théorème 5.36.

(4) Définition 5.6 d'un p-groupe.

(5) Théorème de Sylow 5.11.

(6) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.293.

(7) $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corolaire 19.37.

Thème 10 : produit semi-direct de groupes

(1) Définition 2.47.

(2) Le corolaire 2.49 donne un critère pour prouver qu'un produit NH est un produit semi-direct.

(3) L'exemple 18.170 donne le groupe des isométries du carré comme un produit semi-direct.

(4) Le théorème 5.25 classifie les groupes d'ordre pq (p, q premiers distincts) à grands coups de produit semi-directs.

(5) Le théorème 18.74 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 18.76 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 8.63 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 11 : théorie des représentations

(1) Définition 4.126.

(2) Table des caractères du groupe diédral, section 18.16.

(3) Table des caractères du groupe symétrique S_4, section 16.5.

3. Définition 12.143.
4. Ordre d'un groupe, définition 1.172.

Thème 12 : espaces vectoriels

(1) Les notations $\mathcal{L}(V, W)$, $L(V, W)$, $\mathrm{GL}(V, W)$ ainsi que les notions de morphismes et d'isomorphismes d'espaces vectoriels normés sont dans la définition 11.179.

(2) Existence d'une base. Pour un espace vectoriel quelconque, proposition 4.22.

(3) Théorème de la base incomplète. Pour un espace vectoriel quelconque, théorème 4.23.

Thème 13 : rang

(1) Définition pour une application linéaire : 4.43, pour une matrice : 4.99. L'équivalence est la proposition 4.103.

(2) Le théorème du rang, théorème 4.44

(3) Pour une applicaton linéaire entre deux espaces vectoriels de même dimension finie, il est équivalent d'être injectif, surjectif ou bijectif, c'est le corolaire 4.46.

(4) Pour prouver que des matrices sont équivalentes et pour les mettre sous des formes canoniques, nous avons le lemme 4.105 et son corolaire 4.106.

(5) Tout hyperplan de $\mathbb{M}(n, \mathbb{K})$ coupe $\mathrm{GL}(n, \mathbb{K})$, corolaire 4.106. Cela utilise la forme canonique sus-mentionnée.

(6) Le lien entre application duale et orthogonal de la proposition 9.180 utilise la notion de rang.

(7) Le lemme 9.283 parle de commutant et utilise la notion de rang. Ce lemme sert à prouver diverses conditions équivalentes à être un endomorphisme cyclique dans le théorème 9.284.

Thème 14 : formes bilinéaires et quadratiques

(1) Les formes bilinéaires sont définies en 9.118.

(2) Forme quadratique, définition 9.126.

(3) Matrice d'une forme quadratique, définition 9.131.

(4) Équivalence de forme quadratiques, définition 9.153. Deux formes quadratiques sont équivalentes si et seulement si elles ont même signature, proposition 9.154.

(5) Une isométrie d'une forme bilinéaire est affine ou linéaire, théorème 9.141.

(6) Forme bilinéaire dégénérée, définition 9.121.

(7) Une forme bilinéaire est non-dégénérée si et seulement si sa matrice associée est inversible, c'est la proposition 9.214.

(8) Une isométrie d'une forme bilinéaire est linéaire ou affine par le théorème 9.141.

(9) Toute forme quadratique admet des bases orthogonales, théorème 9.227 pour le cas général ; proposition 9.236 pour le cas de $\mathbb{R}^n$, en se basant sur le théorème spectral.

(10) Théorème de Sylvester à propos de signature (définition 9.147) de forme quadratique réelle : 9.233.

(11) Base q-orthogonale pour une forme quadratique, théorème 9.227.

(12) Le concept de projection orthogonale est la définition 12.138 en dimension finie et la définition 25.7 dans le cas des espaces de Hilbert.

(13) Produit hermitien, définition 9.163. Opérateur hermitien ($A^\dagger = A$), opérateur unitaire ($A^\dagger A = 1$), définition 9.166

Thème 15 : endomorphismes cycliques

(1) Définition 9.98.

(2) Groupe cyclique, définition 1.270.

(3) Son lien avec le commutant donné dans la proposition 9.281 et le théorème 9.284.

(4) Utilisation dans le théorème de Frobenius (invariants de similitude), théorème 9.272.

Thème 16 : extension de corps et polynômes

(1) Définition d'une extension de corps 6.55.

(2) Définition de polynôme minimal : 6.60.

(3) Pour l'extension du corps de base d'un espace vectoriel et les propriétés d'extension des applications linéaires, voir la section 9.14.

(4) Extension de corps de base et similitude d'application linéaire (ou de matrices, c'est la même chose), théorème 9.286.

(5) Extension de corps de base et cyclicité des applications linéaires, corolaire 9.285.

(6) À propos d'extensions de $\mathbb{Q}$, le lemme 6.170.

(7) Corps de rupture : définition 6.107 existence par la proposition 6.113. Il n'y a pas unicité.

(8) Corps de décomposition : définition 6.134. Attention : le plus souvent nous parlons de corps de décomposition d'un seul polynôme. Cette définition est un peu surfaite. Existence par la proposition 6.135 qui le donne même comme extension par toutes les racines, et unicité à isomorphisme près par le théorème 6.137, énoncé de façon plus simple (mais pas indépendante !) en la proposition 6.138.

(9) Si $\mathbb{K}$ est algébrique clos et si $\alpha\colon \mathbb{K} \to \mathbb{L}$ est une extension algébrique, alors $\alpha(\mathbb{K}) = \mathbb{L}$ par le lemme 6.74.

Un trio de résultats d'enfer est :

(1) Dans un anneau principal qui n'est pas un corps, un idéal est maximal si et seulement si il est engendré par un irréductible (proposition 3.81).

(2) Dans un anneau, un idéal I est maximal si et seulement si A/I est un corps (proposition 3.52)

(3) Si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ est principal (lemme 3.132).

Thème 17 : polynômes

Définitions Soient un anneau A, un corps $\mathbb{K}$, une extension $\mathbb{L}$ de $\mathbb{K}$ et un élément $\alpha \in \mathbb{L}$.

(1) La définition la plus formelle est en tant que module produit $A^{(\mathbb{N})}$, définition 1.303. Le produit et l'évaluation sont définis en 1.306 et la formule $(PQ)(x) = P(x)Q(x)$ dans 1.307.

(2) Si $\mathbb{A}$ est commutatif, alors $\mathbb{A}[X]$ est également commutatif, lemme 1.308.

(3) En ce qui concerne la notation $A[X]$, elle ne devrait pas être utilisée, voir 1.16.2. L'ensemble des polynômes sera noté $\mathcal{P}(A)$ et ceux de degré n (définition 1.305), $\mathcal{P}_n(A)$.

(4) $A[X]$, définition 1.303 ; l'anneau $\mathbb{K}[X]$ a même définition parce que c'est un cas particulier. L'évaluation d'un polynôme en un élément de l'anneau, $P(\alpha)$ est définie en 1.306.

(5) Liens entre $\mathbb{K}[\alpha]$, $\mathbb{K}[X]$, $\mathbb{K}(\alpha)$ et $\mathbb{K}(X)$ lorsque α est transcendant, proposition 6.93. Et la proposition 6.96 pour le cas où α est algébrique [5].

(6) Si A est un anneau et si α est un élément d'une extension de A (comme anneau), nous écrivons $A[\alpha]$ pour le plus petit sous-anneau de B contenant A et α. C'est la définition 3.124.

(7) $\mathbb{K}(X)$, le corps des fractions de $\mathbb{K}[X]$, définition 6.78. Si $R = P/Q$ dans $\mathbb{K}(X)$, l'évaluation est $R(\alpha) = P(\alpha)Q(\alpha)^{-1}$, définition 6.79.

(8) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est le plus petit corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, définition 6.80.

(9) À propos de polynômes à plusieurs variables.

— Anneau de polynômes : $A[X_1, \dots, X_n]$ est la définition 3.53.

— Corps engendré : $\mathbb{K}(\alpha_1, \dots, \alpha_n)$ est la définition 6.131.

— Corps des fractions rationnelles : $\mathbb{K}(X_1, \dots, X_n)$ est la définition 6.132.

5. Définition 6.66.

Coefficients dans un anneau commutatif (1) Les polynômes à coefficients dans un anneau commutatif sont à la section 3.10.

Coefficients dans un corps Les polynômes à coefficients dans un corps sont à la section 6.3.

(1) Nous parlons de l'idéal des polynômes annulateurs dans le théorème 6.40.

(2) Le théorème 6.40 dit que $\mathbb{K}[X]$ est un anneau principal et que tous ses idéaux sont engendrés par un unique polynôme unitaire.

(3) Le polynôme minimal est irréductible, proposition 6.63.

(4) Quelques formules sur le pgcd, lemme 6.53.

Polynôme primitif (1) Un polynôme est irréductible sur A si et seulement si il est irréductible et primitif sur le corps des fractions, corolaire 3.152.

Polynôme d'endomorphisme C'est la section 9.6.

Racines et factorisation (1) Si $\mathbb{A}$ est un anneau, la proposition 3.142 factorise une racine.

(2) Si $\mathbb{A}$ est un anneau, la proposition 3.146 factorise une racine avec sa multiplicité.

(3) Si $\mathbb{A}$ est un anneau, le théorème 3.148 factorise plusieurs racines avec leurs multiplicités.

(4) Le théorème 3.148 nous indique que lorsqu'on a autant de racines (multiplicité comprise) que le degré, alors nous avons toutes les racines.

(5) 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 6.94.

(6) Le théorème 6.104 annule un polynôme de degré n ayant $n+1$ racines distinctes.

(7) La proposition 6.173 nous annule un polynôme à plusieurs variables lorsqu'il a trop de racines.

(8) En analyse complexe, le principe des zéros isolés 17.137 annule en gros toute série entière possédant un zéro non isolé.

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

Thème 18 : polynôme d'endomorphismes

(1) Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 9.281.

(2) Racine carrée d'une matrice hermitienne positive, proposition 13.27.

(3) Théorème de Burnside sur les sous-groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.293.

(4) Décomposition de Dunford, théorème 9.245.

(5) Algorithme des facteurs invariants 4.107.

Thème 19 : dualité Ne pas confondre dual algébrique et dual topologique d'un espace vectoriel.

(1) Définition du dual, 4.119.

(2) Le dual d'un espace de Banach est de Banach, proposition 7.232.

(3) Définition de la base duale 4.120.

(4) Base préduale (existence, unicité) : proposition 4.123.

(5) Théorème de représentation de Riez 27.162 : $L^p = (L^q)'$, et en particulier $L^\infty = (L^1)'$.

(6) Il n'est pas vrai que $(L^\infty)' = L^1$, voir la proposition 27.185.

(7) Dans un espace de Banach[6], $\|x\| = \max_{\substack{\varphi \in E' \\ \|\varphi\|=1}} |\varphi(x)|$, proposition 27.152.

Dual topologique et algébrique Ils sont définis par 4.119. 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).

6. Espace de Banach, définition 7.230.

Topologie Une topologie possible sur le dual d'un espace vectoriel topologique est celle $*$-faible de la définition 7.298.

Nous comparons les topologies faibles et de la norme en la section 11.14.

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

(1) Le théorème de représentation de Riesz 25.18 pour les espaces de Hilbert.

(2) La proposition 27.160 pour les espaces $L^p\big([0,1]\big)$ avec $1 < p < 2$.

(3) Le théorème de représentation de Riesz 27.162 pour les espaces L^p en général.

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

Thème 20 : injections

(1) L'espace de Sobolev $H^1(I)$ s'injecte de façon compacte dans $C^0(\bar{I})$, proposition 31.6.

(2) L'espace de Sobolev $H^1(I)$ s'injecte de façon continue dans $L^2(I)$, proposition 31.6.

(3) L'espace $L^2(\Omega)$ s'injecte continument dans $\mathscr{D}'(\Omega)$ (les distributions), proposition 30.24.

Thème 21 : zoologie de l'algèbre Nous listons ici un peu tous les termes qui arrivent en algèbre.

Structures Les structures.

(0a) Algèbre, définition 1.291.

Éléments Pour les éléments, nous avons :

(0a) Élément irréductible en 3.67.

(0b) Élément premier en 1.249.

Anneaux Pour les anneaux, nous avons :

(0a) Anneau factoriel en 3.71.

(0b) Anneau principal en 3.75.

(0c) Anneau intègre en 1.221.

(0d) Anneau noetherien en 3.95.

Idéaux Pour les idéaux, nous avons :

(0a) Idéal principal en 3.74.

(0b) Idéal premier en 3.76.

Thème 22 : 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 14.16 sur un espace mesuré avec une tribu, définition 14.1.

(2) Une mesure extérieure est la définition 14.8

(3) Une algèbre de partie : définition 14.11. Une mesure sur une algèbre de parties : définition 14.9. 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 14.75.

(4) Un λ-système : définition 14.26.

(5) Une mesure complexe : définition 14.208.

En théorie de l'intégration, si X est une partie de $\mathbb{R}^n$, la convention est de considérer des fonctions

$$f\colon \big(X, \mathcal{L}\mathrm{eb}(X)\big) \to \big(\mathbb{R}, \mathcal{B}\mathrm{or}(\mathbb{R})\big).$$

Voir les points 14.111 et 14.155 pour les conventions à ce propos.

À propos d'applications mesurables :

(1) Définition d'une application mesurable, définition 14.38.
(2) Une fonction continue est borélienne, théorème 14.51.
(3) Si les f_n sont mesutables (au sens des boréliens), alors $\sup_n f_n$ est mesurable, lemme 14.93.

À propos de tribu induite :

(1) Définition 14.6.
(2) Les boréliens induits sont bien les boréliens de la topologie induite : $\mathcal{B}\mathrm{or}(Y) = \mathcal{B}\mathrm{or}(X)_Y$, théorème 14.52.

Thème 23 : théorie de la mesure

Mesure À propos de mesure.

(1) Tribu des boréliens, définition 14.45.
(2) Mesure positive, mesure finie et σ-finie, c'est la définition 14.16.
(3) Le produit de tribus est donné par la définition 14.117,
(4) Produit d'une mesure par une fonction, définition 14.195.
(5) le produit d'espaces mesurés est donné par la définition 14.218.
(6) Mesure de Lebesgue sur $\mathbb{R}$, définition 14.132.
(7) Une partie de $\mathbb{R}$ non mesurable au sens de Lebesgue : l'exemple 14.146.
(8) Mesure de Lebesgue sur $\mathbb{R}^N$, définition 14.220.
(9) Mesure à densité, définition 14.195.

Théorèmes d'approximation Il est important de pouvoir approcher des fonctions continues ou L^p par des fonctions étagées, sinon on ne parvient pas à faire tourner la machine de l'intégration de Lebesgue.

(1) Si $(S, \mathcal{A}, \mu)$ est un espace mesuré et si $f\colon S \to [0, +\infty]$ est une fonction mesurable, le théorème fondamental d'approximation 14.110 dit qu'il existe une suite croissante de fonctions étagées qui converge vers f.
(2) Les fonctions simples sont denses dans L^p, proposition 27.47.
(3) Encadrement d'un borélien A par un fermé F et un ouvert V par le lemme 14.211 : $F \subset A \subset V$ avec $\mu(V \backslash F) < \epsilon$.
(4) Approximation L^p de la fonction caractéristique d'un borélien par une fonction continue par le théorème 14.213.

Thème 24 : normes

Définitions (1) Espace vectoriel normé : définition 7.136.

(2) Produit scalaire, définition 9.155.
(3) Norme associée à un produit scalaire (cas réel), théorème 11.1.
(4) Norme associée à une forme sesquilinéaire (cas complexe), proposition 10.95.
(5) Produit scalaire sur $\mathbb{R}^n$, définition 9.160. Norme sur $\mathbb{R}^n$, définition 11.3.
(6) Forme hermitienne sur $\mathbb{C}^n$, définition 9.170. Norme sur $\mathbb{C}^n$, définition 10.96.
(7) La proposition 11.40 donne les normes $\|x\|_1$, $\|x\|_2$ et $\|x\|_\infty$ sur $\mathbb{R}^n$.
(8) Sur $\mathbb{R}^n$, la proposition 17.106 dit que $\|x\|_p$ est une norme.

Topologie (1) Métrique associée à une norme $d(x, y) = \|x - y\|$, définition 7.141.

(2) La topologie d'un espace vectoriel normé est la topologie métrique du théorème 7.98.
(3) Les boules sont les boules métrique définies en (7.91).

Inégalités (1) En général pour les normes $\|.\|_p$, il y a des inégalités dans 17.107 et 17.99.

(2) La proposition 17.112 donne l'inégalité $\|x\|_q \leqslant n^{\frac{1}{q}-\frac{1}{p}}\|x\|_p$ dès que $0 < q < p$.

Équivalence de norme (1) Définition de l'équivalence de norme 11.42.

(2) La proposition 11.43 sur l'équivalence des normes $\|.\|_2$, $\|.\|_1$ et $\|.\|_\infty$ dans $\mathbb{R}^n$.

(3) Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes par le théorème 11.45.

Autres (1) Montrer que le problème $a - b$ est stable dans l'exemple 34.27.

(2) La proposition 12.110 donnant $\rho(A) \leqslant \|A\|$ utilise l'équivalence de toutes les normes sur un espace vectoriel de dimension finie (théorème 11.45.).

(3) La norme $x \mapsto \|x\|$ est une application continue, proposition 7.144.

Norme opérateur et d'algèbre voir le thème 38.

Thème 25 : inégalités Dans $\mathbb{C}$ nous avons $|a + b| \leqslant |a| + |b|$ par la proposition 10.90(6).

Inégalité de Young Nous avons

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} \tag{-2.1}$$

par la proposition 27.30.

Inégalité de Jensen (1) Une version discrète pour $f\big(\sum_i \lambda_i x_i\big)$, la proposition 17.103.

(2) Une version intégrale pour $f\big(\int \alpha d\mu\big)$, la proposition 27.31.

(3) Une version pour l'espérance conditionnelle, la proposition 36.60.

Inégalité de Hölder Il en existe de nombreuses versions et variations.

(1) Hölder pour L^p : $\|fg\|_1 \leqslant \|f\|_p \|g\|_q$, proposition 27.33.

(2) Hölder pour ℓ^p : $\|x\|_q \leqslant n^{\frac{1}{q}-\frac{1}{p}} \|x\|_p$, proposition 17.112.

(3) $\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p} \|x\|_\infty$, théorème 17.107

(4) $\|x\|_p \leqslant n^{1/p} \|x\|_q$, corolaire 17.108.

Inégalité de Minkowsky (1) Pour une forme quadratique[7] q sur $\mathbb{R}^n$ nous avons $\sqrt{q(x+y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$. Proposition 11.10.

(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 27.38.

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

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

ou sous forme compacte

$$\left\| x \mapsto \int_Y f(x,y) d\nu(y) \right\|_p \leqslant \int_Y \|f_y\|_p d\nu(y)$$

C'est la proposition 27.40.

Transformée de Fourier Pour tout $f \in L^1(\mathbb{R}^n)$ nous avons $\|\hat{f}\|_\infty \leqslant \|f\|_1$, lemme 29.12.

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 27.60.

Thème 26 : constructions topologiques

topologie produit Si X et Y sont des espaces topologiques, nous pouvons construire une topologie sur $X \times Y$.

(1) La définition de la topologie produit est 7.14.

(2) Pour les espaces vectoriels normés, le produit est donné par la définition 7.191.

7. Définition 9.126.

(3) L'équivalence entre la topologie de la norme produit et la topologie produit est le lemme 7.191.

(4) Quand V et W sont des espaces métriques, la topologie considérée sur $V \times W$ est celle de la définition 7.191. C'est à la fois la topologie de la norme produit et la topologie produit.

(5) La convergence dans un espace vectoriel est si et seulement si il y a convergence composante par composante, proposition 7.56.

(6) Dans le cas d'espaces normés, la topologie produit est la même que celle de la norme produit, lemme 7.192.

topologie induite Si X est un espace topologique et si A est une partie de X, nous mettons une topologie sur A.

(1) Le topologie induite, définition 7.33.

(2) Si X est un espace vectoriel normé, la topologie induite est celle de la norme restreinte : lemme 7.102.

topologie quotient Si X est topologique et si $\sim$ est une relation d'équivalence, nous définissons une topologie sur $X/\sim$.

(1) La topologie quotient est définie en 7.20.

(2) Si X est vectoriel normé, la topologie sur $X/\sim$ est aussi donnée par une norme quotient de la définition 7.295. La proposition 7.297 donne l'équivalence entre la topologie quotient et la norme quotient.

topologie rendant continues des applications Si nous avons une application $f\colon X \to Y$ et si Y est topologique, nous mettons une topologie sur X rendant f continue.

(1) Définition de la plus petite topologie rendant des applications continues, proposition 7.16.

Thème 27 : espaces métriques, normés

(1) Un espace métrique est un ensemble muni d'une distance, définition 7.97.

(2) La distance entre un point et un ensemble est la définition 7.128.

(3) Le théorème-définition 7.98 donne la topologie sur un espace métrique en disant que les boules ouvertes sont une base de la topologie (définition 7.2).

(4) La définition de la convergence d'une suite est la définition 7.12.

(5) Dans un espace vectoriel normé, une application est continue si et seulement si elle est bornée, proposition 11.61.

(6) Un espace vectoriel topologique [8] qui possède en tout point une base dénombrable de topologie accepte une distance, théorème 7.242.

Thème 28 : limite et continuité

(1) Limite d'une fonction en un point : définition 7.92. Il n'y a pas unicité en général comme le montre l'exemple 7.50 dans un espace non séparé.

(2) Caractérisation de la limite dans $\mathbb{R}$, proposition 12.1.

(3) Unicité de la limite d'une suite dans un espace séparé : proposition 7.53. Unicité de la limite d'une fonction, toujours dans le cas d'un espace séparé : proposition 7.95.

(4) La proposition 7.300 donne l'unicité de la limite dans le cas des espaces duaux pour la topologie $*$-faible. La proposition 7.95 nous dira qu'il y a unicité dès que l'espace d'arrivée est séparé.

(5) Définition de la continuité d'une fonction en un point et sur une partie de l'espace de départ : définition 7.41.

8. Définition 7.148.

(6) Continuité sur une partie si et seulement si continue en chaque point, c'est le théorème 7.170.
(7) Voir l'exemple 12.55 traité en détail pour la (non) continuité d'une fonction qui fait un saut en un point.
(8) La fonction $f(x, y) = x + y$ est continue, lemme 10.27.

Thème 29 : intégration À propos d'intégration.

L'ordre dans lequel les choses sont faites — Nous commençons par considérer des fonctions $f : \Omega \to [0, +\infty]$ dans la définition 14.156.
— Nous donnerons ensuite quelques propriétés restreintes aux fonctions à valeurs positives, par exemple
(1) La convergence monotone 14.166,
(2) Lemme de Fatou 14.170.
(3) (presque) linéarité pour les fonctions positives, théorème 14.171.
— La définition pour les fonctions à valeurs dans $\mathbb{R}$ puis $\mathbb{C}$ est 14.174.
— Pour les fonctions à valeurs dans un espace vectoriel, c'est la définition 14.183.

primitive et intégrale (1) La définition 14.245 donne $\int_a^b f = \int_{]a,b[} f$ lorsque f est intégrable sur $]a, b[$.
(2) Lorsque f n'est pas intégrable sur $]a, b[$ nous pouvons poser $\lim_{x\to b} \int_a^b f$ et dire que c'est une intégrale impropre, définition 14.258.

Quelque résultats (1) Intégrale associée à une mesure, définition 14.156
(2) L'existence d'une primitive pour toute fonction continue est le théorème 12.427.
(3) La définition d'une primitive est la définition 12.198.
(4) Primitive et intégrale, proposition 14.246.
(5) Intégrale impropre, définition 14.258.

Intégrale et mesure (1) L'intégrale de la fonction 1 donne la mesure : $\int_B 1 d\mu = \mu(B)$, c'est le lemme 14.163.
(2) Le théorème de Radon-Nikodym 14.204 donne une densité pour certaines mesures.
(3) Le produit d'une mesure par une fonction donné par la définition 14.195 introduit aussi une densité : $(w \cdot \mu)(A) = \int_A w d\mu$.

Autre résultats (1) Si $A, B \subset \Omega$ sont des parties disjointes, alors $\int_{A\cup B} f = \int_A f + \int_B f$, proposition 14.180.
(2) La σ-additivité dénombrable, $\int_{\bigcup_i A_i} f d\mu = \sum_{i=0}^{\infty} \int_{A_i} f d\mu$ est dans les propositions 14.193 et 14.194.

Thème 30 : connexité

(1) Définition 7.60
(2) L'image d'un connexe par une fonction continue est connexe, proposition 7.184.
(3) Connexité par arcs, définition 10.56.
(4) Une partie de $\mathbb{R}^2$ qui est connexe, mais pas connexe par arcs, proposition 21.57.
(5) Une partie de $\mathbb{R}$ est connexe si et seulement si elle est un intervalle, proposition 10.47.
(6) Le groupe $\mathrm{SL}(n, \mathbb{K})$ est connexe par arcs, proposition 13.18.
(7) Le groupe $\mathrm{GL}(n, \mathbb{C})$ est connexe par arcs, proposition 13.19.
(8) Le groupe $\mathrm{GL}(n, \mathbb{R})$ a exactement deux composantes connexes par arcs, proposition 13.20.
(9) Le groupe $\mathrm{O}(n, \mathbb{R})$ n'est pas connexe, lemme 13.15.
(10) Les groupes $\mathrm{U}(n)$ et $\mathrm{SU}(n)$ sont connexes par arcs, lemme 13.16.
(11) Pour tout $n \geqslant 2$, le groupe $\mathrm{SO}(n)$ est connexe, le groupe $\mathrm{O}(n)$ a deux composantes connexes, proposition 13.4.
(12) Connexité des formes quadratiques de signature donnée, proposition 17.118.

Thème 31 : compacts

Propriétés générales Quelques propriétés de compacts.

(1) La définition d'un ensemble compact est la définition 7.66.

(2) Ne pas confondre le compactifié d'Alexandrov 7.88 avec la droite réelle achevée 12.27.

(3) Si M est un compact de $A \times B$, alors $M \subset K \times L$ où K est compact de A et L de B, proposition 7.86.

(4) Un fermé dans un compact est compact, lemme 7.82

(5) Dans un espace Hausdorff[9], les compacts sont fermés, 7.82(2).

(6) Un espace est compact si et seulement si toute intersection finie de fermé est non vide, théorème 7.91.

(7) Tout compact d'un espace topologique séparé est fermé, lemme 7.82(2).

(8) Dans un espace vectoriel réel de dimension finie, les compacts sont les fermés bornés par le théorème 10.21.

(9) Le théorème de Borel-Lebesgue 10.18 dit qu'un intervalle[10] de $\mathbb{R}$ est compact si et seulement si il est de la forme $[a, b]$.

(10) Théorème des fermés emboîtés dans le cas compact, corolaire 7.79. À ne pas confondre avec celui dans le cas des espaces métrique, théorème 7.259.

(11) L'image d'un compact par une fonction continue est un compact, théorème 7.186.

(12) Suites dans un compact

(12a) Toute suite dans un compact admet une sous-suite convergente, théorème 7.250.

(12b) Dans $\mathbb{R}^n$, toute suite dans un compact admet une sous-suite convergente, théorème 10.50. La démonstration de ce théorème est non seulement plus compliquée que le cas général, mais utilise en plus le cas dans $\mathbb{R}$; lequel cas n'est pas démontré de façon directe dans le Frido.

(12c) Un espace métrique est compact si et seulement si toute suite contient une sous-suite convergente. C'est le théorème de Bolzano-Weierstrass 7.124. La démonstration de ce théorème est indépendante.

(13) Une fonction continue sur un compact est bornée et atteint ses bornes, théorème 7.126.

(14) Une fonction continue sur un compact y est uniformément continue, théorème de Heine 12.78.

(15) Une bijection continue $f : K \to X$ entre un compact et un séparé est un isomorphisme, 7.187.

Produits de compacts À propos de produits de compacts. C'est un compact dans tous les cas métriques[11].

(1) Les produits d'espaces métriques compacts sont compacts. Il s'agit du théorème de Tykhonov que nous verrons ce résultat dans les cas suivants.

— $\mathbb{R}$, lemme 10.20.

— Produit fini d'espaces métriques compacts, théorème 7.275.

— Produit dénombrable d'espaces métriques compacts, théorème 7.277.

9. Hausdorff, définition 7.51.
10. Définition 1.20.
11. Si vous connaissez des exemples non métriques de produits de compacts qui ne sont pas compacts, écrivez-moi.

Thème 32 : densité

(1) Densité de $\mathbb{Q}$ dans $\mathbb{R}$, proposition 10.15.

(2) Densité des polynômes dans $\Big(C^0\big([0,1]\big), \|.\|_\infty\Big)$, théorème de Bernstein 36.145, ou une conséquence de Stone-Weierstrass 12.422.

(3) Densité des polynômes dans $\big(C^0(I), \|.\|_\infty\big)$ lorsque $I = [a,b]$, corolaire 36.146.

(4) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$ pour $1 \leqslant p < \infty$, théorème 27.50.

(5) Densité de $\mathscr{S}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 31.15.

(6) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 31.17.
Cela est utilisé pour le théorème de trace 31.19.

(7) Les applications étagées dans les applications mesurables (qui plus est avec limite croissante), théorème fondamental d'approximation 27.53.

(8) Les fonctions continues à support compact dans $L^2(I)$, théorème 27.54.

(9) Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$. Deux démonstrations indépendantes par le théorème 28.8 et le théorème 27.74.

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 17.128.

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

— Le lemme 12.418 le donne pour la racine carré.

— Le théorème 12.424 donne la densité des polynômes dans les fonctions continues sur un compact.

— Le théorème 12.421 est une généralisation qui donne la densité uniforme d'une sous-algèbre de $C(X, \mathbb{R})$ dès que X sépare les points.

— Le théorème 12.422 donne le même résultat pour la densité dans $C(X, \mathbb{C})$.

— Le lemme 28.1 est une version pour les polynômes trigonométriques.

— Le lemme 12.418 est un cas particulier du théorème 12.424, mais nous en donnons une démonstration indépendante afin d'isoler la preuve de la généralisation 12.422. Une version pour les polynômes trigonométriques sera donnée dans le lemme 28.1.

Le théorème de Stone-Weierstrass est utilisé, entre autres nombreuses choses, pour prouver la densité des polynômes trigonométriques dans les fonctions continues sur S^1, voir la proposition 27.91.

Thème 33 : application réciproque

(1) Définition 7.177.

(2) Dans le cas des réels, des exemples sont donnés en 10.8.

(3) Continuité, proposition 7.179.

(4) Théorème de la bijection 12.52 (qui contient aussi de la continuité).

(5) Dérivabilité, proposition 12.176.

Thème 34 : applications continues et bornées

(1) Une application linéaire non continue : exemple 11.62 de $e_k \mapsto ke_k$. Les dérivées partielles sont calculées en (25.140).

(2) La dérivation sur les polynômes (exemple 11.63) donne un autre exemple d'application linéaire non continue.

(3) Une application linéaire est bornée si et seulement si elle est continue, proposition 11.61.

(4) Une forme sesquilinéaire est bornée si et seulement si elle est continue, proposition 25.2.

Thème 35 : suite de Cauchy, espace complet Nous parlons d'espaces topologiques complets. À ne pas confondre avec un espace mesuré complet, définition 14.63.

(1) Corps complet : définition 1.318(5), espace métrique complet : définition 7.228.

(2) L'image d'une suite de Cauchy est de Cauchy. Si (x_n) est de Cauchy, alors $\big(f(x_n)\big)$ est de Cauchy quand f est une isométrie, lemme 17.132.

(3) La définition 7.227 donne la notion de suite de Cauchy dans un espace métrique.

(4) La définition 7.225 donne la notion de suite de τ-Cauchy dans un espace vectoriel topologique.

(5) Deux espaces métriques (avec une distance) peuvent être isomorphes en tant qu'espaces topologiques, mais ne pas avoir les mêmes suites de Cauchy, exemple 7.231.

(6) La proposition 7.233 donne l'équivalence entre les suites de Cauchy et les suites τ-Cauchy dans le cas des espaces vectoriels topologiques *normés*.

(7) L'exemple 7.231 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 $\mathbb{R}$.

(8) Le théorème 17.136 est un théorème de complétion d'un espace métrique.

(9) Dans $\mathbb{R}$, une suite est convergente si et seulement si elle est de Cauchy, théorème 7.247(2).

(10) Toute suite convergente dans un espace métrique est de Cauchy, proposition 7.234.

Quelques espaces qui sont complets sont listés ci-dessous. Attention : la complétude est bien une propriété de la métrique ; le même ensemble peut être complet pour une distance et pas pour une autre. Souvent, cependant la distance à considérer est donnée par le contexte.

(1) Les réels $\mathbb{R}$, théorème 7.247.

(2) Un espace vectoriel normé sur un corps complet est complet, proposition 7.253.

(3) La proposition 12.363 donne quelques espaces complets. Soit X un espace topologique métrique, (Y, d) un espace métrique complet. Alors les espaces
   (3a) $\big(C_b^0(X, Y), \|.\|_\infty\big)$
   (3b) $\big(C_0^0(X, Y), \|.\|_\infty\big)$
   (3c) $\big(C_0^k(X, Y), \|.\|_\infty\big)$
   sont complets.

(4) Le lemme 12.364 dit que $\big(C^0(A, B), \|.\|_\infty\big)$ est complet dès que A est compact et B est complet.

(5) L'espace $\mathscr{D}(K)$ est complet tant pour la topologie des seminormes que pour la topologie métrique (qui sont les mêmes). C'est la proposition 30.18.

(6) L'espace $\mathscr{S}(\Omega)$ est complet et métrisable par la proposition 30.68.

(7) L'espace $L^p(\Omega, \mathcal{A}, \mu)$ par le théorème 27.43.

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 12.424 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 $\big(C^p(X, Y), \|.\|_\infty\big)$ soit complet en général.

Thème 36 : caractérisations séquentielles Diverses caractérisations séquentielles.

continuité (1) Fonction séquentiellement continue, définition 7.173. Dans un espace métrisable séparé, la continuité séquentielle est équivalente à la continuité, proposition 7.220.

(2) La continuité implique la continuité séquentielle, proposition 7.175 et corolaire 7.117.

(3) Une version spéciale pour $\mathbb{R}^m$ est donnée par le théorème 12.219.

fermeture Fermeture séquentielle, proposition 7.217.

Compacité Un espace topologique séquentiellement compact : toute suite possède une sous-suite convergente, définition 7.75.

Thème 37 : valeurs propres, définie positive

À propos de valeurs propres (1) Définition des valeurs propres d'une forme quadratique : définition 9.234.

(2) Définition de valeur propre et vecteur propre pour un endomorphisme $f\colon V \to V$, définition 9.78.

À propos de choses définies positives (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 9.127.

(2) Un opérateur ou une matrice est défini positif si toutes ses valeurs propres sont positives, c'est la définition 9.215.

(3) Pour une matrice symétrique, définie positive si et seulement si $\langle Ax, x\rangle > 0$ pour tout x. C'est le lemme 9.219.

(4) Une application linéaire est définie positive si et seulement si sa matrice associée l'est. C'est la proposition 9.220.

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

Thème 38 : norme matricielle, norme opérateur et rayon spectral Quelques définitions

(1) Définition de la norme opérateur : définition 11.50.

(2) Définition du rayon spectral 11.56.

La norme matricielle n'est rien d'autre que la norme opérateur de l'application linéaire donnée par la matrice.

(1) Lien entre norme matricielle et rayon spectral, le théorème 12.116 assure que $\|A\|_2 = \sqrt{\rho(A^tA)}$.

(2) Lien entre valeurs propres et norme opérateur : le lemme 12.117 pour les matrices symétriques strictement définies positives donne $\|A\|_2 = \lambda_{max}$.

(3) Pour une matrice diagonale, $\|D\|_2 = \max\{|\lambda_i|\}$, lemme 12.118.

(4) Pour toute norme algébrique nous avons $\rho(A) \leqslant \|A\|$, proposition 12.110.

(5) Dans le cadre du conditionnement de matrice. Voir en particulier la proposition 34.110 qui utilise le théorème 12.116.

(6) Rayon spectral et convergence de méthode itérative, proposition 34.148.

Pour la norme opérateur nous avons les résultats suivants.

(1) La majoration $\|Au\| \leqslant \|A\|\|u\|$ est le lemme 11.58.

(2) Définition d'une algèbre : 1.291 et pour une norme d'algèbre : 11.55.

(3) La norme opérateur est une norme d'algèbre, lemme 11.60.

(4) Pour des espaces vectoriels normés, être borné est équivalent à être continu : proposition 11.61.

(5) Le lemme à propos d'exponentielle de matrice 15.145 donne :

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

La norme opérateur est utilisée pour donner une norme sur les produits tensoriels, définition 11.162.

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 39.

Thème 39 : série de matrices

(1) Rayon spectral et norme opérateur : thème 38.

(2) Exponentielle de matrices : thème 48.

(3) Série entière de matrices : section 15.13.

(4) Pour la série $\sum_k A^k = (1-A)^{-1}$.

— Pour un espace de Banach : proposition 11.201.

— Pour les matrices nilpotentes : proposition 9.199.

— En lien avec le rayon spectral (si et seulement si $\rho(A) < 1$) dans la proposition 15.141.

— Le lemme 15.45 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 34.168 pour prouver qu'une M-matrice irréductible vérifie $A^{-1} > 0$.

Thème 40 : décomposition de matrices

(1) Décomposition de Bruhat, théorème 13.39.

(2) Décomposition de Dunford, théorème 9.245.

(3) Décomposition polaire 13.32 des matrices symétriques et la proposition 17.60 pour la régularité.

Thème 41 : systèmes d'équations linéaires

— Algorithme des facteurs invariants 4.107.

— 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 choisie. Propositions 17.114 pour l'existence et 17.115 pour la méthode.

Thème 42 : réduction, diagonalisation Des résultats qui parlent diagonalisation

(1) Définition d'un endomorphisme diagonalisable : 9.200.

(2) Conditions équivalentes au fait d'être diagonalisable en termes de polynôme minimal, y compris la décomposition en espaces propres : théorème 9.204.

(3) Diagonalisation simultanée 9.207, pseudo-diagonalisation simultanée 11.36.

(4) Diagonalisation d'exponentielle 15.123 utilisant la décomposition de Dunford.

(5) Décomposition polaire théorème 13.32. $M = SQ$, S est symétrique, réelle, définie positive, Q est orthogonale.

(6) Décomposition de Dunford 9.245. $u = s + n$ où s est diagonalisable et n est nilpotent, $[s, n] = 0$.

(7) Réduction de Jordan (bloc-diagonale) 9.275.

(8) L'algorithme des facteurs invariants 4.107 donne $U = PDQ$ avec P et Q inversibles, D diagonale, sans hypothèse sur U. De plus les éléments de D forment une chaine d'éléments qui se divisent l'un l'autre.

Le théorème spectral et ses variantes :

(1) Théorème spectral, matrice symétrique, théorème 9.212. Via le lemme de Schur complexe 12.96.

(2) Théorème spectral autoadjoint (c'est le même, mais vu sans matrices), théorème 11.6

(3) Théorème spectral hermitien, lemme 11.18.

(4) Théorème spectral, matrice normales, théorème 12.98.

Pour les résultats de décomposition dont une partie est diagonale, voir le thème 40 sur les décompositions. Réduction de quadriques :

(1) Réduction de Gauss, théorème 9.226.

Trigonalisation.

(1) Le lemme de Schur complexe 12.96 dit que toute matrice est unitairement équivalente à une matrice triangulaire suppérieure.

Thème 43 : déterminant

(1) Déterminant d'une matrice : définition 4.74.

(2) Déterminant d'un endomorphisme 9.8.

(3) Le lemme 11.5 donne la formule $\det(f) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \langle e_{\sigma(i)}, f(e_i)\rangle$.

(4) Principales propriétés algébriques du déterminant : la proposition 9.9.

(5) La formule $\det(AB) = \det(A)\det(B)$ est la proposition 9.231 pour des matrices et la proposition 9.9(1) pour les applications linéaires.

(6) Déterminant et manipulations de lignes et colonnes, section 4.3.10 et les propositions qui précèdent à partir du lemme 4.76 qui dit que $\det(A) = \det(A^t)$.

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

(8) Déterminant d'une famille de vecteurs 9.4.

(9) Calcul d'un déterminant de taille 2×2 : équation (4.95).

(10) Interprétations géométriques

(10a) 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 11.33.

ii. En particulier le lemme 11.34 nous dit comment un déterminant donne un vecteur orthogonal à une famille donnée de vecteurs.

(10b) Déterminant et aires, volumes

i. Déterminant et mesure de Lebesgue : théorème 14.264.

ii. Aire du parallélogramme : il y a la formule avec le produit vectoriel dans la proposition 18.52, mais l'aire proprement dite, avec une intégrale est dans la proposition 20.31.

iii. Volume du parallélépipède avec le produit mixte et le déterminant 3×3, 18.53.

Tant que nous en sommes dans les interprétations géométriques, 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 (20.198) : $dS = \|T_u \times T_v\|$ qui est la norme du produit vectoriel des vecteurs tangents au paramétrage.

Nous voyons dans l'équation (20.195) 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 du paramétrage.

(11) Le déterminant de Vandermonde est à la proposition 9.11. Il est utilisé à divers endroits :

(11a) Pour prouver que u est nilpotente si et seulement si $\mathrm{Tr}(u^p) = 0$ pour tout p (lemme 9.196)

(11b) Pour prouver qu'un endomorphisme possédant $\dim(E)$ valeurs propres distinctes est cyclique (proposition 9.281).

Thème 44 : espaces de fonctions En ce qui concerne les densités, voir le thème 32.

(1) $C^\infty(A)$ est l'ensemble des fonctions de classe C^∞ sur A. Les éléments de $C^\infty(A)$ peuvent être à valeurs dans $\mathbb{R}$ ou dans $\mathbb{C}$ selon le contexte.

(2) $\mathscr{D}(A)$ est l'ensemble des fonctions de classe C^∞ dont le support est un compact dans A.

Pour les ensembles $\mathscr{S}(A)$, $\mathscr{D}(A)$ et $L^p(A)$, les fonctions sont à valeurs dans $\mathbb{C}$. La raison est que, de toutes façons, le passage à la transformée de Fourier produit en général des fonctions à valeurs dans $\mathbb{C}$ même si les fonctions de départ sont à valeurs dans $\mathbb{R}$.

Topologie Les espaces de fonctions sont souvent munis de topologies définies par des seminormes.

(1) La topologie des seminormes est la définition 7.285.

(2) La définition 30.13 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 30.21.

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) Un rappel de la construction en 27.76.

(2) Le produit scalaire [12] $\langle f,g\rangle$ est donné en (27.419) et la base trigonométrique est (27.420).

(3) La densité des polynômes trigonométriques dans $L^p(S^2)$ est le théorème 27.74 ou le théorème 28.8, au choix.

(4) Une conséquence de cette densité est que le système trigonométrique est une base hilbertienne [13] de L^2 par le lemme 27.117.

L'espace L^2 est discuté en analyse fonctionnelle, dans la section 27.5 et les suivantes parce que l'étude de L^2 utilise entre autres l'inégalité de Hölder 27.33.

Le fait que L^2 soit un espace de Hilbert est utilisé dans la preuve du théorème de représentation de Riesz 27.160.

Si $(\Omega,\mathcal{A},\mu)$ est un espace mesuré, alors $L^p(\Omega,\mathcal{A},\mu)$ est un espace de Banach ; c'est le théorème de Riesz-Fischer 27.44.

Thème 45 : fonctions Lipschitz

(1) Définition : 12.327.

(2) La notion de Lipschitz est utilisée pour définir la stabilité d'un problème, définition 34.26.

(3) Toute fonction Lipschitz est uniformément continue, proposition 12.332.

Thème 46 : suites et séries

Suites Les suites réelles sont en général dans la proposition 10.23 et ce qui s'ensuit. Cette proposition est souvent prise comme définition lorsque seules les suites réelles sont considérées.

(1) Les suites adjacentes, c'est la définition 10.35.

(2) Les séries alternées, théorème 11.124. Il s'agit de dire que $\sum_{k=0}^\infty(-1)^k a_k$ converge quand a_k est décroissante et tend vers zéro.

(3) Le concept de suite adjacente sert à étudier la série de Taylor de $\ln(x+1)$, voir le lemme 15.96 et ce qui l'entoure.

(4) La définition de la convergence absolue est la définition 11.80.

(5) Une suite réelle croissante et majorée converge, proposition 10.31.

(6) Toute suite dans un compact admet une sous-suite convergente, théorème 7.250.

(7) Pour tout réel, il existe une suite croissante de rationnels qui y converge, proposition 10.16.

Produit de Cauchy (1) Dans une algèbre normée, proposition 11.92,

(2) Dans $\mathbb{C}$, théorème 15.30.

Calcul de suites (1) Somme : $x_n+y_n\to x+y$ est la proposition 10.26.

12. Produit scalaire, définition 9.155.
13. Définition 25.27.

Série Les séries sont en général dans la section 11.7.

(1) Quelques séries usuelles en 11.9.3 : série harmonique, géométrique, de Riemann, et la mythique arithmético-géométrique.

(1a) La série est associative : $\sum_k (a_k + b_k) = \sum_k a_k + \sum_k b_k$. C'est la proposition 11.90.

(1b) La série harmonique diverge : $\sum_k \frac{1}{k} = \infty$, exemple 11.117.

(1c) La série géométrique : $\sum_{k=0}^{N} q^k = \frac{1-q^{N+1}}{1-q}$, proposition 11.119.

(1d) Une autre cool série : $\sum_{k=1}^{N} \frac{1}{k(k+1)} = \frac{N}{N+1}$, lemme 11.123.

(2) Critère des séries alternées, théorème 11.124.

(3) Convergence d'une série implique convergence vers zéro du terme général, proposition 11.88.

(4) Dans une algèbre normée : $(\sum_{k=0}^{\infty} a_k)b = \sum_{k=0}^{\infty}(a_k b)$, proposition 11.91.

(5) Produit de Cauchy : théorème 15.30 et proposition 11.92.

Sommes infinies En ce qui concerne les sommes finies, la notation $\sum_{i=1}^{N}$ est définie en 1.81. Pour permuter les termes d'une somme avec un élément du groupe symétrique, nous avons la proposition 1.255.

Voici quelques résultats à propos de sommes infinies.

(1) Une somme indexée par un ensemble quelconque est la définition 11.96.

(2) La proposition 11.100 donne une caractérisation pour le sommes de réels positifs.

(3) La définition de la somme d'une infinité de termes est donnée par la définition 11.77.

(4) Une somme de termes positifs indexée par un ensemble indénombrable est toujours infinie par le lemme 11.106.

(5) si la série converge, on peut regrouper ses termes sans modifier la convergence ni la somme (associativité) ; pour les sommes infinies l'associativité et la commutativité dans une série sont perdues. Néanmoins, il subsiste que

(5a) si la série converge absolument, on peut modifier l'ordre des termes sans modifier la convergence ni la somme (commutativité, proposition 11.94).

(6) Permuter une somme infinie avec une application linéaire : $f(\sum_{i\in I} v_i) = \sum_{i\in I} f(v_i)$, c'est la proposition 11.111.

Série entières (1) Rayon de convergence, définition 15.11.

(2) Convervence absolue à l'intérieur du rayon de convergence, lemme d'Abel 15.17.

(3) La fonction définie par la série entière $z \mapsto \sum_{k=0}^{\infty} a_n z^n$ est holomorphe dans son disque de convergence par la proposition 15.40.

(4) La série entière pour $\frac{1}{1-z^k}$, pour $\frac{1}{\omega - z}$ et pour $\frac{1}{(\omega-z)^k}$ sont dans le lemme 15.45.

Thème 47 : suite de fonctions

(1) Une limite uniforme de fonctions continues est continue, proposition 12.362.

(2) Sous certaines hypothèses, si $f_i \to f$, alors $f_i' \to f'$, théorème 12.379.

(3) Si les f_k sont continues et si la convergence $\sum_k f_k$ est uniforme, alors la somme est continue, théorème 12.375.

Thème 48 : exponentielle Toutes les exponentielles sont définies par la série

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

tant que la somme a un sens.

Réels Voici le plan que nous suivons dans le Frido :

— L'exponentielle est définie par sa série en 15.57.

— Nous démontrons qu'elle vérifie l'équation différentielle $y' = y$, $y(0) = 1$ (théorème 15.73).

— Nous démontrons l'unicité de la solution à cette équation différentielle.

— Nous démontrons qu'elle est égale à $x \mapsto y(1)^x$. Cela donne la définition du nombre e comme valant $y(1)$.

— Nous définissons le logarithme comme l'application réciproque de l'exponentielle (définition 15.78).

— Les fonctions trigonométriques (sinus et cosinus) sont définies par leurs séries. Il est alors montré que $e^{ix} = \cos(x) + i\sin(x)$ (lemme 18.11).

Propriétés — La formule $a^{-x} = 1/a^x$ est la proposition 12.402(3).

— $\exp(x) = e^x$, proposition 15.76.

— Le nombre e est irrationnel, proposition 15.77.

Complexes (1) La définition de $\exp(a + ib)$ est la définition 15.57.

(2) Les principales propriétés, dont $e^{z+w} = e^z e^w$, sont dans la proposition 18.9.

(3) Nous avons $e^{ix} = e^{iy}$ si et seulement si il existe $k \in \mathbb{Z}$ tel que $y = x + 2k\pi$, corolaire 18.23.

(4) Le fait que $e^{i\theta}$ donne tous les nombres complexes de norme 1 est la proposition 18.57.

(5) Le groupe des racines de l'unité est donné par l'équation (19.1).

Algèbre normée commutative Pour la définition c'est la proposition 15.57 et pour la régularité C^∞ c'est la proposition 15.62.

Idem non commutatif Il y a une tentative de théorème 15.63, mais c'est principalement pour les matrices qu'il y a des résultats.

Matrices De nombreux résultats sont disponibles pour les exponentielles de matrices.

(1) $e^{sA}e^{tA} = e^{(s+t)A}$, proposition 15.67.

(2) Si A est une matrice, alors $(e^tA)'(u) = Ae^{uA}$, proposition 15.69.

(3) Les sections 11.13 et 15.5.3 parlent d'exponentielle de matrices.

(4) L'exponentielle donne lieu à une fonction de classe C^∞, proposition 15.142.

(5) Le lemme à propos d'exponentielle de matrice 15.145 donne :

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

(6) La proposition 15.123 : 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.

(7) La section 15.13.3 parle des fonctions exponentielle et logarithme pour les matrices. Entre autres la dérivation et les séries.

(8) Pour résoudre des équations différentielles linéaires : sous-section 32.6.1.

(9) La proposition 15.122 dit que l'exponentielle est surjective sur $\mathrm{GL}(n, \mathbb{C})$.

(10) La proposition 11.204 : si u est un endomorphisme, alors $\exp(u)$ est un polynôme en u.

(11) Calcul effectif : sous-section 15.13.4.

(12) Proposition 13.23 : si $A \in \mathbb{M}(n, \mathbb{C})$ alors $e^{\mathrm{Tr}(A)} = \det(e^A)$.

(13) Les séries entières de matrices sont traitées autour de la proposition 15.139.

Thème 49 : logarithme

(1) Le logarithme pour les réels strictement positifs $\ln\colon]0,\infty[\to \mathbb{R}$ est donné en la définition 15.78 ; c'est l'application réciproque de exp.

(2) Les principales propriétés sont dans la proposition 15.81 : $\ln(xy) = \ln(x) + \ln(y)$ etc.

(3) Dérivée : $\ln'(x) = \frac{1}{x}$, proposition 15.80.

(4) La proposition 15.97 donne la série

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k. \tag{-2.2}$$

(5) L'exemple 20.150 donne l'encadrement $0.644 \leqslant \ln(2) \leqslant 0.846$.

(6) La proposition 15.122 dit que toute matrice complexe admet un logarithme. En particulier une série explicite est donnée pour le logarithme d'un bloc de Jordan [14].

(7) Sur les complexes, le logarithme $\ln\colon \mathbb{C}^* \to \mathbb{C}$ est la définition 26.66. Attention : ce n'est pas la seule définition possible.

(8) La série harmonique diverge à vitesse logarithmique, et la série des inverses des nombres premiers, c'est encore plus lent : théorème 15.111.

Thème 50 : fonction puissance Il y a beaucoup de choses à dire...

Définition Nous considérons, pour $a > 0$, la fonction $g_a\colon \mathbb{R} \to \mathbb{R}$ donnée par $g_a(x) = a^x$. La définition de cette fonction se fait en de nombreuses étapes.

(1) a^n pour $n \in \mathbb{N}$ en la définition 1.222.

(2) a^n pour $n \in \mathbb{Z}$ en la définition 12.383.

(3) $a^{1/n}$ pour $n \in \mathbb{Z}$ en la définition 12.386.

(4) a^q pour $q \in \mathbb{Q}$ en la définition 12.386.

(5) $\sqrt[n]{x}$ en la définition 12.388.

(6) La fonction g_a est Cauchy-continue sur $\mathbb{Q}$, c'est la proposition 12.399.

(7) a^x pour $a > 0$ et $x \in \mathbb{R}$ en la définition 12.401.

(8) a^z pour $a > 0$ et $z \in \mathbb{C}$ en la définition 18.5.

Quelques propriétés (1) Pour tout $q \in \mathbb{Q}$, il y a un $\sqrt{q}$ dans $\mathbb{R}$, proposition 1.404.

(2) Pour $a > 0$ et $x, y \in \mathbb{R}$ nous avons $a^x a^y = a^{x+y}$, proposition 12.402.

(3) Si $a > 0$ et $x, y \in \mathbb{R}$ nous avons $(a^x)^y = (a^y)^x = a^{xy}$ par la proposition 12.412.

(4) La fonction puissance est strictement croissante (en ses deux arguments), proposition 12.407.

(5) La formule $a^{-x} = 1/a^x$ est la proposition 12.402(3).

(6) La fonction puissance $g_a(x) = a^x$ est continue, proposition 12.402.

Croissance (1) La fonction puissance $f_\alpha(x) = x^\alpha$ est strictement croissante, proposition 12.407

(2) La fonction puissance $g_a(x) = a^x$ est strictement croissante, proposition 12.402.

(3) $\lim_{x\to\infty} x^\alpha = \infty$, proposition 12.413.

Continuité, Dérivation Comme toutes les choses sur la fonction puissance, les preuves sont assez différentes selon que l'on parle de a^x ou de x^a.

(1) La fonction $f_\alpha(x) = x^\alpha$ est continue, proposition 12.411.

(2) La fonction a^x est dérivable et sa dérivée vérifie $g_a'(x) = g_a(x)g_a'(0)$, proposition 12.428.

(3) La formule de dérivation pour $x \mapsto x^q$ avec $q \in \mathbb{Q}$ est la proposition 12.437.

14. Jordan, théorème 9.275.

(4) La dérivation de $x \mapsto x^\alpha$ avec $\alpha \in \mathbb{R}$ est la proposition 14.255. Si elle est tellement loin, c'est parce qu'elle nécessite de permuter une limite de fonctions avec une dérivée.

(5) Pour la formule générale de dérivation de $x \mapsto a^x$ demande de savoir les logarithmes (proposition 15.90).

L'équation fonctionnelle L'exponentielle et plus généralement la fonction puissance $g_a(x) = a^x$ peut être introduite au moyen d'une équation fonctionnelle au lieu de l'équation différentielle usuelle. Cette fameuse équation fonctionnelle est

$$f(x+y) = f(x)f(y) \tag{-2.3}$$

en la définition 12.430.

(1) L'équivalence entre l'équation fonctionnelle et l'équation différentielle est donnée par la proposition 12.435.

(2) La fonction $g_a(x) = a^x$ vérifie l'équation fonctionnelle $g_a(x+y) = g_a(x)g_a(y)$ et les conséquences. C'est la définition 12.430 et les choses qui suivent.

(3) L'équation fonctionnelle pour une fonction continue $f\colon \mathbb{R} \to S^1$ est traitée dans la proposition 12.439.

Une définition alternative de la fonction puissance serait de poser directement

$$a^x = e^{x\ln(a)}.$$

De là les propriétés se déduisent facilement. Dans cette approche, les choses se mettent dans l'ordre suivant :

— Définir $\exp(x)$ par sa série pour tout x.

— Démontrer que $\exp(q) = \exp(1)^q$ pour tout rationnel q (première partie de la proposition 15.76).

— Définir $e = \exp(1)$.

— Définir, pour x irrationnel, $a^x = \exp(x\ln(a))$.

— Prouver que $e^x = \exp(x)$ pour tout x.

Thème 51 : sommation finie et infinie La définition du symbole $\sum_{i\in I}$ se fait en trois étapes et deux demi, chacune se basant sur la précédente.

(1) Si $(A,+)$ est un ensemble avec une loi de composition interne, $\sum_{i=0}^n a_i$ est en 1.81.

(2) Si I est fini et si est à valeur dans un groupe commutatif, $\sum_{i\in I} f(i)$ est 1.254.

(3) Enfin si I est un ensemble quelconque, la définition 11.96 introduit la notion de famille sommable dans un espace vectoriel normé.

(4) Si (a_k) est une suite dans un espace vectoriel normé, la somme $\sum_{k=0}^\infty a_k$ est avec les sommes partielles dans la définition 11.77.

(5) La somme au sens de Cesàro est la somme des moyennes partielles, définition 11.125.

Notez que $\sum_{k=0}^\infty a_k$ n'est pas un cas particulier de $\sum_{k\in\mathbb{N}} a_k$. Une différence de taille entre les deux est que pour que $\sum_{k=0}^\infty a_k$ existe, il suffit que les a_k puissent être sommés dans cet ordre. À l'inverse pour que $\sum_{k\in\mathbb{N}} a_k$ existe, il faut que l'ordre de sommation puisse être arbitraire.

Si vous voulez sommer des séries encore moins convergentes, vous pouvez avoir envie d'utiliser la supersomme[2].

Thème 52 : polynôme de Taylor

Énoncés Il existe de nombreux énoncés du théorème de Taylor, et en particulier beaucoup de formules pour le reste.

(1) Énoncé : théorème 12.447.

(2) Une majoration du reste est dans le théorème 15.51

(3) De classe C^2 sur $\mathbb{R}^n$, proposition 12.454.

(4) Avec un reste donné par un point dans $]x, a[$, proposition 12.458.

(5) Avec reste intégral, proposition 20.151 et théorème 20.148 pour le cas simple $\mathbb{R} \to \mathbb{R}$.

(6) Le polynôme de Taylor généralise à l'utilisation de toutes les dérivées disponibles le résultat de développement limité donné par la proposition 12.168.

(7) Pour les fonctions holomorphes, il y a le théorème 26.31 qui donne une série de Taylor sur un disque de convergence.

Utilisation Des polynômes de Taylor sont utilisés pour démontrer des théorèmes par-ci par-là.

(1) Il est utilisé pour justifier la méthode de Newton autour de l'équation (34.95).

(2) On utilise pas mal de Taylor dans les résultats liant extrémum et différentielle/hessienne. Par exemple la proposition 17.73.

Quelques développements Voici quelques développements limités à savoir. Ils sont calculables en utilisant la formule de Taylor-Young (proposition 12.463).

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + x^n \alpha(x) \qquad \text{ordre } n, \text{proposition } 15.92$$

$$\cos(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k}}{(2k)!} + x^{2p+1} \alpha(x) \qquad \text{ordre } 2p+1, \text{proposition } 18.70$$

$$\sin(x) = \sum_{k=0}^{p} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + x^{2p+2} \alpha(x) \qquad \text{ordre } 2p+1, \text{proposition } 18.70$$

$$\ln(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + \alpha(x) x^n \qquad \text{ordre } n, \text{proposition } 15.95$$

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \qquad \text{exact proposition } 15.97$$

$$\ln(2) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \qquad \text{exact proposition } 15.97$$

$$(1+x)^l = \sum_{k=0}^{l} \binom{l}{k} x^k \qquad \text{exact si } l \text{ est entier.}$$

$$(1+x)^l = 1 + \sum_{k=1}^{n} \frac{l(l-1)\dots(l-k+1)}{k!} x^k + x^n \alpha(x) \qquad \text{ordre } n.$$

Dans toutes ces formules, la fonction $\alpha \colon \mathbb{R} \to \mathbb{R}$ vérifie $\lim_{t \to 0} \alpha(t) = 0$.

Le développement limité en 0 d'une fonction paire ne contient que les puissances de x d'exposant pair. Voir comme exemple le développement de la fonction cosinus.

Thème 53 : formule des accroissements finis Il en existe plusieurs formes :

(1) Une version adaptée aux espaces normés de dimension finie, avec hypothèse de différentiabilité, est le théorème 12.325. La formule $\|f(b) - f(a)\|_n \leqslant \sup_{x \in [a,b]} \|df_x\|_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} \|b - a\|_m$.

(2) Une version pour les dérivées partielles est dans le lemme 12.247. Pour rappel, la définition de la dérivation partielle est 12.228.

(3) La formule $f(a+\epsilon e_i) = f(a) + \epsilon(\partial_i f)(a) + \epsilon\alpha(\epsilon)$, proposition 12.248.

(4) L'existence de $c \in]a, b[$ tel que

$$f'(c) = \frac{f(b) - f(a)}{b - a} \tag{-2.5}$$

est le théorème des accroissements finis proprement dit. C'est le théorème 12.192.

(5) Il existe un c entre a et b tel que

$$f(b) = f(a) + (b - a)(\partial_\beta f)(c) \tag{-2.6}$$

où $\beta = b - a$ est la proposition 12.246.

(6) La formule $f(a+h) = f(a) + hf'(a) + \alpha(h)$ pour une fonction $\mathbb{R} \to \mathbb{R}$ en le théorème 12.168.

(7) Une généralisation pour les intervalles non bornés : théorème 12.193.

(8) Espaces vectoriels normés, théorème 11.194

Thème 54 : dérivation

(1) Définition de la dérivée, définition 12.160.

(2) Dérivée de fonction composée, proposition 12.180 dans le cas réel.

(3) Dérivée partielle de fonction composée, théorème 12.313.

(4) $f(\lambda x)' = \lambda f'(x)$, lemme 12.174.

Thème 55 : différentiabilité

Généralités (1) La différentielle est définie en général pour des espaces vectoriels normés par la proposition 11.169

(2) Différentielle d'une application linéaire, lemme 11.176.

(3) Nous parlons de différentielle en dimension finie et donnons une interprétation géométrique en 12.23.1.

(4) La recherche d'extrémums d'une fonction sur $\mathbb{R}^n$ passe par la seconde différentielle, proposition 17.73.

(5) Lien entre différentielle seconde (hessienne) et convexité en la proposition 17.97 et le corolaire 17.99.

(6) La différentielle est liée aux dérivées partielles par les formules données au lemme 12.265

$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot u. \tag{-2.7}$$

Je ne vous cache pas que cette suite d'égalités est une de mes préférées.

Différentielle et dérivées partielles À propos de fonctions de classe C^k, définition 11.170.

(1) Une fonction est de classe C^1 si et seulement si ses dérivées partielles sont continues, théorème 12.306.

(2) Une fonction est C^n si et seulement si ses dérivées partielles sont C^{n-1}, c'est le théorème 12.340.

(3) Différentiabilité en un seul point si les dérivées partielles sont continues en ce point : proposition 12.304.

Fonctions composées À propos de la formule $d(f \circ g)_a = dg_{f(a)} \circ df_a$, il y a deux théorèmes très semblables.

(1) Le théorème 11.183 insiste sur des hypothèses locales.

(2) Le théorème 11.184 fait des hypothèses plus globales pour s'alléger l'esprit, mais fait une récurrence pour dire que $f \circ g$ est de classe C^r si f et g le sont.

Thème 56 : 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 20.38 pour les équations différentielles.
(2) Le théorème de Schauder 20.37 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 17.42) sans la condition Lipschitz
(3) Le théorème de Cauchy-Lipschitz 17.42 est utilisé à plusieurs endroits :
   — Pour calculer la transformée de Fourier de $e^{-x^2/2}$ dans le lemme 29.22.
(4) Théorème de stabilité de Lyapunov 32.38.
(5) Le système proie-prédateur de Lotka-Volterra 32.39
(6) Équation de Schrödinger, théorème 32.46.
(7) L'équation $(x - x_0)^\alpha u = 0$ pour $u \in \mathscr{D}'(\mathbb{R})$, théorème 30.40.
(8) La proposition 32.41 donne un résultat sur $y'' + qy = 0$ à partir d'une hypothèse de croissance.
(9) Équation de Hill $y'' + qy = 0$, proposition 32.43.

Thème 57 : convexité

Fonctions convexes L'essentiel des résultats sur les fonctions convexes sont dans la section 17.11. On a surtout :

(1) Définition des fonctions convexes : 17.79 et 17.95 en dimension supérieure.
(2) En termes de différentielles, 17.96 pour la différentielle première et 17.99 pour la hessienne.
(3) Une courbe paramétrée convexe est la définition 21.91.
(4) L'enveloppe convexe d'une courbe fermée simple et convexe : 21.93.
(5) Courbure et convexité d'une courbe paramétrée : section 21.13.4.
(6) Une courbe paramétrée convexe est localement le graphe d'une fonction convexe par le lemme 21.92.
(7) La convexité est utilisée dans la méthode du gradient à pas optimal de la proposition 17.115.
(8) La fonction $t \mapsto t^p$ est strictement convexe sur les positifs dans le lemme 17.88.

En termes de parties convexes, on a :

Parties convexes (1) Définition 7.134 d'une partie convexe d'un espace vectoriel.

(2) Une boule est convexe, proposition 8.29.
(3) Un polygone convexe est défini en 18.161, et les racines de l'unité forment un polygone convexe par la proposition 18.162.

Thème 58 : espaces de Hilbert, base hilbertienne Beacoup de choses concerant les espaces L^2, et en particulier la base trigonométrique sont dans le thème 44.

(1) Toute partie orthonormale d'un espace de Hilbert est libre, proposition 25.24.
(2) Tout espace de Hilbert possède une base hilbertienne, proposition 25.37.
(3) Unicité de la décomposition dans une base hilbertienne, proposition 25.38.

Thème 59 : analyse complexe, fonctions holomorphes

(1) Le lien entre différentielle et dérivée complexe est donné par les équations

$$df_{z_0}(z) = f'(z_0)z = (\partial_z f)(z_0)z \tag{-2.8}$$

par (12.863) et la proposition 26.4. Cela se résume par la formulation lapidaire $f' = \partial_z f$.
(2) Série de Laurent $f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$, théorème 28.26.

Thème 60 : permuter des limites

Permuter des dérivées partielles Si une fonction est de classe C^2, le théorème de Schwarz 12.349 dit que

$$\partial_k \partial_l f = \partial_l \partial_k f. \tag{-2.9}$$

Fonctions définies par une intégrale Les théorèmes sur les fonctions définies par une intégrale, section 17.4. Nous avons entre autres

(1) $\partial_i \int_B f = \int_B \partial_i f$, avec B compact, proposition 17.27.

(2) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 17.15 pour la continuité de $x \mapsto \int_\Omega f(x,\omega)d\mu(\omega)$.

(3) Pour la fonction $F(x) = \int_\Omega f(x,\omega)d\mu(\omega)$, nous avons la dérivation sous l'intégrale par la formule de Leibnitz

$$F'(a) = \int_\Omega \frac{\partial f}{\partial x}(a,\omega)d\mu(\omega) \tag{-2.10}$$

démontrée en le théorème 17.19.

Des variations avec des dérivées partielles et des différentielles sont dans 17.27 et dans 17.28.

(4) Si $f\colon \mathbb{C} \times \Omega \to \mathbb{C}$ est holomorphe (pour $\mathbb{C}$), alors F est holomorphe et

$$F'(z) = \int_\Omega \frac{\partial f}{\partial z}(z,\omega)d\mu(\omega). \tag{-2.11}$$

(5) Pour des dérivées partielles multiples, nous avons la formule

$$(\partial^\alpha F)(x) = \int_\Omega (\partial^\alpha f_\omega)(a)d\mu(\omega) \tag{-2.12}$$

dans la proposition 17.21.

(6) Si l'intégrale est uniformément convergente, nous avons le théorème 17.16 qui donne la continuité de $F(x) = \int_\Omega f(x,\omega)d\mu(\omega)$.

(7) 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 26.19.

(8) En ce qui concerne le x dans la borne, le théorème 14.246 lie primitive et intégrale en montrant que $F(x) = \int_a^x f(t)dt$ est une primitive de f (sous certaines conditions). Le théorème fondamental de l'analyse 14.247 en est une conséquence.

(9) Si T est une **distribution**, alors nous avons

$$T\big(x \mapsto (\partial_y^\alpha \phi)(x,y_0)\big) = \partial_y^\alpha \Big(T\big(x \mapsto \phi(x,y)\big)\Big)_{y=y_0}. \tag{-2.13}$$

C'est la proposition 30.44.

Limite et intégrale (1) Théorème de la convergence monotone, théorème 14.166.

Fubini 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 [15]. Nous utilisons cette technique pour permuter une somme et une intégrale dans l'équation (26.134).

L'utilisation de Fubini pour permuter des intégrales (sur deux variables différentes) ou deux sommes est expliquée dans 14.273.

C'est par exemple utilisé pour permuter deux sommes dans le cadre des chaines de Markov en 38.8.

— le théorème de Fubini-Tonelli 14.269 demande que la fonction soit mesurable et positive ;

15. Mesure de comptage, définition 14.240.

— le théorème de Fubini 14.272 demande que la fonction soit intégrable (mais pas spécialement positive) ;
— le corolaire 14.271 demande l'intégrabilité de la valeur absolue des intégrales partielles pour déduire que la fonction elle-même est intégrable.

Limite et dérivées, différentielle (1) Permuter limite et dérivée dans le cas $\mathbb{R} \to \mathbb{R}$, théorème 12.379.
(2) Permuter limite et dérivées partielles, théorème 12.382.
(3) Permuter limite et différentielle, théorème 15.8.

Quelques remarques sur les techniques de démonstration.
(1) Le résultat fondamental 12.379 est démontré sans recourir à des intégrales. Une preuve alternative, plus courte, avec des intégrales est donnée en 14.254.
(2) Permuter limite et dérivée partielle, théorème 12.382.
(3) Permuter série et différentielle, théorème 15.8.

Somme et dérivée Permuter somme et différentielle, théorème 15.8.

Limite et mesure Une mesure n'est pas toujours une limite, mais la définition d'une mesure positive sur un espace mesurable parle de permuter limite et mesure : définition 14.16(3).

Thème 61 : 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 14.185 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 27.63.
(3) Et encore un pour L^p dans la proposition 27.165.
(4) Si $\int f\chi = 0$ pour tout χ à support compact alors $f = 0$ presque partout, proposition 30.1.
(5) En utilisant le théorème de représentation de Riesz, on peut prouver que $\int_\Omega f\chi = 0$ imploque $f = 0$ pour tout $f \in L^p$, proposition 27.165.
(6) La proposition 27.21 donne $f = 0$ dans L^p lorsque $\int fg = 0$ pour tout $g \in L^q$.
(7) 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 17.2.

Thème 62 : inversion locale, fonction implicite
(1) Théorème d'inversion locale dans un Banach : théorème 17.50.
(2) Fonction implicite dans un Banach : théorème 17.51.
(3) Utilisé pour montrer que le flot d'une équation différentielle est un C^p-difféomorphisme local, voir 32.34.
(4) Pour le théorème de Von Neumann 17.64.

Thème 63 : points fixes
(1) Il y a plusieurs théorèmes de points fixes.

Théorème de Picard 17.36 donne un point fixe comme limite d'itérés d'une fonction Lipschitz. Il aura pour conséquence le théorème de Cauchy-Lipschitz 17.42, l'équation de Fredholm, théorème 17.41 et le théorème d'inversion locale dans le cas des espaces de Banach 17.50.

Théorème de Brouwer qui donne un point fixe pour une application d'une boule vers elle-même. Nous allons donner plusieurs versions et preuves.
(1a) Dans $\mathbb{R}^n$ en version C^∞ via le théorème de Stokes, proposition 20.35.
(1b) Dans $\mathbb{R}^n$ en version continue, en s'appuyant sur le cas C^∞ et en faisant un passage à la limite, théorème 20.36.
(1c) Dans $\mathbb{R}^2$ via l'homotopie, théorème 26.21. 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 [16].

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

Théorème de Markov-Kakutani 20.40 qui donne un point fixe à une application continue d'un convexe fermé borné dans lui-même.

Théorème de Schauder C'est une version valable en dimension infinie du théorème de Brouwer. Théorème 20.37

(2) Pour les équations différentielles

(2a) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 20.38 pour les équations différentielles.

(2b) Le théorème de Schauder 20.37 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 17.42) 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 34.48.

— La convergence quadratique de la méthode de Newton est donnée par le théorème 34.54.

— En calcul numérique, section 34.5

— Méthode de Newton comme méthode de point fixe, sous-section 34.6.2.

(4) D'autres utilisations de points fixes.

— Processus de Galton-Watson, théorème 38.53.

— Dans le théorème de Max-Milgram 25.61, le théorème de Picard est utilisé.

Thème 64 : changement de variables Il n'existe rien en mathémaitque qui s'appelle « changement de variables ». Il n'existe que des compositions de fonctions. Ce snobisme terminologique étant, voici un certain nombre de résultats de changement de variables.

(1) Dans des intégrales, théorème 14.265.

(2) Dans des limites, le lemme 7.153 donne $\lim_{x\to a} f(x) = \lim_{x\to b} f(x+a-b)$ si la limite existe.

(3) Dans une équation aux dérivées partielles, exemple 33.132.

(4) Limite de fonction composée $\lim_{x\to a}(f\circ g)$, propositions 7.154 et 7.155.

Thème 65 : techniques de calcul Il y en a pour tous les gouts.

Primitives et intégrales Toute la section 17.12 donne des trucs et astuces pour trouver des primitives et des intégrales.

Limite à deux variables Les exemples de limites à plusieurs variables font souvent intervenir des coordonnées polaires (du théorème 18.220) ou autres fonctions trigonométriques. Ils sont donc placés beaucoup plus bas que la théorie.

— Méthode du développement asymptotique, sous-section 18.13.2.

— Méthode des coordonnées polaires, sous-section 18.13.1.

— Utilisation du théorème de la fonction implicite, dans l'exemple 18.222.

Thème 66 : méthodes de calcul

(1) Théorème de Rothstein-Trager 20.98.

(2) Algorithme des facteurs invariants 4.107.

(3) Méthode de Newton, théorème 34.63

(4) Calcul d'intégrale par suite équirépartie 28.10.

Thème 67 : méthode de Newton

(1) Nous parlons un petit peu de méthode de Newton en dimension 1 dans 34.6.
(2) La méthode de Newton fonctionne bien avec les fonctions convexes par la proposition 34.56.
(3) La méthode de Newton en dimension n est le théorème 34.63.
(4) Un intervalle de convergence autour de α s'obtient par majoration de $|g'|$, proposition 34.48.
(5) Un intervalle de convergence quadratique s'obtient par majoration de $|g''|$, théorème 34.54.
(6) En calcul numérique, section 34.6.
(7) Méthode de Newton pour calculer $\sqrt{A}$, exemple 34.57.

Thème 68 : opérations sur les distributions

(1) Convolution d'une distribution par une fonction, définition par l'équation (30.186).
(2) Dérivation d'une distribution, proposition-définition 30.26.
(3) Produit d'une distribution par une fonction, définition 30.25.

Thème 69 : convolution

(1) Définition 27.55, et principales propriétés sur $L^1(\mathbb{R})$ dans le théorème 27.56.
(2) Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 27.60.
(3) $\varphi \in L^1(\mathbb{R})$ et $\psi \in \mathscr{S}(\mathbb{R})$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R})$, proposition 27.198.
(4) Les suites régularisantes : $\lim_{n\to\infty} \rho_n * f = f$ dans la proposition 29.18.
(5) Convolution d'une distribution par une fonction, définition par l'équation (30.186).

Thème 70 : séries de Fourier

— Le système trigonométrique est donné en la définition 27.69.
— Les coefficients de Fourier de $c_n(f)$ sont donnés par 27.71.
— Formule sommatoire de Poisson, proposition 29.15.
— Inégalité isopérimétrique, théorème 28.23.
— Fonction continue et périodique dont la série de Fourier ne converge pas, proposition 28.21.
— 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 Cesàro, théorème de Fejèr 28.7.
  (2) Convergence au sens $L^2\Big([0, 2\pi]\Big)$ dans le théorème 27.118.
  (3) Si f est continue, périodique et si sa série de Fourier converge uniformément, théorème 28.15.
  (4) Si f est périodique et la série des coefficients converge absolument pour tout x, proposition 28.16.
  (5) Si f est périodique et de classe C^1, théorème 28.17.
  (6) Unicité des coefficients de Fourier, corollaire 28.18.

  Il est cependant faux de croire que la continuité et la périodicité suffisent à obtenir une convergence, comme le montre la proposition 28.21.

Thème 71 : transformée de Fourier

(1) Définition sur L^1, définition 29.1.
(2) La transformée de Fourier d'une fonction $L^1(\mathbb{R}^d)$ est continue, proposition 29.11.
(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 continue. Proposition 29.20
(4) L'application $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ est une bijection. Formule d'inversion, proposition 29.26.

Thème 72 : gaussienne

(1) Le calcul de l'intégrale
$$\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}$$
est fait de deux façons dans l'exemple 14.275. Dans les deux cas, le théorème de Fubini 14.272 est utilisé.

(2) Le lemme 29.22 calcule la transformée de Fourier de $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ qui donne $\hat{g}_\epsilon(\xi) = \left(\frac{\pi}{\epsilon}\right)^{d/2} e^{-\|\xi\|^2/4\epsilon}$.

(3) Le lemme 29.23 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 29.26

Thème 73 : lemme de transfert Il y a deux résultats qui portent ce nom. Le premier est dans la théorie de Fourier, le résultat $\hat{f}' = i\xi\hat{f}$.

(1) Lemme 29.19 sur $\mathscr{S}(\mathbb{R}^d)$

(2) Lemme 31.10 pour L^2.

L'autre lemme de transfert est en théorie des tribus, le résultat $\sigma\big(f^{-1}(\mathcal{C})\big) = f^{-1}\big(\sigma(\mathcal{C})\big)$ du lemme 14.43. Celui-ci est d'ailleurs plutôt nommé « lemme de transport ».

Thème 74 : invariants de similitude

(1) Théorème 9.272.

(2) Pour prouver que la similitude d'applications linéaires résiste à l'extension du corps de base, théorème 9.286.

(3) Pour prouver que la dimension du commutant d'un endomorphisme de E est de dimension au moins $\dim(E)$, lemme 9.283.

(4) Nous verrons dans la remarque 9.273 à propos des invariants de similitude que toute matrice est semblable[17] à la matrice bloc-diagonale constituées des matrices compagnon (définition 9.267) de la suite des polynômes minimaux.

Thème 75 : isométries Il y a $(\mathbb{R}^n, \|.\|)$ et $\mathbb{R}^n, d$.

Les isométries de $\|.\|$ sont linéaires, tandis que les isométries de la distance contiennent aussi les translations et les rotations de centre différent de l'origine.

Ne pas confondre une isométrie d'un espace affine avec une isométrie d'un espace euclidien[18]. Les isométries d'un espace euclidien préservent le produit scalaire et fixent donc l'origine (lemme 11.15). Les isométries des espaces affines par contre conservent les distances (définition 8.64) 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).

Général Quelques résultats généraux et en vrac à propos d'isométries.

(1) Définition d'une isométrie pour une forme bilinéaire, 9.139. Pour une forme quadratique : définition 9.138.

(2) Définition du groupe orthogonal 9.36, et le spécial orthogonal $\mathrm{SO}(n)$ en la définition 9.39. Le groupe $\mathrm{SO}(2)$ est le groupe des rotations, par corolaire 18.131.

(3) La rotation $R_A(\theta)$ d'un angle θ autour du point $A \in \mathbb{R}^2$ est donnée par la définition 18.120.

17. Définition 4.104.
18. Définition 9.159.

(4) La proposition 18.132 donne à toute rotation $R_0(\theta)$ une matrice de la forme connue. C'est autour de cela que nous définissons les angles, définition 18.150.

(5) Le groupe orthogonal est le groupe des isométries de $\mathbb{R}^n$, proposition 9.38.

(6) Les isométries de l'espace euclidien sont affines, 9.141.

(7) Les isométries de l'espace euclidien comme produit semi-direct : $\mathrm{Isom}(\mathbb{R}^n) \simeq T(n) \times_\rho \mathrm{O}(n)$, théorème 18.74.

(8) Isométries du cube, section 5.7.

(9) Nous parlons des isométries affines du tétraèdre régulier dans la proposition 18.188.

Groupe diédral Le groupe diédral est un peu central dans la théorie des isométries de $(\mathbb{R}^2, d)$ parce que beaucoup de sous-groupes finis des isométries de $(\mathbb{R}^2, d)$ sont en fait isomorphes au groupe diédral.

(1) Générateurs du groupe diédral, proposition 18.168.

(2) Un sous-groupe fini des isométries de $(\mathbb{R}^2, d)$ contenant au moins une réflexion est isomorphe au groupe diédral par le théorème 18.193.

(3) Le théorème 18.195 dit que le groupe des isométries propres d'une partie quelconque de $(\mathbb{R}^2, d)$ est soit cyclique soit isomorphe au groupe diédral.

Isométries et réflexions Dans un espace euclidien, toute isométrie peut être décomposée en réflexions autour d'hyperplans. Voici quelques énoncés à ce propos.

(1) Définition d'une réflexion dans $\mathbb{R}^2$ 18.114.

(2) La caractérisation en termes de projection orthogonale est le lemme 18.100 ; en termes de médiatrice c'est le lemme 18.106.

(3) Définition d'un hyperplan 9.287.

(4) En dimension 2, une rotation est définie comme composée de deux réflexions en la définition 18.111.

(5) En dimension 2, les réflexions ont un déterminant -1 par le lemme 18.124.

(6) Les isométries du plan $(\mathbb{R}^2, d)$ sont données dans le théorème 18.184, et sont au plus 3 réflexions par le théorème 18.182.

(7) Décomposition des isométries de $\mathbb{R}^n$ en réflexions par le lemme 18.81.

(8) En particulier, les éléments de $\mathrm{SO}(3)$ sont des compositions de deux réflexions par le corolaire 18.83.

(9) Une isométrie de $\mathbb{R}^n$ préserve l'orientation si et seulement si elle est la composition d'un nombre pair de réflexions. C'est le théorème 18.89.

Sous-groupe fini (1) Les sous-groupes finis des isométries de $(\mathbb{R}^2, d)$ sont cycliques, théorème 18.193.

(2) Les sous-groupes finis de $\mathrm{SO}(3)$ sont listés dans 18.217.

(3) Les sous-groupes finis de $\mathrm{SO}(2)$ sont cycliques, lemme 18.135.

Thème 76 : enveloppes

(1) L'ellipse de John-Loewner donne un ellipsoïde de volume minimum autour d'un compact dans $\mathbb{R}^n$, théorème 17.123.

(2) Le cercle circonscrit à une courbe donne un cercle de rayon minimal contenant une courbe fermée simple, proposition 21.90.

(3) Enveloppe convexe du groupe orthogonal 13.38.

(4) Enveloppe convexe d'une courbe fermée plane comme intersection des demi-plans tangents, proposition 21.96.

Thème 77 : intégration sur des variétés

orientation La notion d'orientation commence avec l'orientation des bases d'un espace vectoriel et continue jusqu'à orienter des variétés à partir de ses cartes.

(1) Classe d'orientation sur les bases d'un espace vectoriel, définition 9.21.

(2) Orientation sur une surface, définition 20.63.

(3) Variété orientable, définition 20.16.

théorème de Stokes, théorème de Green et compagnie Tous ces théorèmes sont des conséquences plus ou moins directes de celui de Stokes, et des généralisations du théorème fondamental de l'analyse.

(1) Forme générale, théorème 20.74.

(2) Rotationnel et circulation, théorème 24.8.

Le théorème de Stokes peut être utilisé pour montrer le théorème de Brouwer, proposition 20.35.

Thème 78 : dénombrements

— Coloriage de roulette (18.9.15.1) et composition de colliers (18.9.15.2).

— Nombres de Bell, théorème 15.160.

— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décompositions de fractions en éléments simples [19], théorème 26.53.

Thème 79 : caractérisation de distributions en probabilités

(1) La probabilité conjointe est la définition 36.19.

(2) La fonction de répartition est la définition 36.61.

(3) La fonction caractéristique est la définition 36.63.

Thème 80 : théorème central limite

(1) Pour les processus de Poisson, théorème 40.5.

Thème 81 : probabilités et espérances conditionnelles Les deux définitions de base, sur lesquelles se basent toutes les choses conditionnelles sont :

— L'espérance conditionnelle d'une variable aléatoire sachant une tribu : $E(X|\mathcal{F})$ de la définition 36.37.

Les autres sont listées ci-dessous.

Probabilité conditionnelle .

Plusieurs probabilités conditionnelles.

— D'un événement en sachant un autre : la définition 36.33 donne

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

Cela est la définition de base. L'autre est une définition dérivée.

— D'un événement vis-à-vis d'une variable aléatoire discrète. C'est par la définition 36.55 qui définit la variable aléatoire

$$P(A|X)(\omega) = P(A|X = X(\omega)).$$

Dans le cas continu, c'est la définition 36.56 :

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

19. Éléments simples, lemme 19.18.

— D'un événement par rapport à une tribu. C'est la variable aléatoire

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

Espérances conditionnelles Plusieurs espérances conditionnelles.

— D'une variable aléatoire par rapport à un événement, définition 36.49 :

$$E(X|A) = \frac{E(X\mathbb{1}_A)}{P(A)}. \tag{-2.14}$$

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

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

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 36.37). Cela au sens des projections orthogonales.

— d'une variable aléatoire par rapport à une autre. La définition 36.39 est une variation sur le même thème :

$$E(X|Y) = E(X|\sigma(Y)),$$

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

Quelque formules.

(1) Pour l'espérance conditionnelle d'une variable aléatoire prenant seulement une quantité dénombrable de valeurs : $E(X|A) = \sum_{k=0}^{\infty} y_k P(X = y_k|A)$ par le lemme 36.50.

(2) La probabilité conditionnelle se factorise par rapport à l'union disjointe par le lemme 36.36 : $P\big(\bigcup_{i=0}^{\infty} A_i|B\big) = \sum_{i=0}^{\infty} P(A_i|B)$.

-2.1 Conventions sur les matrices et changement de bases

-2.1.1 Matrices et applications linéaires

Le lien entre matrice et application linéaire est donné par la définition 4.65. L'application d'une matrice à un vecteur est (4.78). Le lien le plus simple entre l'application linéaire et les éléments de matrice est donné par la proposition 4.68. Voici les relations :

$$T_{\alpha i} = T(e_i)_\alpha \tag{-2.15a}$$

$$T(e_i) = \sum_\alpha T_{\alpha i} f_\alpha \tag{-2.15b}$$

$$T(x) = \sum_{i\alpha} T_{\alpha i} x_i f_\alpha \tag{-2.15c}$$

$$T(x)_\alpha = \sum_i T_{\alpha i} x_i. \tag{-2.15d}$$

Cela définit une application $\psi\colon \mathbb{M}(n \times m, \mathbb{K}) \to \mathcal{L}(E, F)$ qui a plein de propriétés.

(1) C'est une bijection, proposition 4.68(4).

(2) C'est un isomorphisme d'algèbre, proposition 4.71.

(3) C'est un isomorphisme d'espaces vectoriels, proposition 4.69.

(4) Isomorphisme d'algèbres et d'anneaux, proposition 4.71.

(5) Isomorphisme d'espaces topologiques, proposition 4.71.

Lorsque nous avons une base orthonormée[20] nous avons aussi les propositions 9.171 et 9.171 qui donnent des formules avec produit scalaire :

(1) $T_{\alpha i} = e_\alpha \cdot T(e_i)$
(2) $x \cdot Ay = \sum_{kl} A_{kl} x_k y_l$.

où le point est le produit scalaire usuel de $\mathbb{R}^n$.

-2.1.2 Le changement de base

Soit un espace vectoriel V muni de deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$. Le lemme 4.108 donne le lien entre les vecteurs de base :

(1) $f_\alpha = \sum_i Q_{i\alpha} e_i$
(2) $e_i = \sum_\alpha Q^{-1}_{\alpha i} f_\alpha$

La proposition 4.109 donne un certain nombre de formules pour les coordonnées des vecteurs :

(1) $y_\alpha = \sum_i Q^{-1}_{\alpha i} x_i$
(2) $x_i = \sum_\alpha Q_{i\alpha} y_\alpha$.
(3) $x_i = (Qy)_i$
(4) $x = Qy$

La transformation de la matrice d'une application linéaire lors d'un changement de base est la proposition 4.112. Soit une application linéaire $T\colon V \to V$ de matrices A et B dans les bases $\{e_i\}$ et $\{f_\alpha\}$. Si les bases sont liées par $f_\alpha = \sum_i Q_{i\alpha} e_i$, alors les matrices A et B sont liées par

$$B = Q^{-1} A Q. \tag{-2.16}$$

-2.1.3 Changement de base : matrice d'une forme bilinéaire

La proposition 9.133 fait le changement de matrice d'une forme bilinéaire lors d'un changement de base. Si la matrice de q dans la base $\{e_i\}$ est A et celle dans la base $\{f_\alpha\}$ est B, alors

$$B = Q^t A Q. \tag{-2.17}$$

Pour comparaison avec la loi de transformation des matrices des applications linéaires, voir la remarque 9.134.

-2.2 Multiindice et liste d'indices

-2.1.
Je crois qu'il y a quelques incohérences de notations/dénominations dans le texte. En principe quand on parle de $\mathbb{R}^n$, un **multiindice** [3] est une vecteur d'entiers positifs à n composantes. Si $\alpha = (2, 1)$ alors nous avons la notation

$$\partial^\alpha f = \frac{\partial^2}{\partial x_1} \frac{\partial}{\partial x_2} f \tag{-2.18}$$

Cette notation pose problème lorsque, par exemple, $\partial_1^2 \partial_2 f \neq \partial_1 \partial_2 \partial_1 f$.

Elle pose également problème lorsque l'on veut faire une récurrence sur l'ordre de dérivation en ajoutant une seule dérivation à la fois.

C'est pourquoi nous introduisons le concept de **liste d'indices**. En parlant de $\mathbb{R}^n$, une liste d'indices est un vecteur arbitrairement long (mais fini) d'entiers dans $\{1, \dots, n\}$. Si, dans $\mathbb{R}^7$, $\alpha = (1, 3, 1, 5)$, alors

$$\partial^\alpha f = \partial_1 \partial_3 \partial_1 \partial_5 f. \tag{-2.19}$$

Si α est une liste d'indices de longueur p, une **queue de** α est une liste d'indices de longueur $0 < k \leqslant p$ de la forme $(\alpha_{p-k+1}, \alpha_{p-k+2}, \dots, \alpha_p)$.

20. Définition 9.157.

-2.3 Anglicismes

Voici quelque anglicismes dont je ne me souviens jamais.

(1) Une σ-algèbre est une tribu, définition 14.1.

(2) Le « uniform boundedness principle » est le théorème de Banach-Steinhaus 11.135.

(3) 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.

Table des matières

Thématique **1**
-2.1 Conventions sur les matrices et changement de bases . . . 34
-2.1.1 Matrices et applications linéaires . . . 34
-2.1.2 Le changement de base . . . 35
-2.1.3 Changement de base : matrice d'une forme bilinéaire . . . 35
-2.2 Multiindice et liste d'indices . . . 35
-2.3 Anglicismes . . . 36

Table des matières **37**

Index **69**

Liste des notations **93**

0 Introduction (Vol 1) **97**
0.1 Auteurs, contributeurs, sources et remerciements . . . 97
0.1.1 Ceux qui ont travaillé sur le Frido . . . 97
0.1.2 Aide directe, mais pas volontairement sur le Frido . . . 98
0.1.3 Des gens qui ont fait un travail qui m'a bien servi . . . 99
0.2 Originalité . . . 99
0.3 Les choses qui doivent vous faire tiquer . . . 100
0.4 Quelques choix qui peuvent provoquer des quiproquos . . . 100
0.5 Autres choix pas spécialement standards . . . 101
0.5.1 Mathématique intéressante . . . 101
0.6 Sage est là pour vous aider . . . 101
0.6.1 Lancez-vous dans Sage . . . 102
0.6.2 Exemples de ce que Sage peut faire pour vous . . . 102
0.7 Comment contribuer et aider ? . . . 103
0.7.1 Des preuves qui manquent . . . 103
0.7.2 Du texte qui manque . . . 103
0.7.3 Des exemples qui manquent . . . 103
0.7.4 Trucs de programmation et de LaTeX . . . 103
0.8 Les politiques éditoriales . . . 104
0.8.1 Licence libre . . . 104
0.8.2 pdflatex . . . 104
0.8.3 utf8 . . . 104
0.8.4 Notations . . . 104
0.8.5 De la bibliographie . . . 104
0.8.6 Faire des références à tout . . . 104
0.8.7 Des listes de liens internes . . . 104
0.8.8 Pas de références vers le futur . . . 104

1 Construction des ensembles de nombres (Vol 1) **105**
1.1 Quelques éléments sur les ensembles . . . 105
1.1.1 Petit mot d'introduction . . . 105

1.1.2 Injection, surjection, bijection . . . 106
1.1.3 Ensemble ordonné . . . 106
1.1.4 Lemme de Zorn . . . 108
1.1.5 Complémentaire . . . 109
1.1.6 Relations d'équivalence . . . 111
1.2 Quelques structures algébriques . . . 112
1.3 Les naturels . . . 113
1.3.1 Applications définies par récurrence . . . 114
1.3.2 Addition sur les naturels . . . 116
1.3.3 Ordre sur les naturels . . . 119
1.3.4 Multiplication dans les naturels . . . 124
1.3.5 Presque unicité des triplets naturels . . . 127
1.3.6 Écriture d'un naturel dans une base . . . 131
1.4 Les entiers . . . 136
1.4.1 Opposé . . . 137
1.4.2 Ordre sur $\mathbb{Z}$ . . . 138
1.5 Quelques résultats de cardinalité . . . 139
1.5.1 Équipotence, surpotence, subpotence . . . 139
1.5.2 Un peu d'infinité . . . 140
1.5.3 Dénombrabilité et ensemble des naturels . . . 146
1.5.4 Théorème de Cantor-Schröder-Bernstein . . . 150
1.5.5 Comparabilité cardinale . . . 153
1.5.6 Théorème de Cantor, ensemble des ensembles . . . 154
1.5.7 Ajouter ou soustraire des cardinalités . . . 155
1.6 Groupes . . . 160
1.6.1 Définition, unicité du neutre . . . 160
1.6.2 Groupe ordonné . . . 161
1.6.3 Classes de conjugaison . . . 161
1.7 Sous-groupe normal . . . 162
1.7.1 Permutations, groupe symétrique . . . 165
1.7.2 Décomposition en cycles . . . 166
1.8 Anneaux . . . 173
1.8.1 Anneau intègre . . . 176
1.8.2 Fonction puissance . . . 177
1.9 Le groupe et anneau des entiers . . . 177
1.9.1 Division euclidienne . . . 177
1.9.2 PGCD, PPCM et Bézout . . . 178
1.9.3 Sous-groupes de $(\mathbb{Z}, +)$ . . . 181
1.9.4 Résultats supplémentaires sur l'anneau des entiers . . . 181
1.10 Corps . . . 181
1.10.1 Définitions, morphismes . . . 181
1.11 Anneau intègre . . . 183
1.11.1 Élément premier . . . 183
1.12 Symbole de sommation . . . 185
1.12.1 Somme à valeurs dans un groupe commutatif . . . 185
1.13 Symbole de produit . . . 188
1.13.1 Sous-groupe engendré . . . 189
1.14 Module sur un anneau . . . 191
1.14.1 Module produit . . . 192
1.14.2 Sous-module . . . 194
1.15 Caractéristique d'un anneau . . . 196
1.15.1 Caractéristique deux . . . 197
1.16 Polynômes . . . 197

1.16.1 Polynômes d'une variable . . . 197
1.16.2 La notation $A[X]$ . . . 200
1.16.3 Action du groupe symétrique . . . 201
1.16.4 Corps des fractions . . . 202
1.16.5 Corps totalement ordonné . . . 204
1.17 Les rationnels . . . 206
1.17.1 Relation d'ordre . . . 207
1.17.2 Caractéristique . . . 207
1.18 Suite de Cauchy dans un corps totalement ordonné . . . 207
1.18.1 Suites de Cauchy dans les rationnels . . . 210
1.19 Insuffisance des rationnels . . . 211
1.20 Les réels . . . 215
1.20.1 L'ensemble . . . 215
1.20.2 Relation d'ordre . . . 217
1.20.3 Complétude . . . 224
1.20.4 Intervalles . . . 227
1.20.5 Maximum, supremum et compagnie . . . 227
1.20.6 Racines . . . 232
1.20.7 Corps valué . . . 233
1.20.8 Partie entière, partie fractionnaire . . . 233
1.21 Les complexes . . . 234

2 Théorie des groupes (Vol 1) 237
2.1 Groupes . . . 237
2.2 Groupe dérivé . . . 237
2.3 Théorèmes d'isomorphismes . . . 238
2.4 Indice d'un sous-groupe et ordre des éléments . . . 240
2.5 Suite de composition . . . 241
2.6 Groupes résolubles . . . 244
2.7 Action de groupes . . . 246
2.8 Produit semi-direct de groupes . . . 252
2.9 Groupe de torsion . . . 254
2.10 Famille presque nulle . . . 254

3 Anneaux (Vol 1) 255
3.1 Inversibles et nilpotents . . . 255
3.2 PGCD, PPCM et éléments inversibles . . . 256
3.2.1 Calcul effectif du PGCD et théorème de Bézout . . . 256
3.2.2 Générateurs . . . 258
3.2.3 Générateurs pour le groupe multiplicatif . . . 259
3.2.4 Décomposition en facteurs premiers . . . 259
3.2.5 Ordre d'un élément dans un groupe fini . . . 263
3.2.6 Écriture des fractions . . . 265
3.2.7 Équation diophantienne linéaire à deux inconnues . . . 266
3.2.8 Quotients . . . 267
3.3 Binôme de Newton et morphisme de Frobenius . . . 268
3.4 Idéal dans un anneau . . . 269
3.5 Polynômes de plusieurs variables . . . 271
3.5.1 Divisibilité et classes d'association . . . 272
3.5.2 PGCD et PPCM . . . 272
3.5.3 Anneaux intègres et corps . . . 273
3.5.4 Élément irréductible . . . 273
3.6 Anneau factoriel . . . 274
3.7 Anneau principal et idéal premier . . . 275

3.7.1 Autour du théorème de Bézout . . . 278
3.7.2 Idéal premier . . . 279
3.7.3 Anneau noethérien . . . 280
3.8 Anneau $\mathbb{Z}/6\mathbb{Z}$ . . . 281
3.9 Anneau euclidien . . . 283
3.9.1 Équations diophantiennes . . . 286
3.9.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$ . . . 287
3.10 Polynômes à coefficients dans un anneau commutatif . . . 289
3.10.1 Monômes . . . 290
3.10.2 Évaluation . . . 290
3.10.3 Polynômes sur un anneau intègre . . . 290
3.10.4 Division euclidienne . . . 291
3.10.5 Polynôme primitif . . . 292
3.10.6 Racines des polynômes . . . 292
3.10.7 Quelques identités . . . 295

4 Espaces vectoriels (début) (Vol 1) 297
4.1 Parties libres, génératrices, bases et dimension . . . 297
4.1.1 Et en dimension infinie . . . 303
4.1.2 Espace librement engendré . . . 305
4.2 Applications linéaires . . . 305
4.2.1 Définition . . . 306
4.2.2 Linéarité et bases . . . 308
4.2.3 Rang . . . 309
4.3 Matrices . . . 313
4.3.1 Définitions . . . 313
4.3.2 Identifier matrices et applications linéaires . . . 315
4.3.3 Déterminant . . . 319
4.3.4 Déterminant en petite dimension . . . 319
4.3.5 Manipulations de lignes et de colonnes . . . 320
4.3.6 Réduction de Gauss . . . 325
4.3.7 Matrices inversibles . . . 327
4.3.8 Inversibilité et déterminant . . . 328
4.3.9 Quelques ensembles de matrices particuliers . . . 329
4.3.10 Déterminant et combinaisons de lignes et colonnes . . . 329
4.3.11 Transvections . . . 330
4.3.12 Mineur, rang . . . 331
4.3.13 Matrices équivalentes et semblables . . . 332
4.3.14 Algorithme des facteurs invariants . . . 334
4.4 Changement de base . . . 335
4.4.1 Changement de base : vecteurs de base . . . 336
4.4.2 Changement de base : coordonnées . . . 336
4.4.3 Changement de base : matrice d'une application linéaire . . . 337
4.5 Espaces de polynômes . . . 338
4.6 Projection et orthogonalité . . . 339
4.7 Dualité . . . 340
4.7.1 Orthogonal . . . 342
4.8 Représentation de groupe . . . 342
4.9 Somme directe d'espaces vectoriels . . . 343
4.9.1 Structure réelle . . . 344

5 Classification de certains groupes (Vol 1) 347
5.1 Théorèmes de Sylow . . . 347
5.2 Groupe monogène . . . 351

5.3 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$ . . . 352
5.4 Groupes abéliens finis . . . 354
5.5 Groupes d'ordre pq . . . 356
5.6 Groupe symétrique, groupe alterné . . . 359
5.6.1 Le groupe alterné . . . 359
5.6.2 Sous-groupes normaux . . . 365
5.6.3 Indice . . . 366
5.7 Isométries du cube . . . 368

6 Corps (Vol 1) 371
6.1 Généralités . . . 371
6.1.1 Corps ordonnés . . . 371
6.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$ . . . 371
6.1.3 Corps premier . . . 373
6.1.4 Petit théorème de Fermat . . . 374
6.1.5 Nombres de Sophie Germain . . . 375
6.2 Théorème des deux carrés . . . 378
6.2.1 Un peu de structure dans $\mathbb{Z}[i]$ . . . 378
6.2.2 Résultats chinois . . . 381
6.3 Polynômes à coefficients dans un corps . . . 384
6.3.1 Irréductibilité . . . 384
6.3.2 Idéaux . . . 385
6.3.3 Identité de Bézout . . . 387
6.3.4 Lemme et théorème de Gauss . . . 388
6.3.5 Polynômes sur un corps et pgcd . . . 389
6.4 Extension de corps . . . 391
6.4.1 Extension et polynôme minimal . . . 392
6.4.2 Extensions algébriques et éléments transcendants . . . 394
6.4.3 Extension algébrique et polynôme minimal . . . 396
6.4.4 Extensions et polynômes . . . 397
6.4.5 Racines de polynômes . . . 407
6.4.6 Corps de rupture . . . 408
6.4.7 Pile d'extensions . . . 410
6.4.8 Clôture algébrique . . . 411
6.4.9 Polynômes à plusieurs variables . . . 417
6.4.10 Corps de décomposition . . . 418
6.4.11 Non irréductible ou pas corps ? . . . 422
6.4.12 Clôture algébrique . . . 422
6.4.13 Extensions séparables . . . 423
6.5 Idéal maximum . . . 428
6.5.1 Idéal maximum . . . 428
6.6 Polynômes symétriques et alternés . . . 429
6.6.1 Polynômes symétriques, alternés ou semi-symétriques . . . 429
6.6.2 Polynôme symétrique élémentaire . . . 429
6.6.3 Relations coefficients racines . . . 432
6.7 Minuscule morceau sur la théorie de Galois . . . 433

7 Topologie générale (Vol 1) 435
7.1 Éléments généraux de topologie . . . 435
7.1.1 Définitions et propriétés de base . . . 435
7.1.2 Base de topologie . . . 435
7.1.3 Fermés . . . 436
7.1.4 Quelques exemples . . . 436
7.1.5 Topologie produit . . . 438

7.2 Topologie rendant continue . . . 439
7.2.1 Topologie quotient . . . 440
7.2.2 Adhérence, fermeture, intérieur, point d'accumulation et point isolé . . . 441
7.3 Suites et convergence . . . 446
7.3.1 Convergence dans un fermé . . . 446
7.3.2 Pour des limites uniques : séparabilité . . . 446
7.3.3 Fonctions équivalentes . . . 448
7.4 Connexité . . . 449
7.5 Compacité . . . 451
7.5.1 Définition et notions connexes . . . 451
7.5.2 Espace localement compact . . . 452
7.5.3 Autres compacité . . . 452
7.5.4 Quelques propriétés . . . 453
7.5.5 Compactifié d'Alexandrov . . . 456
7.5.6 Propriété d'intersection finie . . . 458
7.6 Limite de fonction . . . 458
7.7 Topologie, distances et normes . . . 459
7.7.1 Distance et topologie métrique . . . 460
7.7.2 Topologie métrique et induite . . . 461
7.7.3 Intérieur, adhérence et frontière . . . 461
7.7.4 Boules ouvertes, fermées, sphères . . . 462
7.7.5 Continuité séquentielle . . . 463
7.7.6 Distance à un ensemble . . . 466
7.7.7 Convexité . . . 467
7.7.8 Norme . . . 467
7.8 Espaces vectoriels topologiques . . . 470
7.8.1 Corps topologique . . . 473
7.8.2 Voisinage symétrique et équilibré . . . 473
7.8.3 Limite de suites . . . 475
7.9 Continuité de fonctions . . . 476
7.9.1 Continuité . . . 476
7.9.2 Continuité et topologie induite . . . 479
7.9.3 Continuité et connexité . . . 480
7.9.4 Continuité et compacité . . . 480
7.9.5 Topologie et matrices . . . 481
7.10 Produit fini d'espaces vectoriels normés . . . 481
7.10.1 Distance et norme produit . . . 481
7.11 Topologie réelle en dimension n . . . 484
7.11.1 Ouverts et fermés . . . 484
7.11.2 Point d'accumulation, point isolé . . . 485
7.11.3 Limite de suite . . . 485
7.12 Topologie et distance . . . 486
7.12.1 Suites et espaces métriques . . . 487
7.12.2 Espace métrisable . . . 489
7.13 Suites de Cauchy, métrique et espaces complets . . . 490
7.13.1 Généralités . . . 490
7.13.2 Espace topologique métrique . . . 492
7.13.3 Compacts, fermés . . . 493
7.13.4 Équivalence entre Cauchy et $\tau-$Cauchy . . . 499
7.14 Norme, espace vectoriel normé . . . 501
7.15 Espaces métriques . . . 504
7.15.1 Espaces métrisables . . . 504
7.15.2 Fonctions continues . . . 504

7.15.3 Ensembles enchainés . . . 511
7.15.4 Produit fini d'espaces métriques . . . 512
7.15.5 Équicontinuité . . . 513
7.15.6 Continuité uniforme . . . 514
7.16 Ensembles nulle part denses . . . 514
7.17 Topologie des seminormes . . . 515
7.17.1 Espace dual . . . 521
7.17.2 Espace $C^k(\mathbb{R}, E')$ . . . 521
7.18 Espaces de Baire . . . 522

8 Espaces affines (Vol 1) 525
8.1 Vecteurs agissant sur un espace . . . 525
8.2 Repères cartésiens affines . . . 526
8.3 Classification affine des coniques . . . 526
8.4 Applications affines . . . 529
8.4.1 Autres propriétés . . . 530
8.5 Isomorphismes . . . 531
8.6 Sous espaces affines . . . 532
8.7 Barycentre . . . 534
8.7.1 Sous-espaces affines . . . 534
8.7.2 Enveloppe convexe . . . 536
8.7.3 Applications affines et barycentre . . . 539
8.8 Repères, coordonnées cartésiennes et barycentriques . . . 540
8.8.1 Équation de droite . . . 543
8.8.2 Associativité, coordonnées barycentriques dans un triangle . . . 544
8.9 Applications affines sur $\mathbb{R}^n$ . . . 545
8.9.1 Structure de groupe pour les applications affines . . . 546
8.10 Isométries . . . 547

9 Espaces vectoriels (encore) (Vol 1) 549
9.1 Déterminants . . . 549
9.1.1 Formes multilinéaires alternées . . . 549
9.1.2 Déterminant d'une famille de vecteurs . . . 550
9.1.3 Déterminant d'un endomorphisme . . . 553
9.1.4 Déterminant de Vandermonde . . . 555
9.1.5 Déterminant de Gram . . . 557
9.1.6 Déterminant de Cauchy . . . 558
9.1.7 Matrice de Sylvester . . . 558
9.1.8 Théorème de Kronecker . . . 560
9.2 Orientation . . . 562
9.2.1 Cas vectoriel . . . 562
9.2.2 Cas affine . . . 564
9.3 Hermitien, orthogonal, adjoint . . . 564
9.3.1 Opérateur orthogonal, matrice orthogonale . . . 566
9.4 Topologie . . . 567
9.4.1 Boules et sphères . . . 567
9.4.2 Ouverts, fermés, intérieur et adhérence . . . 569
9.4.3 Point isolé, point d'accumulation . . . 575
9.4.4 Des exemples . . . 576
9.5 Valeur propre et vecteur propre . . . 579
9.5.1 Généralités . . . 579
9.5.2 Dans le vif du sujet . . . 580
9.6 Polynômes d'endomorphismes . . . 581
9.6.1 Polynômes d'endomorphismes . . . 581

9.6.2 Polynôme minimal et minimal ponctuel . . . 584
9.6.3 Polynôme caractéristique . . . 590
9.7 Formes bilinéaires et quadratiques . . . 594
9.7.1 Dégénérescence d'une forme bilinéaire . . . 595
9.7.2 Topologie . . . 595
9.7.3 Isotropie . . . 596
9.8 Formes bilinéaires et quadratiques . . . 596
9.8.1 Généralités . . . 597
9.8.2 Matrice associée à une forme bilinéaire . . . 597
9.8.3 Changement de base : matrice d'une forme bilinéaire . . . 598
9.8.4 Orthogonalité . . . 599
9.8.5 Isométrie, forme quadratique et bilinéaire . . . 599
9.8.6 Isométries . . . 600
9.9 Signature, théorème de Sylvester . . . 602
9.9.1 Équivalence de formes quadratiques . . . 604
9.10 Produit scalaire, produit hermitien . . . 604
9.10.1 Hermitien, unitaire, etc. . . . 606
9.10.2 Éléments de matrice . . . 607
9.10.3 Transposée : pas d'approche naïve . . . 608
9.10.4 Transposée : la bonne approche . . . 610
9.10.5 Polynômes de Lagrange . . . 612
9.10.6 Dual de $\mathbb{M}(n, \mathbb{K})$ . . . 612
9.11 Diagonalisation et trigonalisation . . . 614
9.11.1 Matrices semblables . . . 614
9.11.2 Trace de matrices semblables . . . 615
9.11.3 Endomorphismes nilpotents . . . 615
9.11.4 Endomorphismes diagonalisables . . . 618
9.11.5 Diagonalisation : cas complexe, pas toujours . . . 621
9.11.6 Diagonalisation : cas réel . . . 621
9.11.7 Matrice définie positive . . . 623
9.11.8 Réduction de Gauss . . . 626
9.11.9 Diagonalisation . . . 632
9.11.10 Diagonalisation . . . 633
9.12 Fonctions . . . 634
9.13 Sous espaces caractéristiques . . . 635
9.13.1 Théorèmes de décomposition . . . 637
9.13.2 Valeurs singulières . . . 639
9.14 Extension du corps de base . . . 640
9.14.1 Extension des applications linéaires . . . 640
9.14.2 Projections . . . 642
9.14.3 Rang, polynôme minimal, polynôme caractéristique . . . 645
9.15 Frobenius et Jordan . . . 646
9.15.1 Matrice compagnon . . . 646
9.15.2 Réduction de Frobenius . . . 647
9.15.3 Forme normale de Jordan . . . 650
9.16 Commutant et endomorphismes cycliques . . . 651
9.16.1 Endomorphisme cyclique . . . 651
9.16.2 Commutant : cas diagonalisable . . . 652
9.16.3 Commutant : cas général . . . 654
9.17 Hyperplans et formes linéaires . . . 657
9.17.1 Trouver la matrice d'une symétrie donnée . . . 658
9.18 Théorème de Burnside . . . 660
9.19 Ellipsoïde . . . 662

9.20 Système d'équations linéaires : méthode de Gauss . . . 663

10 Analyse réelle : topologie et continuité (Vol 2) 665
10.1 Intervalles . . . 665
10.2 Application réciproque . . . 666
10.2.1 Définitions . . . 666
10.2.2 Graphe de la fonction réciproque . . . 667
10.3 Topologie sur l'ensemble des réels . . . 667
10.3.1 Compacité pour les réels . . . 668
10.3.2 Conséquence : les fermés bornés sont compacts . . . 670
10.3.3 Suites et limites dans les réels . . . 671
10.3.4 Opérations sur les limites . . . 672
10.3.5 Exemples . . . 673
10.3.6 Limites infinies . . . 673
10.3.7 Suites croissantes et bornées . . . 674
10.3.8 Suites adjacentes . . . 675
10.3.9 Limite supérieure et inférieure . . . 676
10.3.10 Ouverts, voisinage, topologie . . . 677
10.3.11 Intervalles et connexité . . . 678
10.3.12 Recouvrement par des intervalles ouverts . . . 683
10.3.13 Connexité par arcs . . . 684
10.3.14 Un peu de connexité par arcs . . . 685
10.3.15 Des exemples . . . 686
10.3.16 Quelques mots à propos de la droite réelle achevée . . . 688
10.4 Continuité . . . 689
10.4.1 Opération sur la continuité . . . 690
10.4.2 La fonction la moins continue du monde . . . 691
10.4.3 Approche topologique . . . 692
10.4.4 Module sur les nombres complexes . . . 694
10.5 Norme à partir d'un produit scalaire . . . 697
10.5.1 Continuité de la racine carrée, invitation à la topologie induite . . . 697
10.5.2 Second degré . . . 699

11 Espaces vectoriels normés (Vol 2) 701
11.0.1 Norme, produit scalaire et Cauchy-Schwarz (cas réel) . . . 701
11.1 Théorème spectral autoadjoint . . . 704
11.1.1 Inégalité de Minkowski . . . 706
11.1.2 Cauchy-Schwarz etc. cas complexe . . . 708
11.1.3 Diagonalisation : cas complexe, ce qu'on a . . . 709
11.1.4 Projection et orthogonalité . . . 709
11.1.5 Théorème de Pythagore . . . 711
11.1.6 Produit vectoriel . . . 711
11.1.7 Produit mixte . . . 715
11.1.8 Procédé de Gram-Schmidt . . . 716
11.1.9 Pseudo-réduction simultanée . . . 717
11.2 Approximations . . . 718
11.2.1 Quelques exemples de normes sur $\mathbb{R}^n$ . . . 719
11.3 Équivalence des normes . . . 720
11.3.1 En dimension finie . . . 720
11.3.2 Contre-exemple en dimension infinie . . . 724
11.4 Norme opérateur . . . 725
11.4.1 Norme d'algèbre . . . 726
11.5 Application linéaire continue et bornée . . . 728
11.5.1 Suites . . . 730

11.5.2 Continuité du produit de matrices . . . 732
11.6 Applications multilinéaires . . . 732
11.7 Séries . . . 734
11.7.1 Les trois types de convergence . . . 735
11.7.2 Séries dans une algèbre normée . . . 738
11.8 Sommes de familles infinies . . . 740
11.8.1 Convergence commutative . . . 740
11.8.2 Somme non dénombrables . . . 743
11.8.3 Sommes dénombrables . . . 744
11.9 Série réelle . . . 747
11.9.1 Critères de convergence absolue . . . 747
11.9.2 Critères de convergence simple . . . 749
11.9.3 Quelques séries usuelles . . . 749
11.9.4 Séries alternées . . . 752
11.9.5 Moyenne de Cesàro . . . 753
11.9.6 Écriture décimale d'un réel . . . 754
11.9.7 Théorème de Banach-Steinhaus . . . 757
11.9.8 Convergence forte . . . 760
11.10 Application ouverte . . . 761
11.11 Produit tensoriel d'espaces vectoriels . . . 764
11.11.1 Les produits tensoriels . . . 765
11.11.2 Le produit tensoriel . . . 766
11.11.3 Bases . . . 770
11.11.4 Norme . . . 771
11.11.5 Applications bilinéaires, matrices et produit tensoriel . . . 772
11.11.6 Application d'opérateurs . . . 773
11.11.7 Convergence en norme et par composante . . . 773
11.12 Calcul différentiel dans un espace vectoriel normé . . . 774
11.12.1 Définition de la différentielle . . . 774
11.12.2 Différentielle d'applications linéaires . . . 775
11.12.3 Accroissements finis . . . 776
11.12.4 Notations pour les applications linéaires . . . 777
11.12.5 (non ?) Différentiabilité des applications linéaires . . . 777
11.12.6 Dérivation en chaine et formule de Leibnitz . . . 777
11.12.7 Différentiation de produit . . . 782
11.12.8 Formule des accroissements finis . . . 786
11.12.9 Applications multilinéaires . . . 788
11.12.10 Différentielle partielle . . . 790
11.12.11 L'inverse, sa différentielle . . . 791
11.13 Exponentielle de matrice . . . 793
11.14 Espace dual . . . 794
11.14.1 Topologies . . . 794
11.14.2 Module de continuité . . . 796
11.15 Mini introduction aux nombres p-adiques . . . 797
11.15.1 La flèche d'Achille . . . 797
11.15.2 La tortue et Achille . . . 798
11.15.3 Dans les nombres p-adiques, c'est vrai . . . 799

12 Analyse réelle : limites et dérivation (Vol 2) 801
12.1 Limite de fonctions . . . 801
12.1.1 Définition . . . 801
12.1.2 Quelques règles de calcul . . . 803
12.2 Limites pointées et épointées . . . 807
12.2.1 Théorèmes de composition de limites . . . 809

12.2.2 Discussion pointée Vs épointée . . . 812
12.3 Limites en l'infini . . . 815
12.3.1 Limite en des nombres . . . 817
12.3.2 Limites quand tout va bien . . . 817
12.3.3 Limites de fonctions . . . 819
12.3.4 Limite à gauche et à droite . . . 819
12.4 Limite en compactifié d'Alexandrov . . . 820
12.4.1 Prolongement par continuité . . . 821
12.4.2 Prolongement par continuité . . . 822
12.4.3 Théorème de la bijection . . . 823
12.5 Limite et continuité . . . 824
12.5.1 Prolongement des rationnels vers les réels . . . 827
12.6 Espace des fonctions continues . . . 830
12.7 Uniforme continuité . . . 834
12.8 Uniforme continuité . . . 835
12.9 Fonctions sur un compact . . . 838
12.10 Polynômes, théorème de d'Alembert . . . 838
12.10.1 Polynômes sur les réels . . . 838
12.10.2 Polynômes sur les complexes . . . 838
12.11 Trigonalisation . . . 844
12.11.1 Trigonalisation : généralités . . . 844
12.11.2 Trigonalisation : cas complexe . . . 845
12.12 Matrices, spectre et norme . . . 851
12.12.1 Rayon spectral . . . 852
12.12.2 Normes de matrices et d'applications linéaires . . . 856
12.13 Géométrie dans l'espace . . . 860
12.13.1 Droites dans l'espace . . . 861
12.13.2 Projection orthogonale . . . 863
12.13.3 Plan médiateur . . . 865
12.13.4 Tétraèdre . . . 866
12.14 Géométrie dans le plan . . . 867
12.15 Dérivée : exemples introductifs . . . 871
12.15.1 La vitesse . . . 871
12.15.2 La tangente à une courbe . . . 872
12.15.3 L'aire en dessous d'une courbe . . . 872
12.16 Dérivation de fonctions réelles . . . 874
12.16.1 Exemples . . . 875
12.16.2 Interprétation géométrique : tangente . . . 876
12.16.3 Interprétation géométrique : approximation affine . . . 877
12.16.4 Développement limité au premier ordre . . . 878
12.17 Règles de calcul . . . 879
12.17.1 Dérivée de la réciproque . . . 883
12.17.2 Dérivée de fonction composée . . . 884
12.17.3 Dérivée de fonction périodique . . . 885
12.18 Dérivation et croissance . . . 886
12.18.1 Théorèmes de Rolle et des accroissements finis . . . 888
12.18.2 Règle de l'Hospital . . . 890
12.18.3 Dérivée et primitive . . . 892
12.19 Fonctions de plusieurs variables . . . 893
12.19.1 Graphes de fonctions à plusieurs variables . . . 896
12.19.2 Courbes de niveau . . . 897
12.20 Limites à plusieurs variables . . . 901
12.20.1 Caractérisation de la limite par les suites . . . 903

12.20.2 Règle de l'étau . . . 905
12.20.3 Méthode des chemins . . . 906
12.21 Dérivée directionnelle . . . 908
12.21.1 Dérivée partielle et directionnelles . . . 908
12.21.2 Gradient : direction de plus grande pente . . . 912
12.21.3 Gradient : orthogonal au plan tangent . . . 913
12.21.4 Mise en bouche en dimension 2 . . . 914
12.21.5 Accroissements finis et dérivées partielles . . . 915
12.22 Formes différentielles . . . 917
12.22.1 Décomposition dans la base duale . . . 917
12.22.2 L'isomorphisme musical . . . 917
12.23 Différentielle . . . 918
12.23.1 Exemples introductifs . . . 918
12.23.2 Différentielle . . . 919
12.23.3 Matrice de la différentielle . . . 920
12.23.4 Différentielle, dual et forme différentielle . . . 921
12.23.5 Ce n'est pas la différentielle extérieure . . . 922
12.23.6 Continuité, dérivabilité et différentiabilité . . . 923
12.23.7 Calcul de valeurs approchées . . . 926
12.23.8 Différentielle et tangente . . . 927
12.23.9 Prouver qu'une fonction n'est pas différentiable . . . 928
12.23.10 Gradient . . . 932
12.23.11 Linéarité . . . 932
12.24 Produit . . . 933
12.24.1 Difficulté d'ordre supérieur . . . 934
12.24.2 Solution : produit tensoriel . . . 935
12.24.3 Formes bilinéaires . . . 935
12.25 Différentielle de fonction composée . . . 936
12.25.1 Fonctions composées . . . 939
12.26 Autres trucs sur la différentielle . . . 940
12.26.1 Différentielle et dérivées partielles . . . 940
12.26.2 Plan tangent . . . 940
12.26.3 Notes idéologiques quant au concept de plan tangent . . . 940
12.26.4 Gradient et recherche du plan tangent . . . 941
12.27 Jacobienne . . . 945
12.27.1 Rappels et définitions . . . 945
12.28 Fonctions de classe C^1 . . . 946
12.29 Différentielle et dérivée complexe . . . 955
12.29.1 Quelques règles de calcul . . . 958
12.30 Théorèmes des accroissements finis . . . 959
12.31 Fonctions Lipschitziennes . . . 959
12.32 Différentielles d'ordre supérieur . . . 961
12.32.1 Différentielle et dérivées partielles . . . 961
12.32.2 Espaces d'applications multilinéaires et identifications . . . 965
12.32.3 Fonctions différentiables plusieurs fois . . . 969
12.32.4 Différentielle seconde, fonction de classe C^2 . . . 970
12.32.5 Ordre supérieur . . . 973
12.33 Suites et séries : généralités . . . 977
12.33.1 Norme suprémum . . . 977
12.33.2 Convergence uniforme . . . 977
12.33.3 Série de fonctions . . . 981
12.34 Permuter limite et dérivée . . . 983
12.35 La fonction puissance . . . 990

12.35.1 Sur les naturels . . . 990
12.35.2 Sur les rationnels, racines . . . 992
12.36 Densité des polynômes . . . 1006
12.36.1 Théorème de Stone-Weierstrass . . . 1006
12.37 Primitive de fonction continue . . . 1010
12.37.1 Dérivation de la fonction puissance (première) . . . 1011
12.37.2 Équation fonctionnelle . . . 1013
12.37.3 Dérivation de la fonction puissance (seconde) . . . 1016
12.37.4 Vers les complexes . . . 1017
12.38 Polynômes de Taylor . . . 1019
12.38.1 Fonctions « petit o » . . . 1023
12.38.2 Autres formulations . . . 1024
12.38.3 Formule et reste . . . 1025
12.38.4 Reste intégral . . . 1025
12.39 Développement limité autour de zéro . . . 1025
12.39.1 Généralités . . . 1025
12.39.2 Formule de Taylor-Young . . . 1027
12.39.3 Règles de calcul . . . 1028
12.40 Développement ailleurs qu'à l'origine . . . 1030
12.41 Développement au voisinage de l'infini . . . 1031
12.41.1 La fonction puissance : remarques pour la suite . . . 1031
12.42 Fonctions réelles de deux variables réelles . . . 1031
12.42.1 Limites de fonctions à deux variables . . . 1031
12.42.2 Dérivées partielles . . . 1033
12.42.3 Différentielle et accroissement . . . 1033
12.43 Les fonctions à valeurs vectorielles . . . 1034
12.44 Fonctions vectorielles de plusieurs variables . . . 1035
12.45 Limites à plusieurs variables . . . 1035
12.46 Champs de vecteurs . . . 1038
12.46.1 Matrice jacobienne . . . 1039
12.47 Divergence, rotationnel et l'opérateur nabla . . . 1039
12.48 Interprétation de la divergence . . . 1042
12.49 Quelques formules de Leibnitz . . . 1044

13 Analyse sur des groupes (Vol 2) 1045
13.1 Action de groupe et connexité . . . 1045
13.2 Espaces de matrices . . . 1047
13.2.1 Dilatations et transvections . . . 1047
13.2.2 Connexité de certains groupes . . . 1053
13.2.3 Densité . . . 1055
13.2.4 Racine carrée d'une matrice hermitienne positive . . . 1057
13.2.5 Racine carrée d'une matrice symétrique positive . . . 1058
13.2.6 Décomposition polaires : cas réel . . . 1059
13.2.7 Enveloppe convexe . . . 1061
13.2.8 Décomposition de Bruhat . . . 1063
13.3 Sous-groupes du groupe linéaire . . . 1065

14 Tribus, théorie de la mesure, intégration (Vol 2) 1069
14.1 Tribus . . . 1069
14.1.1 Généralités . . . 1069
14.1.2 Tribu induite . . . 1070
14.2 Théorie de la mesure . . . 1071
14.2.1 Mesure sur un ensemble de parties . . . 1071
14.2.2 Mesure sur une algèbre de parties . . . 1071

14.2.3 Mesure sur une tribu, espace mesuré . . . 1073
14.2.4 Mesure extérieure . . . 1080
14.3 Applications mesurables . . . 1082
14.3.1 Propriétés . . . 1082
14.3.2 D'une tribu à l'autre . . . 1082
14.4 Tribu borélienne . . . 1084
14.4.1 Applications continues et boréliennes . . . 1086
14.4.2 Tribu de Baire . . . 1088
14.5 Espace mesuré complet . . . 1090
14.5.1 Partie négligeable . . . 1090
14.5.2 Prolongement . . . 1099
14.5.3 Mesure image . . . 1101
14.5.4 Régularité d'une mesure . . . 1102
14.5.5 Théorème de récurrence . . . 1106
14.6 Mesurabilité des fonctions à valeurs réelles . . . 1107
14.6.1 Fonctions à valeurs réelles sur un espace mesurable . . . 1107
14.6.2 Fonction étagée . . . 1113
14.6.3 Fonctions réelles à variables réelles . . . 1116
14.7 Tribu produit . . . 1117
14.7.1 Produit d'espaces mesurables . . . 1117
14.7.2 Le cas des boréliens . . . 1118
14.8 Mesure de Lebesgue sur $\mathbb{R}$ . . . 1119
14.8.1 Mesure et tribu de Lebesgue . . . 1124
14.8.2 Propriétés de la mesure de Lebesgue . . . 1125
14.8.3 Fonctions mesurables . . . 1130
14.8.4 Ensemble de Vitali (non mesurable) . . . 1130
14.8.5 Ensemble de Cantor . . . 1131
14.8.6 Mesure positive sans intervalle . . . 1133
14.9 Intégrale par rapport à une mesure . . . 1134
14.9.1 Définition pour les fonctions à valeurs positives . . . 1134
14.9.2 Premières propriétés . . . 1136
14.9.3 Propriétés plus avancées . . . 1139
14.9.4 Fonctions à valeurs réelles . . . 1143
14.9.5 Additivité de l'intégrale . . . 1144
14.9.6 Fonctions à valeurs vectorielles (dimension finie) . . . 1145
14.9.7 Quelques propriétés . . . 1147
14.9.8 Permuter limite et intégrale . . . 1148
14.9.9 Additivité de l'intégrale de Lebesgque . . . 1150
14.9.10 Produit d'une mesure par une fonction (mesure à densité) . . . 1152
14.9.11 Mesure et topologie . . . 1154
14.10 Propriétés des intégrales . . . 1156
14.11 Mesure à densité . . . 1157
14.11.1 Théorème de Radon-Nikodym . . . 1157
14.11.2 Mesure complexe . . . 1158
14.11.3 Théorème d'approximation . . . 1159
14.12 Produit de mesures . . . 1162
14.13 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$ . . . 1167
14.13.1 Ensembles négligeables . . . 1168
14.13.2 Parties et fonctions mesurables . . . 1169
14.13.3 Propriétés d'unicité . . . 1170
14.13.4 Régularité . . . 1172
14.14 Propriétés de l'intégrale de Lebesgue . . . 1172
14.14.1 Quelques limites dans les bornes . . . 1173

14.14.2 Mesure de comptage et série . . . 1175
14.14.3 Théorème de la moyenne . . . 1177
14.14.4 Primitives et intégrales . . . 1178
14.14.5 Exemples et applications . . . 1179
14.14.6 Permuter limite et dérivée . . . 1180
14.14.7 Intégrales impropres . . . 1182
14.15 Changement de variables dans une intégrale multiple . . . 1185
14.15.1 Des lemmes . . . 1185
14.15.2 Déterminant et mesure de Lebesgue . . . 1186
14.15.3 Le théorème et sa démonstration . . . 1187
14.15.4 Exemples . . . 1192
14.16 Changement d'espace mesuré . . . 1192
14.17 Théorème de Fubini-Tonelli et de Fubini . . . 1193

15 Suites et séries de fonctions (Vol 2) 1203
15.1 Séries de fonctions . . . 1203
15.1.1 Intégration de séries de fonctions . . . 1203
15.1.2 Différentiabilité . . . 1204
15.2 Séries entières . . . 1209
15.2.1 Disque de convergence . . . 1209
15.2.2 Somme et produit de séries . . . 1214
15.2.3 Convergence normale . . . 1219
15.2.4 Dérivation . . . 1222
15.2.5 Intégration . . . 1225
15.3 Séries de Taylor . . . 1225
15.3.1 Polynôme de Taylor d'une série entière . . . 1225
15.3.2 Une majoration pour le reste . . . 1226
15.3.3 Fonctions analytiques . . . 1227
15.4 Algèbre engendrée par une matrice . . . 1228
15.5 Exponentielle sur une algèbre normée . . . 1229
15.5.1 Définition . . . 1229
15.5.2 Différentielles . . . 1231
15.5.3 Exponentielle de matrice . . . 1234
15.6 Exponentielle et logarithme dans les réels . . . 1237
15.6.1 L'équation différentielle . . . 1237
15.6.2 Existence . . . 1238
15.6.3 Le nombre de Neper e . . . 1239
15.6.4 Application réciproque : logarithme . . . 1240
15.6.5 Approximations numériques de e . . . 1243
15.6.6 Résumé des propriétés de l'exponentielle . . . 1244
15.6.7 Dérivée de la fonction puissance . . . 1245
15.6.8 Dérivée du logarithme . . . 1246
15.6.9 Taylor pour l'exponentielle . . . 1246
15.6.10 Analycité . . . 1247
15.6.11 Autres propriétés et petits calculs . . . 1247
15.6.12 Taylor pour le logarithme . . . 1247
15.6.13 Développements et calcul de limites . . . 1251
15.6.14 Une petite intégrale . . . 1251
15.7 Vitesses des puissances, de l'exponentielle et du logarithme . . . 1252
15.7.1 Un peu de théorie . . . 1252
15.7.2 Nombres premiers . . . 1257
15.7.3 Quelques limites . . . 1259
15.8 Trigonométrie hyperbolique . . . 1260
15.9 Séries entières de matrices . . . 1262

15.9.1 Différentiabilité . . . 1262
15.10 Exponentielle de matrices . . . 1264
15.10.1 Diagonalisabilité d'exponentielle . . . 1265
15.11 Étude d'asymptote . . . 1267
15.12 Développement en série . . . 1268
15.12.1 Série génératrice d'une suite . . . 1268
15.12.2 Développement en série et Taylor . . . 1268
15.12.3 Resommer une série . . . 1270
15.13 Séries entières de matrices . . . 1275
15.13.1 Rayon de convergence . . . 1275
15.13.2 Convergence et rayon spectral . . . 1275
15.13.3 Exponentielle et logarithme de matrice . . . 1277
15.13.4 Calcul effectif de l'exponentielle d'une matrice . . . 1280
15.14 Lemme de Borel . . . 1281
15.14.1 Fonctions plateaux, Urysohn, partition de l'unité . . . 1281
15.14.2 Lemme de Urysohn . . . 1284
15.14.3 Partition de l'unité . . . 1285
15.14.4 Le lemme de Borel . . . 1286
15.15 Nombres de Bell . . . 1288

16 Représentations et caractères (Vol 3) 1293
16.1 Représentations et caractères . . . 1293
16.1.1 Crochet de dualité et transformée de Fourier . . . 1295
16.1.2 Groupes non abéliens . . . 1296
16.1.3 Représentations linéaires des groupes finis . . . 1296
16.1.4 Module . . . 1298
16.1.5 Structure hermitienne . . . 1299
16.1.6 Caractères . . . 1300
16.2 Équivalence de représentations et caractères . . . 1300
16.2.1 Représentation régulière . . . 1303
16.2.2 Caractères et représentations : suite et fin . . . 1304
16.3 Représentation produit tensoriel . . . 1306
16.4 Exemple sur le groupe symétrique . . . 1307
16.5 Table des caractères du groupe symétrique S_4 . . . 1307
16.5.1 Calculs à partir de rien ou presque . . . 1307
16.5.2 À propos de la représentation ρ_s . . . 1309
16.5.3 À propos de la représentation ρ_u . . . 1310

17 Encore de l'analyse (et c'est pas fini) (Vol 3) 1313
17.1 Densité des polynômes . . . 1313
17.2 Primitive et intégrale . . . 1313
17.2.1 Théorème taubérien de Hardy-Littlewood . . . 1314
17.2.2 Théorème de Müntz . . . 1317
17.3 Intégrales convergeant uniformément . . . 1320
17.3.1 Définition et propriété . . . 1320
17.3.2 Critères de convergence uniforme . . . 1321
17.4 Fonctions définies par une intégrale . . . 1321
17.4.1 Continuité sous l'intégrale . . . 1322
17.4.2 Le coup du compact . . . 1323
17.4.3 Dérivabilité sous l'intégrale . . . 1323
17.4.4 Absolue continuité . . . 1326
17.4.5 Différentiabilité sous l'intégrale . . . 1328
17.5 Deux théorèmes de point fixe . . . 1331
17.5.1 Points fixes attractifs et répulsifs . . . 1331

17.5.2 Picard . . . 1332
17.6 Théorèmes de point fixes et équations différentielles . . . 1335
17.6.1 Théorème de Cauchy-Lipschitz . . . 1335
17.7 Théorèmes d'inversion locale et de la fonction implicite . . . 1342
17.7.1 Mise en situation . . . 1342
17.7.2 Théorème d'inversion locale . . . 1343
17.7.3 Théorème de la fonction implicite . . . 1346
17.7.4 Exemple . . . 1348
17.8 Décomposition polaire (régularité) . . . 1349
17.9 Théorème de Von Neumann . . . 1351
17.10 Recherche d'extrémums . . . 1354
17.10.1 Extrema à une variable . . . 1354
17.10.2 Extrema libre . . . 1355
17.10.3 Extremums et Hessienne . . . 1356
17.10.4 Un peu de recettes de cuisine . . . 1358
17.10.5 Extrema liés . . . 1359
17.11 Fonctions convexes . . . 1361
17.11.1 Inégalité des pentes . . . 1361
17.11.2 Convexité et régularité . . . 1362
17.11.3 Dérivées d'une fonction convexe . . . 1363
17.11.4 Graphe d'une fonction convexe . . . 1365
17.11.5 Convexité et hessienne . . . 1370
17.11.6 Quelques inégalités . . . 1376
17.11.7 Norme l^p . . . 1378
17.11.8 Hölder . . . 1379
17.12 Trucs et astuces de calcul d'intégrales . . . 1383
17.12.1 Quelques intégrales « usuelles » . . . 1383
17.12.2 Reformer un carré au dénominateur . . . 1386
17.12.3 Décomposition en fractions simples . . . 1386
17.13 Algorithme du gradient à pas optimal . . . 1387
17.14 Formes quadratiques, signature, et lemme de Morse . . . 1391
17.15 Ellipsoïde de John-Loewner . . . 1395
17.16 Prolongement de fonctions . . . 1401
17.16.1 Encore du prolongement . . . 1403
17.17 Complétion d'un espace métrique . . . 1405
17.17.1 Principe des zéros isolés . . . 1409
17.18 Un petit extra . . . 1410

18 Trigonométrie, isométries (Vol 3) 1413
18.1 Trigonométrie . . . 1413
18.1.1 Définitions, périodicité et quelques valeurs remarquables . . . 1413
18.1.2 Fonction puissance (pour les complexes) . . . 1414
18.1.3 Formules de trigonométrie . . . 1416
18.2 Très modeste approximation de π . . . 1426
18.3 Cercle trigonométriques . . . 1427
18.3.1 Les fonctions tangente et arc tangente . . . 1428
18.3.2 La fonction arc sinus . . . 1430
18.3.3 La fonction arc cosinus . . . 1432
18.3.4 Une meilleure approximation de π . . . 1434
18.3.5 Angle entre deux vecteurs . . . 1435
18.3.6 Aire du parallélogramme . . . 1436
18.4 Paramétrisation du cercle . . . 1437
18.4.1 Bijection continue . . . 1438
18.4.2 Inverse . . . 1440

18.4.3 Cercle trigonométrique . . . 1440
18.4.4 Du point de vue de la tribu, mesure et co. . . . 1441
18.5 Exemples trigonométriques . . . 1444
18.5.1 Quelques équations trigonométriques . . . 1445
18.5.2 Développements en série . . . 1445
18.6 Isométries de l'espace euclidien . . . 1447
18.6.1 Structure du groupe $\mathrm{Isom}(\mathbb{R}^n)$ . . . 1447
18.7 Isométries dans $\mathbb{R}^n$ . . . 1449
18.7.1 Préserver l'orientation . . . 1453
18.8 Groupes finis d'isométries . . . 1454
18.8.1 Points fixés par une affinité . . . 1456
18.9 Classification des isométries dans $\mathbb{R}^2$ . . . 1456
18.9.1 Projection orthogonale . . . 1456
18.9.2 Réflexions . . . 1457
18.9.3 Segment, plan médiateur et équidistance . . . 1458
18.9.4 Translations et réflexions . . . 1460
18.9.5 Rotations . . . 1460
18.9.6 Rotation d'un angle donné . . . 1463
18.9.7 Rotations vectorielles . . . 1465
18.9.8 Matrice des transformations orthogonales . . . 1467
18.9.9 Rotations, SO(2) et matrice de rotation . . . 1469
18.9.10 Rotation et application affine . . . 1470
18.9.11 Angle orienté . . . 1471
18.9.12 Angles et nombres complexes . . . 1475
18.9.13 Polygone convexe . . . 1480
18.9.14 Groupe diédral . . . 1481
18.9.15 Applications : du dénombrement . . . 1488
18.9.16 Classification . . . 1489
18.9.17 Classification des isométries de $\mathbb{R}$ . . . 1492
18.9.18 Isométries du tétraèdre régulier . . . 1493
18.9.19 Représentation de S_4 via les isométries du tétraèdre . . . 1494
18.10 Transformations de Lorentz . . . 1495
18.10.1 Sous-groupe fini d'isométries du plan . . . 1498
18.10.2 Relations trigonométriques dans un triangle rectangle . . . 1502
18.10.3 Pavages du plan . . . 1504
18.11 Un peu de structure de $\mathrm{O}(n)$ . . . 1519
18.11.1 Valeurs propres dans $\mathrm{O}(n)$ . . . 1519
18.11.2 Sous-groupes finis de SO(3) . . . 1522
18.12 Systèmes de coordonnées . . . 1531
18.12.1 Coordonnées polaires . . . 1531
18.12.2 Coordonnées cylindriques . . . 1538
18.12.3 Coordonnées sphériques . . . 1539
18.13 Calcul de limites . . . 1540
18.13.1 Méthode des coordonnées polaires . . . 1541
18.13.2 Méthode du développement asymptotique . . . 1544
18.14 Quelques intégrales avec de la trigonométrie . . . 1545
18.14.1 Changement de variables . . . 1545
18.14.2 Coordonnées polaires . . . 1546
18.14.3 Coordonnées cylindriques . . . 1548
18.14.4 Coordonnées sphériques . . . 1549
18.14.5 Un autre système utile . . . 1549
18.15 Aire d'une surface de révolution . . . 1551
18.16 Table de caractères du groupe diédral . . . 1552

18.16.1 Représentations de dimension un . . . 1552
18.16.2 Représentations de dimension deux . . . 1553
18.16.3 Le compte pour n pair . . . 1555
18.16.4 Le compte pour n impair . . . 1555

19 Corps finis, racines de l'unité (Vol 3) 1557
19.1 Le groupe des racines de l'unité dans les nombres complexes . . . 1557
19.1.1 Le groupe . . . 1557
19.1.2 Fonction indicatrice d'Euler (première partie) . . . 1559
19.1.3 Introduction par les racines de l'unité . . . 1560
19.1.4 Générateurs . . . 1561
19.1.5 Fonction indicatrice d'Euler (propriétés) . . . 1561
19.1.6 Décomposition en éléments simples . . . 1562
19.2 Chiffrement RSA . . . 1563
19.2.1 Mise en place par Bob . . . 1563
19.2.2 Chiffrement . . . 1563
19.2.3 Déchiffrement . . . 1563
19.2.4 Une imprudence à ne pas commettre . . . 1564
19.2.5 Problèmes calculatoires . . . 1564
19.2.6 La solidité de RSA . . . 1565
19.2.7 Note non mathématique pour doucher l'enthousiasme . . . 1565
19.3 Polynômes cyclotomiques . . . 1565
19.3.1 Définitions et propriétés . . . 1565
19.3.2 Nombres premiers . . . 1569
19.4 Corps finis . . . 1571
19.4.1 Théorème de Wedderburn . . . 1571
19.4.2 Existence, unicité . . . 1573
19.4.3 Symboles de Legendre et carrés . . . 1575
19.4.4 Théorème de Chevalley-Warning . . . 1581
19.4.5 Contenu d'un polynôme . . . 1583
19.4.6 Théorème de l'élément primitif . . . 1584
19.4.7 Construction de $\mathbb{F}_{p^n}$ . . . 1588
19.4.8 Exemple : étude de $\mathbb{F}_{16}$ . . . 1590
19.4.9 Polynômes irréductibles sur $\mathbb{F}_q$ . . . 1592
19.4.10 Matrices . . . 1595
19.5 Constructions à la règle et au compas . . . 1596
19.5.1 Quelques constructions . . . 1597
19.5.2 Nombres constructibles . . . 1599
19.5.3 Polygones constructibles . . . 1600

20 Intégration sur des variétés (Vol 3) 1607
20.1 Variétés . . . 1607
20.1.1 Introduction . . . 1607
20.1.2 Définition, carte . . . 1607
20.1.3 Ancienne définition . . . 1608
20.1.4 Espace tangent . . . 1610
20.2 Intégration . . . 1610
20.2.1 Le problème pour une intégration globale . . . 1610
20.2.2 Intégrale sur une carte . . . 1610
20.2.3 Quelques expressions pour l'élément de volume . . . 1613
20.3 Intégrale sur une variété . . . 1614
20.3.1 Mesure sur une carte . . . 1614
20.3.2 Intégrale sur une carte . . . 1615
20.3.3 Exemples . . . 1616

20.3.4 Orientation . . . 1616
20.3.5 Formes différentielles . . . 1618
20.3.6 Intégrale d'une fonction sur une sous-variété . . . 1619
20.4 Longueur, aire, volumes etc. . . . 1620
20.4.1 Quelques aires faciles . . . 1620
20.5 Autres théorèmes de points fixes . . . 1622
20.5.1 Brouwer . . . 1622
20.5.2 Théorème de Schauder . . . 1624
20.5.3 Théorème de Cauchy-Arzella . . . 1625
20.5.4 Théorème de Markov-Kakutani . . . 1626
20.6 Intégrales curvilignes . . . 1627
20.6.1 Chemins de classe C^1 . . . 1627
20.6.2 Intégrer une fonction . . . 1627
20.6.3 Intégrer un champ de vecteurs . . . 1628
20.6.4 Intégrer une forme différentielle sur un chemin . . . 1629
20.6.5 Intégration d'une forme différentielle sur un chemin . . . 1629
20.6.6 Interprétation physique : travail . . . 1631
20.6.7 Intégrer un champ de vecteurs sur un bord en $2D$ . . . 1632
20.6.8 Intégrer une forme différentielle sur un bord en $2D$ . . . 1632
20.6.9 Intégrer une forme différentielle sur un bord en $3D$ . . . 1632
20.6.10 Intégrer un champ de vecteurs sur un bord en $3D$ . . . 1633
20.6.11 Dérivées croisées et forme différentielle exacte . . . 1633
20.7 Surfaces paramétrées . . . 1634
20.7.1 Graphe d'une fonction . . . 1635
20.7.2 Intégrale sur une partie de $\mathbb{R}^m$ . . . 1636
20.8 Intégrales de surface . . . 1637
20.8.1 Intégrale d'un champ de vecteurs . . . 1637
20.9 Intégrales de surface . . . 1638
20.9.1 Aire d'une surface paramétrée . . . 1638
20.9.2 Intégrale d'une fonction sur une surface . . . 1639
20.9.3 Intégrale d'une 2-forme . . . 1640
20.10 Flux d'un champ de vecteurs à travers une surface . . . 1640
20.11 Divergence, Green, Stokes . . . 1643
20.11.1 Théorème de la divergence . . . 1643
20.11.2 Formule de Green . . . 1644
20.11.3 Formule de Stokes . . . 1646
20.12 Résumé des intégrales vues . . . 1647
20.12.1 L'intégrale d'une fonction sur les réels . . . 1647
20.12.2 Intégrale d'une fonction sur un chemin . . . 1647
20.12.3 Intégrale d'une fonction sur une surface . . . 1647
20.12.4 Intégrale d'une fonction sur un volume . . . 1648
20.12.5 Conclusion pour les fonctions . . . 1649
20.12.6 Circulation d'un champ de vecteurs . . . 1649
20.12.7 Flux d'un champ de vecteurs . . . 1649
20.12.8 Conclusion pour les champs de vecteurs . . . 1650
20.12.9 Attention pour les surfaces fermées ! . . . 1650
20.13 Formes différentielles exactes et fermées . . . 1652
20.14 Théorème d'Abel angulaire . . . 1654
20.14.1 Passage à la limite sous le signe intégral . . . 1657
20.14.2 Intégrale en dimension un . . . 1657
20.14.3 Intégrales convergentes . . . 1658
20.14.4 La méthode de Rothstein-Trager . . . 1658
20.15 Rappel sur les intégrales usuelles . . . 1664

20.16 Intégrales le long de chemins . . . 1664
20.16.1 Circulation d'un champ de vecteur . . . 1664
20.17 Circulation d'un champ conservatif . . . 1666
20.18 Intégration de fonction à deux variables . . . 1668
20.18.1 Intégration sur un domaine rectangulaire . . . 1668
20.18.2 Intégration sur un domaine non rectangulaire . . . 1669
20.18.3 Changement de variables . . . 1671
20.19 Les intégrales triples . . . 1671
20.19.1 Volume . . . 1673
20.20 Un petit peu plus formel . . . 1674
20.20.1 Intégration sur un domaine non rectangulaire . . . 1674
20.21 Aire et primitive . . . 1676
20.21.1 Longueur d'arc de courbe . . . 1676
20.21.2 Aire de révolution . . . 1677
20.22 L'aire en dessous d'une courbe . . . 1677
20.23 Propriétés des intégrales . . . 1678
20.24 Techniques d'intégration . . . 1679
20.24.1 Intégration par parties . . . 1680
20.24.2 Changement de variables – pour trouver des primitives . . . 1681
20.24.3 Changement de variables – pour calculer des intégrales . . . 1683
20.24.4 Intégrations des fractions rationnelles réduites . . . 1685
20.24.5 Quelques formules à connaitre . . . 1685
20.24.6 Approximation de ln(2) . . . 1686
20.25 Constructions plus naïves de l'intégrale dans le cas réel . . . 1689
20.25.1 Mesure de Lebesgue, version rapide . . . 1689
20.25.2 Pavés et subdivisions . . . 1690
20.25.3 Intégrale d'une fonction en escalier . . . 1693
20.25.4 Intégrales partielles . . . 1693
20.25.5 Réduction d'une intégrale multiple . . . 1694
20.25.6 Propriétés de l'intégrale . . . 1695
20.25.7 Intégrales multiples, cas général . . . 1696
20.25.8 Réduction d'une intégrale multiple . . . 1697
20.25.9 Intégrales sur des parties de $\mathbb{R}^2$ . . . 1698
20.25.10 Intégrales sur des parties de $\mathbb{R}^3$ . . . 1701
20.25.11 Fonctions et ensembles non bornés . . . 1703
20.25.12 Lemme de Morse . . . 1704
20.26 Autres intégrales sympathiques . . . 1706
20.26.1 Intégrale de Wallis . . . 1706
20.26.2 Formule de Stirling . . . 1710
20.26.3 La fonction sinus cardinal, intégrale de Dirichlet . . . 1713

21 Arcs paramétrés (Vol 3) 1723
21.1 Définitions . . . 1723
21.2 Longueur d'arc . . . 1723
21.3 Abscisse curviligne . . . 1726
21.3.1 Formule intégrale de la longueur . . . 1727
21.4 Suite du chapitre . . . 1733
21.5 Autres exemples . . . 1734
21.6 Élément de longueur . . . 1735
21.6.1 Élément de longueur : cartésiennes . . . 1735
21.6.2 Élément de longueur : polaires (1) . . . 1735
21.6.3 Élément de longueur : polaires (2) . . . 1736
21.6.4 Approximation de la longueur par des cordes . . . 1738
21.7 Arc géométrique . . . 1739

21.7.1 Abscisse curviligne et paramétrage normal . . . 1741
21.7.2 Tangente à une courbe paramétrée . . . 1746
21.8 Un peu de topologie . . . 1747
21.9 Repère de Frenet . . . 1750
21.9.1 Torsion . . . 1752
21.10 Hors des coordonnées normales . . . 1753
21.11 Tracer des courbes paramétriques dans $\mathbb{R}^2$ . . . 1756
21.12 Courbes planes . . . 1757
21.12.1 Angle . . . 1757
21.12.2 Courbure signée . . . 1758
21.12.3 Degré, indice et homotopie . . . 1761
21.13 Courbes fermées planes . . . 1767
21.13.1 Cercle circonscrit . . . 1767
21.13.2 Description locale . . . 1769
21.13.3 Enveloppe convexe . . . 1770
21.13.4 Courbure et convexité . . . 1773
21.13.5 Théorème des quatre sommets . . . 1774
21.13.6 Espace topologique normal . . . 1777
21.13.7 Théorème d'Urysohn . . . 1777
21.13.8 Le théorème de Jordan . . . 1782

22 Géométrie hyperbolique (Vol 3) 1785
22.1 Inversion . . . 1785
22.1.1 Cercles perpendiculaires . . . 1785
22.1.2 Inversion . . . 1786

23 Espaces projectifs (Vol 3) 1791
23.1 Sous espaces projectifs . . . 1791
23.2 Espace projectifs comme « complétés » d'espaces affines . . . 1793
23.3 Théorème de Pappus . . . 1796
23.4 Homographies . . . 1797
23.4.1 Homographies . . . 1797
23.4.2 Le groupe projectif . . . 1798
23.4.3 Repères projectifs . . . 1799
23.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$ . . . 1803
23.4.5 Birapport . . . 1804
23.5 Coordonnées homogènes . . . 1810
23.5.1 Curiosité : matrice de translation . . . 1810
23.5.2 Dualité . . . 1811
23.5.3 Polynômes . . . 1813
23.6 La sphère de Riemann $P_1(\mathbb{C})$ . . . 1814
23.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$ . . . 1815
23.6.2 Homographies . . . 1818
23.6.3 Birapport . . . 1824
23.6.4 Division harmonique . . . 1827
23.6.5 Groupe circulaire . . . 1831
23.6.6 Action du groupe modulaire . . . 1833

24 Analyse vectorielle (Vol 3) 1839
24.1 Le théorème de Green . . . 1839
24.2 Théorème de la divergence dans le plan . . . 1843
24.2.1 La convention de sens de parcours . . . 1843
24.2.2 Théorème de la divergence . . . 1844
24.3 Théorème de Stokes . . . 1844

24.4 Théorème de Gauss . . . 1846
24.5 Coordonnées curvilignes . . . 1848
24.5.1 Base locale . . . 1848
24.5.2 Importance de l'orthogonalité . . . 1848
24.5.3 Coordonnées polaires . . . 1850
24.5.4 Coordonnées cylindriques . . . 1850
24.5.5 Coordonnées sphériques . . . 1851
24.5.6 Gradient en coordonnées curvilignes . . . 1851
24.5.7 Divergence en coordonnées curvilignes . . . 1852
24.5.8 Laplacien en coordonnées curvilignes orthogonales . . . 1855
24.5.9 Rotationnel en coordonnées curvilignes orthogonales . . . 1855
24.6 Les formules . . . 1856
24.6.1 Coordonnées polaires . . . 1856
24.6.2 Coordonnées cylindriques . . . 1856
24.6.3 Coordonnées sphériques . . . 1857

25 Espaces de Hilbert (Vol 3) 1859
25.1 Espaces de Hilbert . . . 1859
25.1.1 Sous-espace vectoriel fermé ? ? ? . . . 1861
25.2 Théorème de la projection . . . 1861
25.3 Systèmes orthogonaux et bases . . . 1864
25.3.1 Orthogonal d'une partie . . . 1864
25.3.2 Dual, théorème de représentation de Riesz . . . 1865
25.3.3 Séparabilité . . . 1867
25.3.4 Base hilbertienne . . . 1869
25.3.5 Décomposition dans une base hilbertienne . . . 1874
25.3.6 Digression sur les normes opérateurs . . . 1878
25.3.7 Applications linéaires et continuité . . . 1879
25.4 Théorème de Kochen-Specker . . . 1881
25.5 Théorème de Lax-Milgram . . . 1882

26 Analyse complexe (Vol 3) 1887
26.1 Fonctions holomorphes . . . 1887
26.1.1 Équations de Cauchy-Riemann . . . 1887
26.1.2 Intégrales sur des chemins fermés . . . 1890
26.1.3 Lacets, indice et homotopie . . . 1892
26.1.4 Théorème de Cauchy et analycité . . . 1893
26.1.5 Théorème de Brouwer en dimension 2 . . . 1896
26.1.6 Principe des zéros isolés . . . 1897
26.1.7 Prolongement de fonctions holomorphes . . . 1899
26.1.8 Théorème de Runge . . . 1900
26.2 Intégrales de fonctions holomorphes . . . 1902
26.2.1 Mesure de Radon . . . 1906
26.3 Conditions équivalentes à l'holomorphie . . . 1909
26.4 Singularités, pôles et méromorphe . . . 1909
26.5 Dénombrement des solutions d'une équation diophantienne . . . 1913
26.6 Fonctions d'Euler . . . 1919
26.6.1 Euler et factorielle . . . 1922
26.7 Exponentielle et logarithme complexe . . . 1922
26.7.1 Propriétés de l'exponentielle . . . 1922
26.7.2 Intégrale de Fresnel . . . 1923
26.7.3 Logarithme complexe . . . 1926
26.8 Théorème de Weierstrass . . . 1935
26.9 Analyse complexe en plusieurs variables . . . 1936

26.9.1 Inverse de fonctions analytiques . . . 1936

27 Analyse fonctionnelle (Vol 4) 1937
27.1 Théorème d'isomorphisme de Banach . . . 1937
27.2 Théorème d'Ascoli . . . 1937
27.3 Espaces de Lebesgue L^p . . . 1942
27.3.1 Généralités . . . 1942
27.3.2 Un peu de convergence de suites . . . 1945
27.3.3 L'espace L^∞ . . . 1947
27.3.4 Quelques identifications . . . 1948
27.3.5 Inégalité de Young, Jensen, Hölder et de Minkowski . . . 1949
27.3.6 Ni inclusions ni inégalités . . . 1958
27.3.7 Complétude . . . 1960
27.3.8 Théorèmes d'approximation . . . 1965
27.3.9 Densité des fonctions infiniment dérivables à support compact . . . 1966
27.3.10 Approximation . . . 1969
27.4 Convolution . . . 1970
27.4.1 Approximation de l'unité . . . 1973
27.4.2 Densité des polynômes trigonométriques . . . 1976
27.5 Espaces L^2, généralités . . . 1977
27.6 L'espace $L^2(\mathbb{R}^d)$ . . . 1979
27.7 L'espace $L^2(S^1)$ . . . 1980
27.7.1 Espace mesuré . . . 1980
27.7.2 Topologie . . . 1981
27.7.3 Système trigonométrique . . . 1986
27.7.4 Convolution . . . 1987
27.7.5 Approximation de l'unité . . . 1990
27.7.6 Base hilbertienne (suite des polynômes trigonométriques) . . . 1994
27.7.7 Convolution, bis . . . 1995
27.8 L'espace $L^2\big([a,b]\big)$ . . . 1995
27.9 Sur $[0,2\pi[$ . . . 1999
27.10 Sur $[-T,T[$ . . . 1999
27.10.1 Le cas dans $[0,2\pi]$ . . . 1999
27.11 Théorème de la projection normale . . . 2001
27.11.1 Espace uniformément convexe . . . 2001
27.11.2 Des inégalités . . . 2004
27.11.3 Inégalités de Clarkson . . . 2010
27.11.4 Uniforme convexité des espaces de Lebesgue . . . 2014
27.11.5 Théorème de la projection normale . . . 2014
27.12 Théorèmes de Hahn-Banach . . . 2016
27.12.1 Applications $\mathbb{R}$-linéaires et $\mathbb{C}$-linéaires . . . 2016
27.12.2 Hahn-Banach . . . 2019
27.12.3 Prolongement de fonctionnelles (dimension finie) . . . 2021
27.12.4 Prolongement de fonctionnelles (dimension infinie) . . . 2024
27.13 Théorème de Tietze . . . 2026
27.14 Dualité, réflexivité et théorème de représentation de Riesz . . . 2029
27.15 Topologie faible . . . 2043
27.15.1 Espace de Banach réflexif . . . 2047
27.15.2 Espaces L^∞ . . . 2048
27.16 Espace de Schwartz . . . 2048
27.16.1 Topologie . . . 2050
27.16.2 Produit de convolution . . . 2052
27.17 Théorème de Montel . . . 2052
27.18 Espaces de Bergman . . . 2054

28 Séries de Fourier (Vol 4) 2059
28.1 Densité des polynômes trigonométriques . . . 2059
28.1.1 Convergence pour les fonctions continues (via Weierstrass) . . . 2059
28.1.2 Convergence pour les fonctions continues (via Fejér) . . . 2059
28.1.3 Densité dans L^p . . . 2063
28.1.4 Suite équirépartie, critère de Weyl . . . 2063
28.2 Fonctions de Dirichlet . . . 2066
28.3 Coefficients et série de Fourier . . . 2067
28.3.1 Le contre-exemple que nous attendions tous . . . 2070
28.3.2 Inégalité isopérimétrique . . . 2073
28.3.3 À propos des coefficients . . . 2074
28.4 Série de Laurent . . . 2076

29 Transformation de Fourier (Vol 4) 2083
29.1 Transformée de Fourier sur $L^1(\mathbb{R}^d)$ . . . 2084
29.1.1 Formule sommatoire de Poisson . . . 2087
29.2 Suite régularisante . . . 2090
29.3 Transformée de Fourier dans l'espace de Schwartz . . . 2091
29.3.1 Quelques transformées de Fourier . . . 2094
29.3.2 Formule d'inversion . . . 2096
29.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$ . . . 2100
29.4.1 Le problème . . . 2100
29.4.2 Extension de $L^1 \cap L^2$ vers L^2 . . . 2101
29.4.3 Une formule de Leibnitz . . . 2103

30 Distributions (Vol 4) 2105
30.1 Dérivée faible . . . 2106
30.1.1 Dérivée partielle au sens faible . . . 2106
30.1.2 Dérivée faible partielle . . . 2108
30.2 Topologie et convergence sur des espaces de fonctions . . . 2109
30.3 Distributions . . . 2112
30.3.1 Multiplication d'une distribution par une fonction . . . 2113
30.3.2 Dérivée de distribution . . . 2113
30.3.3 Ordre et support d'une distribution . . . 2115
30.4 L'espace $C^\infty(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d))$ . . . 2117
30.4.1 Dérivation . . . 2120
30.5 Une équation de distribution . . . 2120
30.6 Localisation, principe de recollement . . . 2121
30.7 Permuter distributions, dérivées et intégrales . . . 2126
30.8 Distributions tempérées . . . 2131
30.8.1 Topologie . . . 2132
30.8.2 Distributions associées à des fonctions . . . 2133
30.8.3 Composition avec une fonction . . . 2133
30.8.4 Transformée de Fourier d'une distribution tempérée . . . 2133
30.8.5 Convolution d'une distribution par une fonction . . . 2134
30.8.6 Approximation de la distribution de Dirac . . . 2134
30.8.7 Peigne de Dirac . . . 2137
30.9 L'espace $C^\infty(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d))$ . . . 2138
30.9.1 Propriétés générales . . . 2139
30.9.2 Dérivation . . . 2141

31 Espaces de Sobolev, équations elliptiques (Vol 4) 2145
31.1 Espaces de Sobolev . . . 2145
31.1.1 Sur un intervalle de $\mathbb{R}$ . . . 2145

31.1.2 Sur un ouvert de $\mathbb{R}^n$ . . . 2150
31.1.3 Espace de Sobolev fractionnaire . . . 2152
31.2 Trace . . . 2154
31.3 Théorème de plongement . . . 2157

32 Équations différentielles ordinaires (Vol 4) 2163
32.1 Équation homogène, solution particulière . . . 2164
32.2 Que faire avec $f(z)dz = g(t)dt$? . . . 2165
32.3 Équations linéaires du premier ordre . . . 2166
32.3.1 Pourquoi la variation des constantes fonctionne toujours ? . . . 2167
32.4 Équations à variables séparées . . . 2168
32.4.1 La méthode rapide . . . 2168
32.4.2 La méthode plus propre . . . 2169
32.4.3 Les théorèmes . . . 2169
32.5 Équations linéaires d'ordre supérieur . . . 2171
32.5.1 Équations et systèmes linéaires à coefficients constants . . . 2171
32.5.2 Si les coefficients ne sont pas constants ? . . . 2172
32.6 Système d'équations linéaires . . . 2172
32.6.1 La magie de l'exponentielle... . . . 2172
32.6.2 ...mais la difficulté . . . 2173
32.6.3 La recette . . . 2173
32.6.4 Système d'équations linéaires avec matrice constante . . . 2174
32.6.5 Système d'équations linéaires avec matrice non constante . . . 2174
32.7 Réduction de l'ordre . . . 2174
32.8 Autour de Cauchy-Lipschitz . . . 2176
32.8.1 Fuite des compacts et explosion en temps fini . . . 2176
32.8.2 Écart entre deux conditions initiales . . . 2178
32.8.3 Flot d'un champ de vecteurs . . . 2179
32.8.4 Stabilité de Lyapunov . . . 2193
32.8.5 Système proies-prédateurs de Lotka-Volterra . . . 2197
32.9 Équation du second ordre . . . 2200
32.9.1 Wronskien . . . 2200
32.9.2 Avec second membre . . . 2201
32.9.3 Équation $y'' + q(t)y = 0$ . . . 2201
32.9.4 Équation de Hill . . . 2203
32.10 Différents types d'équations différentielles . . . 2206
32.10.1 Équation homogène . . . 2206
32.10.2 Équation de Bernoulli . . . 2206
32.10.3 Équation de Riccati . . . 2206
32.10.4 Équation différentielle exacte . . . 2207
32.11 Distributions pour les équations différentielles . . . 2208
32.11.1 Équation de Schrödinger . . . 2208
32.12 Équations différentielles du premier ordre . . . 2211
32.13 Premier ordre, variables séparables . . . 2213
32.14 Équations différentielles linéaires du premier ordre . . . 2216
32.14.1 Méthode de variation de la constante . . . 2218
32.15 Équations différentielles linéaires du second ordre . . . 2219
32.15.1 Équations différentielles linéaires du second ordre homogènes à coefficients constants . . . 2219
32.15.2 Linéaires du second ordre à coefficients constants, non homogènes . . . 2220
32.16 Fonction de Green . . . 2222

33 Équations aux dérivées partielles (Vol 4) 2225
33.1 Symbole principal, équation des caractéristiques . . . 2225

33.2 Méthode des caractéristiques pour l'ordre 1 . . . 2225
33.2.1 Un exemple complet un peu minimal . . . 2226
33.2.2 Un théorème d'existence et d'unicité . . . 2228
33.3 Méthode des caractéristiques pour l'ordre 2 . . . 2231
33.3.1 Principe général . . . 2231
33.3.2 Exemple : l'équation d'onde . . . 2232
33.4 Classification des équations du second ordre . . . 2234
33.4.1 Problème au limite . . . 2235
33.5 Principe du maximum . . . 2236
33.6 Quelques exemples . . . 2239
33.6.1 Un changement de variables . . . 2239

34 Numérique (Vol 4) 2243
34.1 Introduction . . . 2243
34.2 Représentations numériques . . . 2243
34.2.1 Entier relatif en complément à deux (binaire) . . . 2243
34.2.2 Représentation en virgule flottante . . . 2245
34.2.3 Simple précision, IEEE-754 . . . 2245
34.3 Problèmes pour écrire des nombres . . . 2248
34.3.1 Troncature : la base . . . 2248
34.3.2 Troncature : le drift . . . 2249
34.3.3 Quelques bonnes règles . . . 2250
34.3.4 Erreur de "cancellation" . . . 2250
34.3.5 Calcul d'une dérivée . . . 2252
34.3.6 Erreur d'absorption . . . 2252
34.4 Conditionnement et stabilité . . . 2253
34.4.1 Comment choisir et penser le K ? . . . 2256
34.5 Un peu de points fixes . . . 2257
34.5.1 Choix de la fonction à point fixe . . . 2257
34.5.2 Convergence quadratique . . . 2258
34.5.3 Convergence . . . 2260
34.6 Méthode de Newton . . . 2261
34.6.1 « Justification » par la formule par Taylor . . . 2261
34.6.2 « Justification » par points fixes . . . 2262
34.6.3 Convergence de la méthode de Newton . . . 2262
34.6.4 Racine carré par la méthode de Newton . . . 2264
34.6.5 Formalisation de l'algorithme . . . 2264
34.6.6 Caractéristiques . . . 2265
34.6.7 Exemple de la racine carrée . . . 2266
34.6.8 Si multiplicité . . . 2267
34.6.9 Et la dérivée ? . . . 2267
34.6.10 Méthode de Newton : le cas général . . . 2267
34.7 Estimation de l'ordre de convergence . . . 2269
34.8 Autres méthodes . . . 2270
34.8.1 Méthode de Schröder . . . 2270
34.8.2 Halley . . . 2270
34.9 Méthode des sécantes variables . . . 2270
34.9.1 Aitken . . . 2271
34.10 Équations algébrique . . . 2272
34.10.1 Résoudre un système linéaire . . . 2272
34.10.2 Caractéristiques . . . 2272
34.10.3 Définitions . . . 2273
34.11 Équations non linéaires . . . 2273
34.11.1 Méthode de bisection . . . 2275

34.12 Efficacité . . . 2277
34.13 Exemples sous forme d'exercices . . . 2277
34.14 Approximations de fonctions . . . 2281
34.14.1 Critère d'interpolation . . . 2281
34.14.2 Base de Newton . . . 2282
34.14.3 Méthode des minimums quadratiques . . . 2283
34.14.4 Notre espace de Hilbert . . . 2284
34.14.5 Droite de régression . . . 2285
34.15 Conditionnement d'une matrice . . . 2286
34.15.1 Perturbation du vecteur . . . 2288
34.15.2 Perturbation de la matrice . . . 2290
34.16 Système linéaires (généralités) . . . 2291
34.16.1 Les méthodes directes . . . 2291
34.16.2 Méthodes itératives . . . 2292
34.17 Système linéaires (méthodes directes) . . . 2292
34.17.1 Inversion de matrice triangulaire . . . 2292
34.17.2 Transformation gaussienne . . . 2293
34.17.3 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 2294
34.17.4 Méthode de Gauss sans pivot (décomposition LU) . . . 2295
34.17.5 Matrice de permutation élémentaire . . . 2299
34.18 Méthode de Gauss avec pivot partiel (décomposition PLU) . . . 2300
34.18.1 L'idée . . . 2300
34.18.2 Le théorème . . . 2301
34.18.3 D'un point de vue algorithmique . . . 2304
34.18.4 Exemples . . . 2306
34.19 Résolution de systèmes linéaires (suite) . . . 2309
34.19.1 Déterminant . . . 2309
34.19.2 Plusieurs termes indépendants . . . 2309
34.19.3 Cholesky . . . 2310
34.20 Système linéaire (méthodes itératives) . . . 2313
34.20.1 La méthode générale . . . 2314
34.20.2 Jacobi . . . 2314
34.20.3 Gauss-Seidel . . . 2314
34.20.4 Autres . . . 2314
34.21 Indices connectés, matrice irréductible . . . 2314
34.22 Localisation des valeurs propres . . . 2316
34.22.1 Matrices à diagonale dominante . . . 2318
34.22.2 M-matrice . . . 2321

35 Méthode des différences finies (Vol 4) 2325
35.1 Problèmes de dimension un . . . 2325
35.1.1 Un schéma à cinq points . . . 2326
35.1.2 Exemple . . . 2331
35.2 Problèmes de dimension deux . . . 2331
35.2.1 Discrétisation en croix . . . 2332
35.2.2 Discrétisation en carré . . . 2333
35.2.3 Résolution de la discrétisation en croix . . . 2334
35.3 Consistance, convergence . . . 2336
35.3.1 Définitions, mise en place . . . 2336
35.3.2 Exemple . . . 2337
35.3.3 Consistance, stabilité et convergence . . . 2339
35.3.4 Exemple : schéma à cinq points, laplacien en croix . . . 2340
35.4 Autres laplaciens . . . 2341
35.4.1 Travail avec le laplacien à 9 points . . . 2344

36 Variables aléatoires et théorie des probabilités (Vol 4) 2345
36.1 Espace de probabilité . . . 2345
36.2 Variables aléatoires . . . 2346
36.2.1 Indépendance . . . 2346
36.2.2 Lois conjointes et indépendance . . . 2349
36.2.3 Somme et produit de variables aléatoires indépendantes . . . 2350
36.2.4 Espérance . . . 2352
36.2.5 Variance . . . 2353
36.2.6 Covariance . . . 2354
36.2.7 Probabilité conditionnelle : événements . . . 2355
36.2.8 Espérance conditionnelle . . . 2357
36.2.9 Probabilité conditionnelle : tribu . . . 2364
36.2.10 Variables de Rademacher indépendantes . . . 2367
36.2.11 Un petit paradoxe . . . 2368
36.2.12 Inégalité de Jensen . . . 2375
36.2.13 Fonction de répartition . . . 2375
36.2.14 Fonction caractéristique . . . 2376
36.2.15 Fonction génératrice des moments, transformée de Laplace . . . 2377
36.2.16 Loi d'une variable aléatoire . . . 2378
36.2.17 Changement de variables . . . 2379
36.3 Convergence . . . 2379
36.4 Loi des grands nombres, théorème central limite . . . 2385
36.4.1 Loi des grands nombres . . . 2385
36.4.2 Théorème central limite . . . 2387
36.4.3 Marche aléatoire . . . 2390
36.5 Les lois usuelles . . . 2391
36.5.1 Loi de Bernoulli . . . 2391
36.5.2 Loi binomiale . . . 2392
36.5.3 Loi multinomiale . . . 2393
36.5.4 Loi géométrique . . . 2393
36.5.5 Loi de Poisson . . . 2394
36.5.6 Loi exponentielle . . . 2394
36.5.7 Approximation de la binomiale par une Poisson . . . 2397
36.5.8 Loi de Poisson et loi exponentielle . . . 2398
36.5.9 Loi normale . . . 2399
36.5.10 Vecteurs gaussiens . . . 2401
36.5.11 Variable aléatoire de Rademacher . . . 2405
36.5.12 Loi de Student . . . 2407
36.5.13 Indépendance, covariance et variance de somme . . . 2407
36.6 Estimation des grands écarts . . . 2408
36.7 Simulations de réalisations de variables aléatoires . . . 2411
36.7.1 Générateur uniforme . . . 2411
36.7.2 Simulation par inversion . . . 2412
36.7.3 Algorithme de Box-Muller . . . 2413
36.7.4 Méthode du rejet . . . 2414
36.7.5 Simuler une loi géométrique à l'ordinateur . . . 2415
36.7.6 Simuler une loi exponentielle à l'ordinateur . . . 2415
36.7.7 Simuler une loi de Poisson à l'ordinateur . . . 2416
36.8 Sage . . . 2416
36.8.1 Loi exponentielle . . . 2416
36.8.2 Inverser des lois . . . 2417
36.9 Monte-Carlo . . . 2417
36.9.1 Intervalle de confiance . . . 2418

36.10 Résultats qui se démontrent avec des variables aléatoires . . . 2420
36.10.1 Nombres normaux . . . 2420
36.10.2 Théorème de Bernstein . . . 2422

37 Statistiques (Vol 4) 2427
37.1 Notations et hypothèses . . . 2427
37.2 Modèle statistique . . . 2427
37.3 Modèles d'échantillonnages . . . 2430
37.4 Estimation ponctuelle . . . 2433
37.5 Statistiques et estimateurs . . . 2435
37.5.1 Qualité des estimateurs . . . 2435
37.5.2 Méthode des moments . . . 2436
37.5.3 Méthode de substitution . . . 2437
37.5.4 Méthode du maximum de vraisemblance . . . 2438
37.5.5 Exemples sous forme d'exercices . . . 2438
37.5.6 Estimation d'une fonction de répartition . . . 2440
37.5.7 Exemples sous forme d'exercices . . . 2441
37.5.8 Espérance et variance d'un estimateur . . . 2442
37.6 Estimation par intervalle de confiance . . . 2443
37.6.1 Région de confiance . . . 2446
37.6.2 Fonction pivotale . . . 2446
37.6.3 Sondage de proportion . . . 2449
37.7 Estimer une densité lorsqu'on ne sait rien . . . 2450
37.7.1 Distance entre des mesures . . . 2451
37.7.2 Estimateur par fenêtres glissantes . . . 2452
37.8 Test d'hypothèses, prise de décision . . . 2454
37.8.1 Exemple : qualité des pièces d'usine . . . 2454
37.8.2 Exemple : la résistance d'un fil . . . 2454
37.8.3 Vocabulaire et théorie . . . 2455
37.8.4 Risque de première et seconde espèce . . . 2456
37.8.5 Modèle paramétrique de loi gaussienne . . . 2457
37.9 Tests paramétriques . . . 2458
37.10 Tests d'adéquation . . . 2459

38 Chaînes de Markov à temps discret (Vol 4) 2465
38.1 Généralités . . . 2465
38.2 Chaînes de Markov sur un ensemble fini . . . 2468
38.3 Marche aléatoire sur $\mathbb{Z}$ . . . 2470
38.3.1 Chaînes de Markov homogènes . . . 2472
38.3.2 Graphe de transition . . . 2473
38.3.3 Chaîne de Markov définie par récurrence . . . 2474
38.4 Classification des états . . . 2476
38.4.1 Chaînes irréductibles . . . 2478
38.4.2 Nombre de visites . . . 2480
38.5 Mesure invariante . . . 2484
38.6 Convergence vers l'équilibre . . . 2486
38.7 Processus de Galton-Watson . . . 2489

39 Martingales (Vol 4) 2493
39.1 Convergence de martingales . . . 2493
39.2 Temps d'arrêt et martingale terminée . . . 2496
39.3 Décomposition de martingales . . . 2498
39.4 Problème de la ruine du joueur . . . 2500
39.4.1 Le cas où la pièce est truquée . . . 2501

39.4.2 Le cas où la pièce est non truquée . . . 2504
39.4.3 Un petit complément . . . 2506

40 Processus de Poisson (Vol 4) 2509
40.1 Processus de Poisson . . . 2509
40.2 Quelques trucs sur la simulation . . . 2512
40.2.1 Le théorème central limite pour Markov . . . 2513
40.2.2 Feuille 5 . . . 2513
40.2.3 Feuille 6 . . . 2513
40.2.4 Feuille 7 . . . 2514
40.2.5 Simuler des lois conditionnelles . . . 2514

41 Langages (Vol 4) 2515
41.1 Langages . . . 2515
41.1.1 Alphabets et mots . . . 2515
41.1.2 Langage . . . 2516

42 Utilisation dans les autres sciences (Vol 4) 2519
42.1 Démystification du MRUA . . . 2519
42.1.1 Preuve de la formule . . . 2519
42.1.2 Interprétation graphique . . . 2520
42.2 Relativité en mécanique newtonienne . . . 2520
42.2.1 Relativité du mouvement . . . 2520
42.2.2 Bob et Alice . . . 2520
42.3 Invariance de la vitesse de la lumière . . . 2521
42.3.1 Champ de gravitation et électrique . . . 2521
42.3.2 Support du champ : pas d'éther . . . 2521
42.3.3 Le problème . . . 2522
42.4 Conséquences . . . 2522
42.4.1 Ligne d'univers . . . 2522
42.4.2 Transformations de Lorentz . . . 2523
42.4.3 Conditions d'existence . . . 2526
42.4.4 La notion d'intervalle . . . 2527
42.4.5 Le cône de lumière d'un point . . . 2527
42.4.6 Contraction des longueurs . . . 2528
42.4.7 Dilatation des intervalles de temps . . . 2528
42.4.8 Invariance de l'intervalle . . . 2529
42.4.9 Vitesse limite . . . 2534
42.5 Applications . . . 2534
42.5.1 Le GPS . . . 2534
42.5.2 Les ondes électromagnétiques . . . 2534
42.6 Mécanique relativiste . . . 2534
42.6.1 Des problèmes, toujours des problèmes . . . 2535
42.6.2 Loi d'addition des vitesses . . . 2535
42.6.3 L'action d'une force . . . 2536
42.6.4 Équivalence entre la masse et l'énergie . . . 2537
42.7 Principe de correspondance . . . 2537

43 Exemples avec Sage (Vol 4) 2539
43.1 Graphiques . . . 2539
43.1.1 Autres . . . 2539

44 Épilogue : la constante de Weiner (Vol 4) 2559

45 Développements possibles (Vol 4) 2561

45.1 Algèbre et géométrie 2561
45.2 Analyse 2565
45.3 Anciennes leçons 2568

46 GNU Free Documentation License (Vol 4) 2577

Bibliographie 2585

Index

λ-système, 1077
p-Sylow, 347
p-groupe, 347
q-orthogonal, 597

Abel
  angulaire, 1654
  convergence radiale, 1220
abscisse
  curviligne, 1741
absolument continue, 1326
absorbant, 2476
abélianisé, 238
accroissement, 1033
accroissements finie
  dérivée partielle, 911
accélération d'un chemin, 1723
action, 201
  adjointe, 246
  de groupe
    Wedderburn, 1571
  domaine fondamental, 248
  fidèle, 246
  libre, 252
  transitive, 252
action de groupe, 556
  sur des matrices, 1704
action fidèle, 246
action transitive, 252
adjoint, 564
affine
  application, 529
  espace, 525
  sous-espace, 532
aire, 1700
aire dans $\mathbb{R}^2$, 1620
algorithme, 2273
  consistant, 2273
  facteurs invariants, 334
  fortement consistant, 2273
  stable, 2273
algorithme convergent, 2273
algèbre, 195
  de parties, 1071
  engendrée, 195
algèbre de Banach, 491
algèbre engendrée, 1228
algébrique, 394
algébriquement
  indépendant, 433
alignement
  dans un espace projectif, 1792
alphabet, 2515
alterné
  groupe, 359
  polynôme, 429
alternée
  forme linéaire, 549
analytique
  au sens complexe, 1019
angle
  d'une courbe, 1760
  entre deux droites, 1477
  orienté de vecteurs, 1472
angle entre deux vecteurs, 1435
Anneau
  $\mathbb{Z}/n\mathbb{Z}$
    polynôme cyclotomique, 1567
anneau, 113
  $\mathbb{Z}/n\mathbb{Z}$, 357, 1563, 1571, 1574
  de séries formelles, 1288
  euclidien
    facteurs invariants, 334
  factoriel, 274
  noethérien, 280
  principal, 279, 379, 588
    utilisation, 380
  quotient par un idéal, 175
  à division, 371
anneau commutatif, 113
anneau intègre, 176
anneau topologique, 473
application
  de classe C^k, 775
  différentiable, 774, 775, 923, 959, 1344, 1704
    extrémums lié, 1359
  définie positive, 597
  en escalier, 1692

linéaire
théorème de Banach-Steinhaus, 759
mesurable, 1075
multilinéaire, 732
ouverte, 483
semi-définie positive, 597
tangente, 927
application affine, 601
application bilinéaire, 594
application ouverte, 761
application quotient, 441
application réciproque, 478
applications linéaires semblables, 614
approximation
de fonctions
par des polynômes, 2423
de l'unité, 1973
par polynômes, 1315
polynomiale, 1900
apériodique
chaine de Markov, 2488
état d'une chaine de Markov, 2487
arc
géométriques, 1740
paramétré, 1723
arc cosinus, 1433
arc sinus, 1431
arc tangente, 1429
archimédien, 124
associée
subdivision, 1692
associés
éléments d'un anneau, 272
asymptotiquement pivotale, 2446
attractif
point fixe, 1331
automorphisme, 445
d'espace vectoriel, 307
automorphisme de groupes, 112
axiome
du choix, 106

Baire
espace, 522
théorème, 514, 522
tribu, 1088
barycentre
cas affine, 534
cas vectoriel, 1454
enveloppe convexe, 538
base, 298
d'un module, 194
de Newton, 2282
de topologie
dénombrable, 463
espace métrique, 463
duale, 340
espace préhilbertien, 1869
hilbertienne, 1869
utilisation, 2073
locale, 1848
base associée à un repère cartésien, 526
base canonique de $\mathbb{R}^m$, 302
base de topologie, 435
base de topologie et continuité, 445
base de voisinages, 436
base orthonormée, 604
base préduale, 341
Bergman (espace), 2054
Bernoulli, 2391
somme, 2409
Berry-Esséen (borne), 2390
Bessel
inégalité, 1868
biais
d'estimateur, 2436
bien
conditionné, 2254
enchainé, 511
biholomorphe, 1936
bijection, 106, 478
bilinéaire, 732
binormale, 1751
birapport, 1805
birapport dans $\mathbb{C} \cup \{\infty\}$, 1824
birégulier
point sur une courbe, 1738
Bolzano-Weierstrass
espaces métriques, 464
bon
ordre, 108
bord, 461
borné, 463
partie de V, 567
temps d'arrêt, 2496
bornée
différentielle, 946
partie de $\mathbb{R}^m$, 834
suite, 674
borélienne
fonction, 1086
tribu, 1084
boréliens, 1084
boule
avec seminormes, 515
ouverte, 460
boule dans un corps, 204

boule fermée, 462
Bruhat (décomposition), 1063
Burnisde
    formule, 249
Bézout
    anneau principal, 279
    calcul effectif, 258
    nombres entiers, 178
    polynômes, 387

canonique
    base, 302
    décomposition, 112
    espace affine, 526
Cantor
    ensemble, 1131
caractère, 1300
    abélien, 1293
    de S_4, 1307
    groupe diédral, 1552
    irréductible, 1300
caractéristique
    d'un anneau, 196
    d'une équation différentielle, 2225
    polynôme, 590
    sous-groupe, 161
cardinal, 143
cardioïde, 1737
carré
    dans un corps fini, 1575
carrée
    matrice, 314
carte, 1608
catégorie
    ensemble de première, 514
Cauchy
    critère
        uniforme, 977
    déterminant, 558
    formule, 1893
    suite, 500
    théorème, 347
Cauchy-continue, 827
Cauchy-Riemann, 1887
Cauchy-Schwarz, 701, 1860
Cayley
    théorème, 347
cellule d'un pavage, 1692
centrale (application), 1300
centralisateur, 161, 174
centre
    d'un anneau, 174
    d'un groupe, 161
    d'une rotation, 1461
cercle
    circonscrit à une courbe, 1767
    dans la sphère de Riemann, 1816
cercle-droite, 1816
cercles
    perpendiculaires, 1785
Cesàro
    moyenne, 753
chaine, 511
    de Markov, 2465
        apériodique, 2488
        convergence, 2488
        finie, 2468
        homogène, 2465
        irréductible, 2473
        récurrente positive, 2483
        régulière, 2468
champ
    conservatif, 1666
    de vecteurs, 1628
champ de vecteur conservatif, 1631
champ dérivant d'un potentiel, 1666
changement de variable, 1740
Chasles, 525
Chemin
    classe C^2, 1627
chemin, 1627, 1644, 1723
    dans $\mathbb{R}^p$, 1723
chemin de Jordan, 1644
chemin régulier, 1644
circulation, 1664
classe
    d'association, 272
    de conjugaison, 161
classe d'association, 272
classe de conjugaison
    dans S_4 , 171
clôture algébrique, 395
codimension, 303
coefficients
    de Fourier, 1976
coefficients binomiaux, 268
coefficients de Fourier, 1987
coercion, 1882
coercive, 1387
colinéarité, 1791
combinaison
    convexe, 534
combinatoire, 1488
commutant, 651
commutateur
    dans un groupe, 237
compacité, 464, 507, 512, 1904, 1935

sous-groupes du groupe linéaire, 1066
théorème de Dini, 979
utilisation, 1397
théorème de Montel, 2053
compact, 451, 570
arc paramétré, 1723
boule unité, 671
et fonction continue, 465, 681
fermé et borné, 670
intervalle $[a, b]$, 670
le coup du, 1323
localement, 452
opérateur, 1937
produit dénombrable, 513
produit fini, 512
quasi, 451
suite exhaustive, 509
séquentiellement, 452
compacts et fermés, 454
complet
$\mathbb{R}$
corps, 226
espace métrique, 500
corps, 204
espace mesuré, 1090
espace topologique, 490
métrique, 491
complexe conjugué, 235
complète
famille de projecteurs, 194
complément
à deux, 2244
complémentaire, 109
complétion
projective, 1794
complétion d'espace métrique, 1406
complétude, 1403, 1407, 1960
espaces L^p, 1961
complété
d'un espace métrique, 1407
composante, 1034
composante connexe, 450
composition
suite de, 241
concaténation
de langages, 2516
de mots, 2515
concave, 1361
log-concave, 1368
condition initiale, 2213
conditionnement
absolu, 2253
d'une matrice inversible, 2287
relatif asymptotique, 2273
conjugué hermitien, 606
conjugués
éléments d'une extension, 394
connectés
indices d'une matrice, 2314, 2315
connexe par arc, 684
connexité, 512
définition, 449
et intervalles, 679
fonction holomorphe, 1410
indice d'une courbe, 1892
le groupe $\mathrm{GL}^+(n, \mathbb{R})$, 1054
par arc
fonction différentiable, 959
points d'accumulation, 507
prolongement analytique, 1410
signature d'une forme quadratique, 1392
théorème de Runge, 1900
théorème des valeurs intermédiaires, 693
utilisation
Brouwer, 1896
consistance
estimateur, 2435
ordre, 2336
constructible
angle, 1600
point, 1596
réel, 1596
construction
des réels, 216
contenu, 292
continue
fonction entre espaces métriques, 505
fonction entre espaces topologiques, 445
forme différentielle, 917
sur espace métrique, 698
continuité
fonction définie par une intégrale, 1322
séquentielle, 478
contraction, 1332
convergence
commutative, 740
dans un espace vectoriel normé, 672
de martingales, 2497
en loi, 2380
en probabilité, 2380
ordre, 2269
presque surement, 2379
quadratique, 2258
rapidité, 1258, 2088, 2089, 2268, 2409
suite
dans un corps, 204

suite dans $\mathbb{R}^m$, 485
suite numérique, 671, 1315, 2063
Abel angulaire, 1654
uniforme, 977
intégrale, 1320
théorème de Dini, 979
convergence absolue, 735
convergence de suite, 438
convergence forte, 760
convergence normale, 735
convergence uniforme
série de fonctions, 736
convergent
estimateur, 2435
convexe
courbe plane, 1769
fonction sur $\mathbb{R}^n$, 1370
convexité
barycentre, 536
enveloppe de $O(n)$, 1062
fonction, 1361
inégalité de Jensen, 1376
locale, 2020
méthode de Newton, 2264
utilisation, 1397
convolution, 2351, 2491
convolution sur S^1, 1988
coordonnées
cartésiennes
dans un espace affine, 541
curvilignes, 1848
cylindrique, 1538
dans un espace affine, 526
homogène, 1810
sphériques, 1539
coordonnées barycentrique, 542
coordonnées polaires, 1533
corps, 181
complet, 204
de décomposition, 418
de rupture, 408
polynôme cyclotomique, 1567
des fractions, 202, 203
extension, 424, 431
fini, 1574, 1578, 1581
Wedderburn, 1571
formellement réel, 371
ordonné, 204
premier, 373
corps algébriquement clos, 395
corps valué, 233
cosinus, 1413
hyperbolique, 1260
courbe, 1723
efficacité, 2456
fermée, 1757
simple, 1757
étude métrique, 2073
courbe de Jordan, 1644
courbe de niveau, 894, 897
courbe simple, 1644
courbure, 1751
signée, 1758
totale, 1759
covariance, 2354
critique
Galton-Watson, 2491
point, 1355
point d'un arc, 1738
région, 2455
valeur, 2456
critère
Abel, 1210
Abel pour intégrales, 1321
Cauchy
uniforme, 977
de Cauchy, 226, 500
Weierstrass, 1321
série de fonctions, 982
critère du quotient, 748
cycle, 166
cyclique
endomorphisme, 586
matrice, 586
cycloïde
coordonnées normales, 1743
longueur, 1737
Cône de lumière, 2527

degré
application $S^1 \to S^1$, 1761
d'une représentation, 342
extension de corps, 391
degré d'un polynôme, 198
demi-plan, 868
dense, 443
nulle part, 514
densité, 1407
conjointe, 2350
d'une variable aléatoire, 2346
dans un espace de fonction
critère de Weyl, 2063
de $\mathbb{Q}$ dans $\mathbb{R}$, 668
utilisation, 1367
de $\mathrm{GL}(n, \mathbb{R})$ dans $\mathbb{M}(n, \mathbb{R})$, 1055
de $\mathscr{D}(\mathbb{R}^n)$ dans $L^1(\mathbb{R}^n)$, 2051
de $C_c^\infty(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$, 1967

de $L^2([0,1])$ dans $L^p([0,1])$, 1968
de $S^+(n,\mathbb{R})$ dans $S^{++}(n,\mathbb{R})$, 1059
des fonctions étagées dans L^p, 1967
des polynômes
dans $C_c^0[0,1]$, 2423
matrices diagonalisables dans $\mathbb{M}(n,\mathbb{C})$, 1055
mesure, 1152
points extrémaux dans $\mathcal{L}$, 1061
prolongement, 1403
densité d'une mesure, 1152
diagonale
dominante, 2318
fortement dominante, 2319
strictement dominante, 2319
diagonalisable, 618
et polynôme minimum scindé, 618
exponentielle, 1265
diagonalisation
cas complexe, 845
cas réel, 622
endomorphisme autoadjoint, 704
simultanée, 620
diamètre, 834
difféomorphisme, 775
de classe C^k, 969
différence
centrée, 2325
divisée, 2282
progressive, 2325
régressive, 2325
différentiabilité, 959
différentielle, 774
de $u \mapsto u^{-1}$, 791
partielle, 790
totale, 1033
dilatation, 1048
dilatation (matrice), 330
dimension, 302
n-formes multilinéaires alternées, 549
définition, 302
sous espace affine, 532
utilisation, 538
direction, 909, 1746
sous-espace affine, 532
Dirichlet
noyau, 2060
théorème, 2060
théorème (sur les nombres premiers), 1571
disque de convergence, 1209
distance, 460
associée à une norme, 469
compatible, 494
entre deux mesures de probabilités, 2451
invariante, 495
point et ensemble, 487
distance discrète, 487
distance produit, 481
distingué
sous-groupe, 161
distribution, 2112
de Dirac, 2131
produit par une fonction, 2113
tempérée, 2131
équation de Schrödinger, 2208
divergence, 1040
diviseur
dans un anneau, 176
de zéro, 176
de zéro à droite, 176
polynôme, 291
diviseur de zéro, 176
divisible, 130
division
euclidienne, 177, 291
harmonique, 1827
diédral, 1488
domaine, 665
fondamental d'une action, 248
dominé
modèle statistique, 2434
dominée
convergence (Lebesgue), 1149
mesure, 1157
droite
dans la sphère de Riemann, 1815
projective, 1791
droite affine, 543, 861
droite réelle achevée, 688
droite vectorielle, 861
droites parallèles, 861
droites perpendiculaires, 861
dual
d'un espace de Hilbert, 1865
de $\mathbb{M}(n,\mathbb{K})$, 612
de $L^p(\Omega)$, 2039
de L^p, 2034
dual algébrique, 340
dual topologique, 794, 1859
Dunford
décomposition, 638
dyadique, 1168
développement, 223
décalage, 2245
décomposition
Bruhat, 1063
canonique, 112

corps, 418
Dunford, 638
application, 1265
exponentielle de matrice, 1005
Jordan
et exponentielle de matrice, 1264
polaire, 1060
primaire, 637
sous-espaces caractéristiques, 637
spectrale, 637
décomposition décimale, 756
décomposition polaire, 1439
dénombrable, 145
à l'infini, 452
dénombrement, 1488
partitions de $\{1, \ldots, n\}$, 1288
déplacement, 1504
dérivabilité
fonction définie par une intégrale, 1323
lemme de Borel, 1286
dérivable, 874
au sens complexe, 955
fonction, 2163
dérivation
au sens des distribution
Sobolev, 2148
dérivé
groupe, 237
dérivée
dans Sobolev $H^1(I)$, 2145
directionnelle, 908, 912
distributionnelle, 2113
faible, 2108
fonction à valeurs dans E', 522
partielle, 908, 1033
seconde, 875
dérivée directionnelle, 909
dérivée faible, 2106
dérivée partielle, 908
déterminant, 549
Cauchy, 1318
d'un endomorphisme, 553
d'une famille de vecteurs, 550
de Cauchy, 558
et inversibilité, 553
forme linéaire alternée, 549
Gram, 557, 1318
interprétation géométrique, 1186
matrice, 319
résultant, 559, 1659
utilisation, 1397
Vandermonde, 555
détermination
logarithme, 1932
principale, 1932
développable
en série entière, 1268
développement
asymptotique, 1031
limité
en zéro, 1025
fonction holomorphe, 1887
premier ordre, 878
Taylor, 1704

écart-type, 2354
échantillon, 2428, 2430
effectif
empirique, 2460
efficacité
courbe, 2456
d'une méthode itérative, 2277
élément
de surface, 1639
de torsion, 254
élément de surface, 1614
ellipsoïde, 662
elliptique
équation aux dérivées partielles, 2234
endomorphisme, 307
cyclique, 586
diagonalisable, 661, 1057, 2204
Dunford, 638
diagonalisation, 622
décomposition
polaire, 1060
nilpotent
Dunford, 638
préservant une forme quadratique, 1068
sous-espace stable, 638, 2204
endomorphisme direct, 563
endomorphisme préserve l'orientation, 563
engendré, 269
λ-système, 1077
corps, extension, 400
idéal dans un anneau, 269
sous-espace affine, 532
sous-groupe, 189
tribu
par une variable aléatoire, 2347
ensemble
archimédien, 124
de Cantor, 1131
différence symétrique, 110
infini, 140
ensemble connexe, 449
ensemble des mots, 2515

ensemble fini, 140
ensemble ordonné, 107
ensemble quotient, 111
ensembles
  disjoints, 105
entier, 136
entrelacement, 1300
enveloppe
  convexe, 536
équation
  différentielle
    linéaire, 2166
    ordinaire d'ordre 1, 2163
    variables séparées, 2168
  générale de degré n, 433
équi-intégrable, 2497
équivalence
  arcs paramétrés, 1739
  chemin, 1732
  classe de fonctions, 1942
  de représentations, 1300
  de suites, 673
  homotopie, 1892
  suite de composition, 243
erreur, 2283, 2313
  assignation, 2248
  de consistance, 2336
  discrétisation, 2337
  quadratique, 2286
  troncature, 2248
erreur relative, 2248
escalier, 1114
Espace
  de Sobolev, 2150
espace
  L^2
    Sobolev, 2148
  L^p, 1945
  de Baire, 522
  de Bergman, 2054
  de fonctions
    L^p, 1961
    Sobolev H^1, 2148
  de Hilbert
    espace de Sobolev H^1, 2148
  de probabilité, 2345
  de Schwartz, 2048, 2131
  de Sobolev, 2145, 2152
  euclidien, 716
  mesurable, 1069
  mesuré, 1073
    complété, 1092
  métrique, 460
    base de topologie, 463
  projectif, 1791
  propre, 580
  séparé, 447
  tangent, 1610
  topologique
    métrisable, 504
  vectoriel, 192
    dimension, 549
  vectoriel topologique
    métrisable, 489
espace de Banach, 491
espace topologique, 435
espace topologique normal, 1777
espace vectoriel
  topologique, 470
espace vectoriel normé, 467
espérance, 2352
  conditionnelle, 2357, 2359
    variable aléatoire, 2363
    événement, 2359
estimateur, 2435
  biais, 2436
  consistant, 2435
  convergent, 2435
  de fonction de répartition, 2440
  maximum de vraisemblance, 2438
estimation
  des grands écarts, 2409
étrangers
  polynômes, 291
Euclide
  algorithme étendu, 256
  lemme, 259
euclidien
  anneau, 283
  espace, 605
Euler
  indicatrice, 1560
événement, 2345
exact
  intervalle de confiance, 2444
excès
  intervalle de confiance, 2444
exhaustive (suite de compacts), 509
exponentielle, 1244
  convergence, 751
  de matrice, 793, 1264, 1265, 1277
    utilisation, 1353
  existence, 1238
  rapide, 1565
  unicité, 1013
exposant, 354, 661

d'un groupe, 164
extension
corps de base, 640
de corps, 391, 431
algébrique, 404
finie, 1584
simple, 400
utilisation, 1601
isométrie, 1404
séparable, 425
extension algébrique, 395
extension algébriquement clos, 395
extrapolation, 2282
extrémité
d'un intervalle, 678
extrémum, 1704
lié, 1359
local
relatif, 1359
extrémum local, 1354
extrémums, 1357
volume d'un ellipsoïde, 1397

facteur
intégrant, 2207
factoriel
anneau, 274
factorisation
de polynôme, 293, 407
faisceau de droites, 1813
famille
sommable, 741
famille trigonométrique
sur S^1, 1986
Fatou, 1140
Fejér
noyau, 2060
fermeture séquentielle, 488
fermé, 436, 570
filtration, 2493
fine
subdivision, 1724
fixateur, 246
flot, 2180, 2225
flux
d'un champ de vecteur, 1643
flux d'un champ de vecteurs, 1641
fonction, 665
Γ d'Euler, 1919
borélienne, 1086
caractéristique, 1694
d'une variable aléatoire, 2376
continue
par morceaux, 1117
égales, 489
convexe, 1361, 1367
croissante, 666
de Dirichlet, 2066
de Möbius, 1592
de répartition, 2375
décroissante, 666
définie par une intégrale, 1282, 1321, 1327, 1328, 1935
Γ d'Euler, 1919
utilisation, 2396
en escalier intégrable, 1693
génératrice, 2377
holomorphe, 955, 1935
théorème de Montel, 2053
image, 665
monotone, 666
par morceaux, 1117
méromorphe
Γ d'Euler, 1919
valeurs vectorielles, 1034
à décroissance rapide, 2049
étagée, 1966
fonction continue en un point, 445
fonction de classe C^1, 946
fonction dérivée, 874
fonction périodique, 885
fonctionnelle
énergie, 1886
fonctions équivalentes, 448
fondamental
domaine d'une action, 248
forme
bilinéaire, 594
non dégénérée, 595
différentielle, 917
exacte, 1652
fermée, 1652
linéaire
différentielle, 1359
quadratique, 1392, 1394, 1704
groupe orthogonal, 1068
forme bilinéaire symétrique, 594
forme canonique
fonction simple, 1114
forme linéaire, 340
alternée, 549
forme quadratique, 597
formule
Bayes, 2356
Burnside, 249
d'expulsion (produit vectoriel), 714
de Cauchy, 1893

Hadamard, 1213
inversion Möbius, 1593
probabilité totales, 2356
sommatoire de Poisson, 2088
Taylor
reste intégral, 1688
utilisation, 2268
Formule de Leibnitz, 2108
formule de Stirling, 1711
Fourier, 2088
série
utilisation, 2073
transformée
groupe abélien fini, 1295
fraction
rationnelle
intégration, 1659
fraction dyadique, 223, 1780
fraction rationnelle, 1685
fractions
rationnelles, 202
fractions (corps), 203
Fredholm
équation, 1334
Frenet
formules, 1752
Fresnel
intégrale, 1923
Frobenius
morphisme, 197
frontière, 443, 461, 574
fréquence
empirique, 2460, 2480
Fubini
théorème
dans $\mathbb{R}^n$, 1201

Galton-Watson
sous-critique, 2491
sur-critique, 2491
Gauss
lemme
polynômes, 388
somme de, 1576
Gershgorin
disque, 2316
gradient, 925, 932, 945
Gram (déterminant), 557
Gram-Schmidt, 716
graphe, 665, 876
de transition (chaine de Markov), 2473
fonction, 894
fonction de deux variables, 896
groupe, 112
$\mathrm{GL}(n, \mathbb{R})$, 1394
p-groupe, 347
action, 1834
utilisation, 1068
agissant sur un ensemble
diédral, 1481
alterné, 359
circulaire, 1831
de Galois, 433
de permutation, 1552
de permutations, 1488
caractères de S_4, 1307
de torsion, 254
des isométries
espace métrique, 487
des symétries, 1501
diédral, 1481, 1488
générateurs (preuve), 1482
générateurs (utilisation), 1552
dérivé, 237
de $\mathrm{GL}(n, \mathbb{K})$, 1053
de $\mathrm{SL}(n, \mathbb{K})$, 1053
du groupe alterné, 362
du groupe symétrique, 361
en géométrie, 1481
et géométrie, 549, 1488, 1834
isométries du cube, 368
fini, 349, 357, 1488, 1563, 1571, 1574
alterné, 361
diédral, 1481
Wedderburn, 1571
linéaire, 1063
décomposition polaire, 1060
enveloppe convexe de $\Omega(n)$, 1062
hyperplan, 613
sous-groupes compacts, 1066
modulaire, 1833
orthogonal, 566
d'une forme quadratique, 1068
partie génératrice, 361, 1563, 1834
permutation, 549, 556, 1063, 1563
diédral, 1481
projectif, 1798
quotient, 241
résoluble, 244
simple, 162
spécial orthogonal, 567
symétrique, 165
action sur un triangle, 1297
groupe abélien, 112
groupe commutatif, 112
groupe cyclique, 190
groupe de pavage, 1504

groupe dérivé
de $\mathrm{GL}(n, \mathbb{C})$, 1047
groupe ordonné, 161
groupe simple, 162
Grönwall (lemme), 2163, 2164
générateur, 190
génératrice
partie d'un module, 194
géométrie
avec des groupes, 1482, 1834
avec nombres complexes, 1482, 1834
géométrique
avec des nombres complexes, 2073

Hadamard
conditions, 2236
formule, 1213
Hardy-Littlewood (théorème), 1315
Hausdorff, 447
Heine (théorème), 835
hermitienne, 606
hessienne, 971
Hilbert
espace, 1859
holomorphe, 955, 1936
sur un compact, 955
homographie, 1797, 1834
sur $\mathbb{C} \cup \{\infty\}$, 1820
homogène
chaine de Markov, 2465
homotopie, 1763
homéomorphisme, 445
hyperbolique
équation aux dérivées partielles, 2234
hyperplan, 657, 1449
de $\mathbb{M}(n, \mathbb{K})$, 614
sépare
au sens strict, 2020
séparer
au sens large, 2020
hypothèse
alternative, 2455
composite, 2455
multiple, 2455
nulle, 2455
simple, 2455

identifiable, 2434
identité de polarisation, 597
idéal
bilatère, 174
dans un anneau, 174
maximal, 271
principal
à droite, 275
à gauche, 275
idéal à droite, 174
image, 665
incompressible
champ de vecteur, 1042
indicatrice d'Euler, 1560
indice, 240
d'une courbe dans $\mathbb{C}$, 1892
de rotation, 1762
indice d'inertie, 602
inductif, 108
induite
topologie, 571
tribu, 1070
indécomposable
module, 195
indépendance, 2349
affine, 540
algébrique, 433
projective, 1799
sous tribus, 2346
variables aléatoires, 2347
événements, 2346
utilisation, 2421, 2500
infimum, 228
inférence statistique, 2427
injection, 106, 478
interpolation, 2282
Intervalle, 2527
intervalle, 108, 665, 684
longueur, 1120
intervalle de confiance
asymptotique, 2448
intégrable, 1143
fonction non en escalier, 1699
fonction positive, 1703
fonction à valeurs vectorielles, 1146
intégrale
calcul, 2063
convergente, 1183, 1658
d'une fonction sur une carte, 1615
d'une fonction sur une variété, 1619
d'une forme différentielle, 1630
fonction en escalier, 1693
Fresnel, 1923
impropre, 1182, 1183
sur un chemin, 1627
intégrale d'une fonction, 1134
intégrale d'une forme différentielle, 1629
intégrale de Dirichlet, 1721
intégrale sur une carte, 1611
intégration

fraction rationnelle, 1659
intéressante, 101
intérieur, 441
d'un ensemble, 569
point, 569
invariance cyclique
trace, 315
invariant
de similitude, 648
invariante
mesure
pour une chaine de Markov, 2484
inverse
dans un groupe, 160
inverse généralisé, 2412
inversion, 1786
dans le groupe symétrique, 172
inversion dans $\mathbb{C} \cup \{\infty\}$, 1818
involution, 620
inégalité
arithmético-géométrique, 1377
Bessel, 1868
Cauchy-Schwarz, 701, 1860
de Khintchine, 2405
de la moyenne, 959
des pentes, 1362
Hölder, 1950
utilisation, 2353, 2423
isopérimétrique, 2073
Jensen, 1376
espérance conditionnelle, 2375
pour une somme, 1376
version intégrale, 1949
Kantorovitch, 1377
Markov, 2385
Minkowski, 1953
triangulaire, 460, 467
irrationalité
$\sqrt{2}$, 211
irréductible
chaine de Markov, 2473
dans un anneau, 273
module, 195
représentation, 1298
isobarycentre, 534
isomorphisme, 765
$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, 1574
d'anneaux, 173
de corps, 182
espace affine, 531
espace vectoriel normé, 777
espaces vectoriels, 307
isomorphisme d'espace topologique, 445
isométrie
de forme quadratique, 599
de l'espace euclidien $\mathbb{R}^2$, 1482
espace euclidien
isométries du cube, 368
groupe, 487
isométrie (forme bilinéaire), 599
isométrie d'espaces métriques, 486
isométrie positive, 1453
isotrope
cône, 596
totalement, 596
isotrope (vecteur), 596

jacobien, 945, 1039
Jordan
réduction, 650
Jordan-Hölder, 242

Kronecker
symbole, 302

lacet, 1644, 1892
lacet de Jordan, 1644
Lagrange
multiplicateur, 1359
polynôme, 612
théorème, 241
lagrangien, 1359
langage, 2516
itéré, 2517
itéré strict, 2517
unité, 2516
vide, 2516
Laplace
transformée, 2377
laplacien, 1855
Legendre
symbole, 1575
Leibnitz, 879
applications entre espaces vectoriels normés, 789
lemme
Borel, 1286
d'Euclide, 260
de Borel-Cantelli, 2383
de Gauss
contenu de polynôme, 1583
pour des entiers, 259
de Morse, 1704
de Schreider, 243
de Slutsky, 2382
de transport, 1083
de Zorn, 108

des noyaux, 581
Fatou, 1140
Gauss
dans un anneau principal, 279
polynômes, 388
Grönwall, 2163, 2164
Hadamard, 1330
regroupement, 2349
Schur complexe, 845
Schur réel, 621
lettres, 2515
Levi-Civita, 1849
libre
action, 252
partie, 297
partie d'un module, 194
librement engendré, 305
Ligne d'univers, 2522
limite, 634
d'ensembles, 1076
d'une fonction, 458
de fonctions holomorphes, 1935
de suite
espace topologique, 438
inférieure, 676
inversion, 1288, 1315, 1935, 1945
permutation
utilisation, 2377
suite dans $\mathbb{R}^m$, 485
suite numérique, 671
supérieure, 676
unicité, 459
limite inférieure, 676
limite pointée, 807
limite à droite, 819
limite épointée, 807
linéaire
application, 306
linéairement indépendant
module, 192
Lipschitz, 1332
localement, 960
Lipschitzienne, 960
liste d'indices, 35
localement
intégrable, 1182
log-concave, 1368
logarithme
complexe, 1928
dans $\mathbb{C}$, 1926
de matrice, 1278
sur les réels positifs, 1240
loi
χ^2, 2407
binomiale
comportement asymptotique, 2409
conjointe, 2349
d'une variable aléatoire, 2378
de Poisson, 2394
des grands nombres
forte, 2386
pour les chaines de Markov, 2486
processus de Poisson, 2510
utilisation, 2409, 2421
marginale, 2349
normale
vecteur gaussien, 2401
parente, 2427, 2428
parente d'un échantillon, 2430
réciprocité quadratique, 1578
sans mémoire, 2396
Student, 2407
longueur
arc géométrique, 1740
d'un arc paramétré compact, 1724
d'un intervalle, 1120
d'une arrête, 1690
élément de, 1735
longueur d'arc, 1726
Lotka-Volterra, 2197

M-matrice, 2321
maigre, 1088
maigre (ensemble), 514
majorant, 107, 227
essentiel, 1947
Markov
inégalité, 2385, 2433
martingale, 2493
bornée dans $L^2(\Omega)$, 2494
matrice, 313, 1063, 1834
compagnon, 646
creuse, 2291
cyclique, 586
d'une application linéaire, 316
de dilatation, 330
de permutation, 330
de similitude, 955
de Sylvester, 558
de transvection, 330
dense, 2291
hermitienne
racine carrée, 1057
jacobienne, 932, 945
normale, 845
orthogonale, 329, 566
permutation

élémentaire, 2299
racine carrée, 1057
réductible, 2315
semblable, 1057
semblables, 1394
stochastique, 2467
symétrique, 1394
réelle, 1392
trigonalisable, 844
équivalence, 332
dans le groupe linéaire, 1068
matrice d'une forme bilinéaire, 597
matrice de transition, 2466
matrice-colonne, 314
matrice-ligne, 314
matrices
similitude, 332
matrices semblables, 614
matrices équivalentes, 614
maximale
partie orthonormale, 1870
maximum, 107, 122
global, 1354
local, 1354
maximum local, 1354
mesurable
application, 1075
au sens de m^*, 1081
ensemble, 1074
fonction, 1082
Lebesgue, 1689
mesure
σ-finie, 1073
absolument continue, 1157
angle entre vecteurs, 1473
complexe, 1158
dans une carte, 1614
de Borel, 1104
de comptage, 1175, 1200
de Lebesgue, 1124
de Radon, 1104, 1907
externe, 1689
extérieure, 1071
finie, 1073
image, 1101
positive, 1073
probabilité, 2345
produit, 1165
régulière, 1104
extérieure, 1104
intérieure, 1104
sur un ensemble de parties, 1071
mesure σ-finie, 1071
mesure finie, 1071
mesure à densité, 1152
minimum, 122
ensemble ordonné, 107
minimum local, 1354
minorant, 107, 227
modulaire (groupe), 1833
module
de continuité, 796
indécomposable, 195
irréductible, 195
simple, 195
à gauche, 191
module d'un nombre complexe, 694
module produit, 192
modèle
paramétrique, 2428
statistique, 2427
échantillonnage, 2428, 2430
moment, 2352
fonction génératrice, 2377
monogène, 190
extension de corps, 400
monotonie, 2202
monoïde, 378
monôme, 290
morceau
fonction continue ou monotone, 1117
morphisme
d'algèbres, 195
d'anneaux, 113
de corps, 182
espace vectoriel normé, 777
Frobenius, 197
morphisme de groupes, 112
morphisme de modules, 191
morphisme de produits tensoriels, 765
mot, 2515
longueur d'un mot, 2515
mot vide, 2515
nombre d'occurrences, 2515
mot non vide, 2515
moyenne
de Cesàro, 753
empirique, 2354
empirique d'un échantillon, 2430
quadratique, 2354
multiindice, 35
multiplicateur
de Lagrange, 1359
multiplication, 124
multiplicité
racine d'un polynôme, 293

racine de $f(x) = 0$, 2261
valeur propre
algébrique, 843
géométrique, 843
méromorphe, 1912
Méthode
de Newton, 2262
méthode
des chemins, 906
Newton, 2268
cas convexe, 2264
métrisable
espace vectoriel topologique, 489

nabla, 941
neutre
dans un groupe, 160
Newton
méthode, 2268
nilpotent, 176
niveau de confiance, 2444
nombre
complexe
norme 1, 1571
de Fermat, 1601
dénormalisé, 2246
normal, 2420
normalisée, 2246
premier, 357, 379, 1258, 1563, 1571, 1574
dans leur ensemble, 178
deux nombres entre eux, 178
théorème des deux carrés, 380
tours d'une courbe plane, 1761
nombre premier
polynôme cyclotomique, 1567
nombres complexes, 234
non dénombrable, 145
normal
arc paramétré, 1741
endomorphisme, 845
nombre, 2420
sous-groupe, 161
normal extérieur
vecteur, 1643
normale
loi réduite, 2400
principale, 1751
normalisateur, 161
norme, 467
d'algèbre, 726
d'application linéaire, 725
d'une application linéaire, 725
euclidienne
dans $\mathbb{R}^m$, 705
supremum, 719
sur $\mathbb{Z}[i\sqrt{5}]$, 280
norme maximum, 481
norme quotient, 519
norme suprémum, 977
normes équivalentes, 721
normé
espace vectoriel, 469
noyau
application vers un groupe, 238
d'une forme bilinéaire, 596
Dirichlet, 2060
Fejér, 2060
vers un espace vectoriel, 307
nulle part dense, 514
négatif, 219
négligeable
partie d'un espace mesuré, 1090

observation, 2345
opposés
chemins, 1732
opérateur
défini positif, 623
opérateur hermitien, 607
opérateur unitaire, 607
opérateurs
compatibles, 1881
orbite
d'un point sous une action, 246
orbite d'un élément sous une permutation, 165
ordre, 106
bon ordre, 108
d'un polynôme, 292
d'un élément, 163
d'une matrice carrée, 314
dans un corps, 204
de convergence d'un schéma, 2337
distribution, 2115
partiel, 107
sur un anneau factoriel, 274
total, 107
ordre d'un groupe, 163
orientation, 562, 1616, 1637
orientation affine, 564
origine
abscisse curviligne, 1741
repère affine, 541
orthogonal, 604, 1864
coordonnées curviligne, 1848
famille de projecteurs, 194
matrice, 566
opérateur, 566
sous-espace, 342

orthonormé, 705
    système, 1867
oscillation
    d'une fonction, 830
    d'une fonction en un point, 830
osculateur (cercle), 1756
ouvert, 435, 570

parabolique
    équation aux dérivées partielles, 2234
parallèle
    sous-espaces affines, 532
parallélogramme, 1621
paramétrage, 1732
    normale, 1741
paramétrages
    admissible, 1740
Parseval, 1874
partie
    génératrice, 190, 297
    régulière, 1025
    totale, 1867
partie convexe, 467
partie entière, 233
partie fractionnaire, 233
partie libre
    module, 192
partie linéaire, 529
partie symétrique, 473
partie équilibrée, 473
partition
    dénombrable mesurable, 1109
partition de l'unité, 1285
passage aux classes, 441
pavable, 1690
pavage du plan, 1504
pavé, 1689, 1690
Pearson
    théorème, 2460
peigne de Dirac, 2137
permutation, 165
    matrice, 330
permutation impaire, 171
permutation paire, 171
permuter
    différentielle et intégrale
        $\mathbb{R}^n$, 1329
    dérivée et intégrale
        $\mathbb{R}^n$, 1328
        dans $\mathbb{R}$, 1324
        dans $\mathbb{R}$ avec les bornes, 1327
    dérivée et limite, 986
    intégrale
        et série, 1200
    limite et intégrale
        convergence dominée, 1149
        convergence monotone, 1139
        espace mesuré, 1322, 1323
    somme et intégrale, 1140, 1203
    série entière et dérivation, 1223
    série entière et intégration, 1225
permuter limite et intégrale, 1321
petit théorème de Fermat, 374
pgcd
    calcul effectif, 257
    dans un anneau intègre, 175
    polynômes, 390
pivotale, 2446
plan
    projectif, 1791
    tangent, 944, 1034
plan affine, 861
plan médiateur, 865
plan vectoriel, 861
Plancherel, 1874
plongement, 1403
Poincaré (demi-plan), 1833
point
    d'équilibre
        stable, 2193
    pondéré, 534
point adhérent, 442
point critique
    définition, 1706
point d'accumulation, 445
point d'équilibre, 2193
point extrémal dans un convexe, 1061
point fixe, 2491
    attractif, 1331
    Brouwer, 1622
    Picard, 1332
    Schauder, 1624
point intérieur, 441
point isolé, 445
Poisson
    formule sommatoire, 2088
    processus, 2509
polydisque, 1936
polygone, 1480
polygone convexe, 1480
polynôme
    alterné, 429
    annulateur, 581, 585
    caractéristique, 590, 850
    contenu, 292
    cyclotomique, 1565
        irréductibilité, 1567

propriétés, 1566
d'endomorphisme, 1057
décomposition de Dunford, 638
de Bernstein, 2422
irréductible
sur $\mathbb{F}_q$, 1593
séparable, 423
minimal, 396
d'un élément d'une extension, 392
ponctuel, 585
racines, 431
scindé, 385
semi-symétrique, 429
symétrique, 429, 431, 556, 561, 1581
élémentaire, 429, 432
séparable, 423
trigonométrique, 1976
à plusieurs indéterminées, 561, 1581
polynôme de Lagrange, 612
polynôme de plusieurs variables, 271
polynôme de Taylor, 1019
polynôme scindé, 385
polynôme trigonométrique, 1986
polynômes, 198
portée
mesure, 1157
positif, 219
potentiel, 1634, 1666
ppcm
dans un anneau intègre, 175
premier
corps, 373
deux polynômes entre eux, 291
deux éléments d'un anneau principal, 278
idéal, 275
premier temps d'atteinte, 2476
premier type
région solide, 1701
presque
nulle, 254
partout, 1075
surjective, 2026
primitif
polynôme, 292
au sens du pgcd, 292
racine, 1587
triplet pythagoricien, 287
élément d'un corps, 1574
élément d'une extension de corps, 400
primitive, 873, 1664
de fonction continue, 1011
et intégrale, 1184
fonction, 893
primitive et intégrale, 1178
principal
anneau, 275
idéal, 275
principe
prolongement analytique, 1410
zéros isolés, 1409
Principe de correspondance, 2538
probabilité
conditionnelle, 2355
problème
aux limites d'évolution, 2235
aux limites stationnaires, 2235
bien posé, 2235
limite de Dirichlet, 2235
limite de Von Neumann, 2235
problème de Cauchy, 2213
processus
adapté à une filtration, 2493
arrêté, 2498
croissant prévisible, 2498
Galton-Watson, 2489
Poisson, 2509
sans mémoire, 2396
produit, 124
d'une mesure par une fonction, 1152
de convolution, 1970
et Fourier, 2085
de langages, 2516
de mots, 2515
distribution et fonction, 2113
espaces mesurés, 1167
espaces topologiques, 438
mixte, 712, 715
scalaire
en général, 604
sur $\mathbb{M}(n, \mathbb{R})$, 857
semi-direct, 252
tensoriel
de représentations, 1306
vectoriel, 711
produit de Cauchy, 1216, 1218
produit hermitien, 606
produit pseudo-scalaire, 604
produit remarquable, 1901
produit scalaire sur $\mathbb{R}^n$, 605
produit tensoriel, 765, 766
projecteur
dans un module, 194
projectif
complétion, 1794
droite, 1791, 1792
espace, 1791

groupe, 1798
hyperplan, 1792
plan, 1791
repère, 1799
sous-espace, 1791
projection
orthogonale, 1863
projection normale, 2002
projection orthogonale, 1456
prolongement
analytique, 1410
utilisation, 2053
de fonctions, 1403
lemme de Borel, 1286
méromorphe de la fonction Γ, 1919
par continuité, 822
dans $H^1(I)$, 2148
par densité, 1403
théorème de Hahn, 1099
prolongement de fonctionnelle linéaire, 2024
propriété d'intersection finie, 458
propriété d'intersection non vide, 453
propriété universelle, 765
prébase, 437
précision
simple, 2245, 2246
préhilbertien, 1859
préserve l'orientation, 564, 1453
puissance
d'un langage, 2517
d'un mot, 2516
d'un point, 1785
d'un test, 2456
d'une inversion, 1818
période, 885
pôle, 1523, 1909

quasi-compact, 451
quaternion, 428
queue de liste d'indices, 35
quotient, 177, 291
dans une suite de composition, 242
de groupe, 241
de groupes, 356
quotient d'un chemin, 1782
quotient d'un espace vectoriel, 764

racine
carré
de matrice hermitienne, 1057
carré de matrice
hermitienne positive, 1057
d'un polynôme, 292
de l'unité, 661, 1482, 1567, 1571, 1834
primitive, 1559
utilisation, 1566
de polynôme, 407
multiple, 2261
primitive, 1587
simple, 2261
racine carrée, 693
racine de l'unité, 1557
raffinement, 1724
subdivision d'un pavé, 1691
rang, 309, 549, 631
classe d'équivalence, 333
diagonalisation, 622
différentielle, 1359
utilisation, 1866
rang d'une forme quadratique, 602
rang d'une matrice, 331
rare, 1088
rationnels, 206
rayon
de convergence, 1209
de courbure, 1751
de torsion, 1752
spectral, 726, 1005
recouvrement, 451
rectangle
produit de tribus, 1117
rectifiable, 1724
arc géométrique, 1740
rejet
région dans une prise de décision, 2455
relation binaire, 106
relation d'équivalence, 111
relations
coefficient-racines, 432
de Chasles, 525, 1678
relativement compact, 452
relèvement, 1761
Représentation
virgule flottante normalisée, 2245
représentation, 342
de groupe fini
caractères de S_4, 1307
fidèle, 342
groupe diédral, 1552
irréductible, 1298
produit tensoriel, 1306
régulière gauche, 1303
virgule fixe, 2245
repère
affine, 541
cartésien
espace affine, 526

de Frenet, 1751
projectif, 1799
reste, 177, 291
d'un développement limité, 1026
restriction d'une distribution, 2122
Riemann
fonction, 1255
risque
première espèce, 2456
quadratique, 2435
seconde espèce, 2456
rotation
en dimension 2, 1460
rotation d'angle θ, 1464
rotation-homothétie, 1817
rupture
corps, 408
règle du produit nul, 176
réciproque
continuité, 479
dérivabilité, 888
récurrent
nul, 2477
point d'un système dynamique, 1106
positif, 2477
état, 2477
réduction
d'endomorphisme, 638
Jordan, 650
réduction de Frobenius, 648
réel, 216
réflexif, 2029
réflexion, 1492
dans $\mathbb{R}^2$, 1457
glissée, 1490
par rapport à un hyperplan, 1449
région
critique, 2455
de confiance exact, 2446
de rejet, 2455
régularité
d'une mesure, 1104
extérieure de la mesure de Lebesgue, 1127
intérieure de la mesure de Lebesgue, 1129
régulier
arc, 1738
point d'un arc, 1738
régulier à droite, 176
régulière
surface, 1635
répulsif
point fixe, 1331
résidu
méthode itérative, 2313
résolvante, 2174
résultant, 558
utilisation, 561, 1659

Schrödinger, 2208
Schur (théorème), 1301
schéma
consistant, 2336
schéma discret convergent, 2337
schéma numérique, 2336
section, 909
de graphe, 894
propriété des, 1162
segment
dans $\mathbb{R}^p$, 678
dans un espace affine, 534
semi-défini positif, 623
semi-simple
endomorphisme, 587
semi-symétrique
polynôme, 429
seminorme, 515
seminorme quotient, 519
sesquilinéaire, 606
signature, 172
forme quadratique, 602
similitude, 955
simple
extension de corps, 400
fonction, 1114
module, 195
singularité, 1909
singularité effaçable, 1909
singularité isolée, 1909
sinus, 1413
hyperbolique, 1260
sinus cardinal, 1385, 1713
solfège, 1558
solution
générale, 2211
particulière, 2211
somme
inférieure, 1696
partielle, 734
supérieure, 1696
somme directe, 343
somme directe (de représentations), 1297
somme directe topologique, 471
somme partielles
Abel angulaire, 1654
sommet, 1775
Sophie Germain, 375
sous arc, 1723

sous-additivité
sur algèbre de parties, 1072
sous-anneau, 270
sous-corps premier, 373
sous-espace
affine engendré par une partie, 532
caractéristique, 635
sous-groupe
caractéristique, 161
distingué, 357
dans le groupe alterné, 361
engendré, 189
normal, 241, 356
sous-groupe distingué, 161
sous-martingale, 2493
sous-module, 194
sous-module engendré, 192
sous-suite, 834
spectre
matrice hermitienne, 709
matrice symétrique réelle, 622
spectre d'un endomorphisme, 580
sphère, 462
de Riemann, 1814
stabilisateur, 246
stabilité
d'un point d'équilibre, 2193
Lyapunov, 2193
stable, 2253
schéma numérique, 2339
stathme
sur $\mathbb{Z}[i]$, 378
stathme euclidien, 283
stationnaire
chaine de Markov, 2485
statistique, 2435
statistiques
descriptives, 2427
strictement
convexe
sur $\mathbb{R}^n$, 1370
strictement convexe, 1361
structure
complexe, 1758
structure d'anneau canonique, 174
structure réelle, 344
Student, 2407, 2448
subdivision, 1691
associée à une fonction, 1692
d'un intervalle, 1724
subordonnée
norme, 725
subpotent, 139
suite, 254
arithmético-géométrique, 751
de Cauchy, 490
dans un corps, 204
de fonctions, 1945
théorème de Montel, 2053
de fonctions intégrables, 1321, 1935
de Jordan-Hölder, 242
définie par itération, 2268
exacte, 252
régularisante, 2090
équirépartie, 2063
critère de Weyl, 2063
suite de composition, 241
suite de variables aléatoires de Bernoulli, 2489
suites adjacentes, 675
support, 1693
d'une permutation, 165
distribution, 2115
famille d'éléments, 254
supremum, 228
d'une suite d'ensembles, 1069
sur-martingale, 2493
surface paramétrée, 1634
surjection, 106, 478
surpotent, 139
Sylow
p-Sylow, 347
Sylvester (matrice), 558
symbole
de Legendre, 1575
symbole principal, 2225
symétrique
polynôme, 429
système
fondamental, 2171
orthonormé, 1867
trigonométrique, 1867, 1976
système trigonométrique, 1998
séparable, 1867
espace topologique, 447
extension de corps, 425
polynôme irréductible, 423
polynôme non constant, 423
élément d'une extension, 425
sépare
les points, 1007
séquentiellement fermé, 488
série
dans un espace vectoriel normé, 734
de Fourier, 2067, 2088
utilisation, 2073
de puissance, 1209

donnant $(1-A)^{-1}$, 791
entière, 1209, 1288, 2088
Abel angulaire, 1654
processus de Markov, 2491
utilisation, 2377
fonctions, 1315, 2088
génératrice d'une suite, 1268
géométrique, 750
nombres, 1315
numérique, 1258, 1288
Riemann, 751
Taylor, 1269
série convergente, 734
série de Laurent, 2076
série de Taylor, 1019
série divergente, 734
série entière
fonctions holomorphes, 1894
série harmonique, 749

tangent
vecteur unitaire, 1751
tangente, 1428, 1746
Tangente hyperbolique, 2532
tangente hyperbolique, 1262
tangente à un chemin, 1610
taubérien, 1314
taux d'accroissement, 1361
Taylor
série entière, 1268
Taylor avec reste intégral, 1688
temps d'arrêt, 2496
temps de retour, 2476
terminée
martingale, 2497
test, 2455
bilatéral, 2456
unilatéral, 2456
théorème
accroissements finis, 787
dans $\mathbb{R}$, 889
forme générale, 959
Ascoli, 1938
Baire, 514
Banach-Steinhaus, 759
avec seminormes, 1938
base incomplète, 301
Beppo-Levi, 1139
Bolzano-Weierstrass, 464
Borel-Cantelli, 2345
Borel-Lebesgue, 670
Brouwer, 1623
dimension 2, 1896
Bézout
polynômes, 387
utilisation, 559
Carathéodory, 538
Cauchy
groupe, 263
Cauchy-Arzela, 1626
Cauchy-Lipschitz, 1335
Cayley-Hamilton, 592, 1056
central limite, 2387
processus de Poisson, 2511
Chevalley-Warning, 1581
chinois, 383
anneau des polynômes, 384
anneau principal, 278
Cochran, 2432
Cochrane, 2433
convergence
dominée de Lebesgue, 1149
monotone, 1139
de Baire, 522
de Jordan, 1783
de représentation de Riesz, 1866
des deux carrés, 380
version faible, 379
Dini, 979
Dirichlet, 2060
forme faible, 1571
Doob, 2497
du rang, 310
décomposition des noyaux
et exponentielle de matrice, 1280
élément primitif, 1584
extension d'isométrie, 1404
Fejér, 2061
fonction implicite dans $\mathbb{R}^n$, 1347
fonction implicite dans Banach, 1346
Fubini
dans $\mathbb{R}^n$, 1201
espace mesuré, 1196
version compacte dans $\mathbb{R}^2$, 1668
Fubini-Tonelli, 1194
fuite des compacts, 2177
Gauss
polynômes, 388
Gauss-Wantzel, 1601
Glivenko-Cantelli, 2440
Hahn-Banach, 2019
Hardy-Littlewood, 1315
Heine, 835
incidence, 1792
inversion locale, 1344
utilisation, 1359, 1704
isomorphisme

second, 238
troisième, 239
isomorphisme de Banach, 1937
Kronecker, 561
Lagrange, 241
Lie-Kolchin, 848
Lotka-Volterra, 2197
Markov-Takutani, 1626
Montel, 2053
Pappus
affine, 1796
projectif, 1796
Pearson, 2460
petit de Fermat, 374
Picard, 1332
point fixe
Brouwer, 1896
projection
cas vectoriel, 1863
partie fermée convexe, 1861
prolongement de Hahn, 1099
prolongement de Riemann, 1910
Radon-Nikodym, 1157
complexe, 1158
Rolle, 888
Rothstein-Trager, 1659
Runge, 1900
Schauder, 1624
Schur, 1301
spectral, 637
autoadjoint, 704
matrice symétrique, 622
matrices normales, 845
stabilité de Lyapunov, 2193
Stone-Weierstrass, 1007, 1010
Sylvester, 631
taubérien, 1314
taubérien faible, 1656
transfert, 2378
Tykhonov, 511
dénombrable, 513
fini, 512
valeurs intermédiaires, 693
Von Neumann, 1353
Wedderburn, 1571
Weierstrass, 465
élément primitif, 427, 1585
théorème des extrémums liés, 1359
théorème fondamental du calcul intégral, 1178
topologie, 435, 571
$*$-faible, 521, 2112
p-adique, 799
discrète, 437
engendrée par une famille, 437
et seminormes, 515
forte, 726
grossière, 436
métrique, 460
produit, 438
sur $\mathscr{D}(\Omega)$, 2109
sur $\mathscr{D}(K)$, 2109
sur $C^\infty(\Omega)$, 2109
sur dual topologique, 521
usuelle sur $\mathbb{R}^n$, 571
topologie faible, 2043
topologie induite, 443
topologie quotient, 440
torsion, 1752
d'un groupe, 254
totale, 1867
trace, 2154
dual de $\mathbb{M}(n, \mathbb{K})$, 612
endomorphisme, 616
matrice, 615
produit scalaire sur $\mathbb{M}(n, \mathbb{R})$, 857
unicité pour la propriété de trace, 613
transcendant, 394
transformation
Fourier, 2088
gaussienne, 2293
transformée
de Cauchy, 1907
de Fourier, 2083, 2376
continuité, 2086
groupe abélien fini, 1295
Fourier
distribution tempérée, 2133
Laplace, 2377
transient
état, 2477
transition
probabilité, 2465
transitoire
état, 2477
transposition, 171
transposé, 565
transposée, 610
transvection, 1048
transvection (matrice), 330
transversale, 248
tribu, 1069
borélienne, 1084
de Baire, 1088
de Lebesgue, 1124
engendrée
par un événement, 2347

par une application, 1083
par une variable aléatoire, 2347
induite, 1070
produit, 1117
tribu de Lebesgue sur S^1, 1443
tribu engendrée, 1070
trigonalisation
et polynôme caractéristique, 844
simultanée, 847, 848
triplet
pythagoricien, 287
triplet naturel, 113
type
fini
en algèbre, 428
espace vectoriel, 300
tétraèdre régulier, 866

unicité
des mesures, 1078
uniformément continue, 514
uniformément convexe, 2001
unipotent, 176
unitaire
normale principale, 1751

valeur
principale (distribution), 2132
propre, 580
singulière, 639
valeur absolue, 233
p-adique, 799
valeur principale, 1926
valeur propre
d'une forme quadratique, 633
forme quadratique, 632
valuation
p-adique, 799
valuation d'un polynôme, 198
Vandermonde (déterminant), 555
variable
de décision, 2456
variable aléatoire, 2346
absolument continue, 2346
Bernoulli
marche aléatoire, 2470
utilisation, 2423
binomiale
utilisation, 2500
centrée, 2353
de Bernoulli
utilisation, 2500
de Rademacher, 2405
variance, 2353
empirique, 2354, 2431
empirique corrigée, 2431
vecteur gaussien, 2401
variation des constantes, 2167, 2172
variété, 1359, 1607
variété
orientée, 1617
variété orientable, 1617
vecteur
cyclique, 586
gaussien, 2401
propre, 580
unitaire normal, 1751
unitaire tangent, 1746
Vitali (ensemble), 1130
vitesse d'un chemin, 1723
voisinage, 436, 571
volume
d'une région solide, 1702
région bornée dans $\mathbb{R}^3$, 1700
volume dans $\mathbb{R}^3$, 1620
vraisemblance, 2438

Weiner
constante, 2559
Wronskien, 2200

Éther, 2522
Événement, 2527
élément
inversible
dans un anneau, 176
élément maximal, 107
élément minimal, 107
élément premier, 183
élément régulier, 116
élément régulier à gauche, 116
élémentaire
polynôme symétrique, 429
équation
aux variations, 2185
de Riccati, 2206
des classes, 249
des orbites, 248
différentielle
Hill, 2203
homogène, 2206
système, 2204
étude qualitative, 2204
diophantienne, 266, 286, 288
Fredholm, 1334
orbite-stabilisateur, 247
équation de droite, 867
équation différentielle

linéaire du premier ordre, 2216
linéaire du premier ordre, homogène, 2216
linéaire du second ordre, 2219
linéaire du second ordre, homogène, 2219
premier ordre, 2211
second ordre, 2212
variables séparables, 2214
équation exponentielle, 1013
équation fonctionnelle, 1013
équation homogène associée, 2216
équicontinu, 513
équipotent, 139
équivalence de forme quadratiques, 604
équivalence de normes, 721
étagée
fonction, 1114
état
apériodique, 2487
récurrent, 2477
récurrent positif, 2477
transitoire, 2477
étoile de Kleene, 2517
étranger
dans leur ensemble, 291
évaluation
polynôme plusieurs variables, 271

Liste des notations

Algèbre

$[\mathbb{L} : \mathbb{K}]$ degré d'une extension de corps, page 391

(p) idéal engendré par p, page 269

$\mathcal{L}(E, F)$ Ensemble des applications linéaires de E dans F, page 306

$\mathbb{F}_p$ lorsque p est premier, page 373

$\mathbb{F}_{p^n}$ corps fini à p^n éléments, page 1573

$\mathbb{K}(A)$ corps contenant $\mathbb{K}$ et A, page 400

$\mathbb{K}[A]$ anneau contenant $\mathbb{K}$ et A, page 400

$\mathrm{Frac}(A)$ Le corps des fractions de l'anneau A, page 203

$\mathrm{Fun}(X, Y)$ les applications de X vers Y, page 174

$\mathrm{S}(E)$ Les opérateurs autoadjoints de E, page 607

∇f gradient de la fonction f, page 932

$\mathtt{proj}_V$ projection de $V \times W$ sur V, page 484

$\mathrm{res}(P, Q)$ résultant des polynômes P et Q, page 558

$\mathrm{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A, page 297

$\sqrt{A}$ racine d'une matrice hermitienne positive, page 1057

$\theta_\alpha(P)$ la multiplicité de α par rapport à P, page 293

$A[X]$ tous les polynômes de degré fini à coefficients dans A, page 200

$A_n[X]$ les polynômes à coefficients dans A et de degré inférieur à n, page 290

$C(P)$ matrice compagnon, page 647

$D \mid P$ D divise P, page 291

$df_a(u)$ Application de la différentielle de f sur le vecteur u, page 919

$E_\lambda(T)$ Espace propre de T, page 580

$f^{(n)}$ La n-ième dérivée de la fonction f, page 1021

$mat_{\mathcal{B}}(q)$ matrice de q dans la base $\mathcal{B}$, page 1395

$o(x)$ fonction tendant rapidement vers zéro, page 1023

$U(A)$ ensemble des inversibles, page 176

U_n Le groupe des racines n^{e} de l'unité, page 1557

Ensembles de matrices

$S^+(n, \mathbb{R})$ matrices symétriques définies positives, page 624

$S^{++}(n, \mathbb{R})$ matrices symétriques strictement définies positives, page 624

$\mathrm{Aut}(E)$ automorphisme de l'espace vectoriel E, page 307

$L(E, F)$ applications linéaires bornées (continues), page 777

$\mathrm{End}(E)$ les endomorphismes de E, page 307

$\mathrm{O}(n, \mathbb{R})$ le groupe des matrices orthogonales, page 567

$\Omega(E)$ formes quadratiques non dégénérées, page 1391

$Q(E)$ formes quadratiques réelles sur E, page 597

$Q^+(E)$ formes quadratiques positives, page 1391

$Q^{++}(E)$ formes quadratiques strictement définies positives, page 1391

$S_n^{p,q}(\mathbb{R})$ matrices symétriques réelles de signature (p, q), page 1391

$\beta(s)$ Vecteur unitaire de la binormale, page 1751

$\gamma \sim g$ Équivalence d'arcs paramétrés, page 1739

$\nu(s)$ Vecteur unitaire de la normale principale, page 1751

$c(s)$ rayon de courbure, page 1751

$t(s)$ Torsion, page 1752

Géométrie

$(x_0 : \ldots : x_n)$ coordonnées homogènes dans un espace projectif, page 1810

$\mathcal{C}\mathrm{onv}(A)$ enveloppe convexe, page 536

$\mathbb{C}[G]$ combinaisons d'éléments de G à coefficients dans $\mathbb{C}$, page 1298

$\mathrm{PGL}(E)$ groupe projectif, page 1798

B^o orthogonal dans le dual, page 342

$P(E)$ l'espace projectif de E, page 1791

$P_1(\mathbb{C})$ sphère de Riemann, page 1814

Chaînes de Markov

$\pi(x)$ lié au temps de retour, page 2479

Probabilités et statistique

$\sigma(X)$ La tribu engendrée par la variable aléatoire X, page 2347

$a \wedge b$ $\min(a, b)$, page 2497

K_X matrice de covariance d'un vecteur gaussien, page 2401

$m(\mathcal{A})$ Ensemble des fonctions $\mathcal{A}$-mesurables, page 1086

Théorie des groupes

$(G/H)_g$ classes à gauche, page 246

$[a]_p$ ensemble des $a + kp$, page 267

$[G, G]$ groupe dérivé, page 237

$[g, h]$ commutateur dans un groupe, page 237

$\mathrm{Aff}(\mathbb{R}^n)$ Le groupe des applications affines bijectives de $\mathbb{R}^n$., page 546

gr_G groupe engendré, page 189

$\hat{G}$ groupe des caractères de G, page 1293

σ_x réflexion par rapport à x, page 1492

A_n groupe alterné, page 359

$D(G)$ groupe dérivé, page 237

D_n groupe diédral, page 1481

G^{ab} groupe abélianisé de G, page 238

$N \times_\phi H$ produit semi-direct, page 252

S_n le groupe symétrique, page 165

$N \lhd G$ Le sous-groupe N est normal dans G, page 161

Topologie et théorie des ensembles

$\mathrm{Adh}(A)$ adhérence de A, page 442

$\complement A$ Le complémentaire de l'ensemble A, page 109

$\mathrm{Diam}(A)$ Diamètre de la partie A, page 834

∂A La frontière de l'ensemble A, page 574

$A \Delta B$ différence symétrique, page 110

A^c complémentaire de A, page 109

Analyse

$C^\infty(I, \mathscr{D}'(\mathbb{R}^d))$ fonctions à valeurs dans les distributions, page 2118

$(S, \hat{\mathcal{F}}, \hat{\mu})$ complété de l'espace mesuré $(S, \hat{\mathcal{F}}, \hat{\mu})$, page 1092

$\mathcal{L}(E, F)$ Les applications linéaires de E vers F, page 777

$\arg(z)$ La valeur principale de l'argument de $z \in \mathbb{C}$, page 1926

$\mathbb{D}_b$ l'ensemble des écritures décimales en base b, page 754

$\mathbb{R}$ l'ensemble des réels, page 216

$\mathbb{R}^+$ les réels positifs ou nuls, page 219

exp exponentielle, page 1244

$\mathscr{H}'$ dual, page 1865

$\mathrm{Isom}(X)$ Le groupe des isométries de X, page 487

$\liminf a_n$ limite inférieure, page 676

$\limsup a_n$ limite supérieure, page 676

$\mathcal{L}^p$ espace de Lebesgue, sans les classes, page 1943

$\mu \ll \nu$ La mesure μ est absolument continue par rapport à la mesure ν., page 1157

$\mu \perp \nu$ mesures perpendiculaires, page 1157

μ^* La mesure extérieure associée à la mesure μ, page 1080

$\partial_z, \partial_{\bar{z}}$ dérivées partielles d'une fonction complexe, page 1887

$\mathtt{proj}_K(x)$ projection orthogonale de x sur y, page 1863

$\sigma(\mathcal{A})$ tribu engendrée par $\mathcal{A}$, page 1070

$\mathscr{D}(\Omega)$ Les fonctions C^∞ à support compact sur Ω, page 2112

$\mathscr{S}'(\mathbb{R}^d)$ espace des distributions tempérées, page 2131

$A^2(\Omega)$ espace de Bergman, page 2054

$A^\perp$ orthogonal d'une partie., page 1864

$C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ Fonctions à valeurs dans les distributions., page 2208

$C_c(I)$ fonctions continues à support compact dans I, page 1969

$f \sim g$ fonctions ayant des limites équivalentes, page 905

$H^1(\Omega)$ espace de Sobolev sur Ω, page 2150

$H^1(I)$ espace de Sobolev, page 2145

$H^m(M)$ espace de Sobolev, page 2152

$L^1_{loc}(I)$ fonctions intégrables sur les compacts de I, page 2145

L^p espace de Lebesgue avec les classes, page 1944

$M_i\varphi$ La fonction $x \mapsto x_i\varphi(x)$, page 2049

$S_n f$ somme partielle de série de Fourier, page 2000

Chapitre 0

Introduction

0.1 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.

0.1.1 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 19.8.

(8) Danarmk pour des réponses à des questions dans les commentaires (allongement pour éviter un Overfull hbox) `http://linuxfr.org/nodes/110155/comments/1675589`. Et aussi pour une discussion à propos de la topologie sur $\mathscr{D}(\Omega)$.

(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 [1]

(11) Bertrand Desmons pour plusieurs patchs 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.

(13) ybailly pour avoir détecté un bon nombre de coquilles dans la partie sur les ensembles de nombres.

(14) Éric Guirbal pour le remplacement de `frenchb` par `french`.

1. `http://linuxfr.org/nodes/110155/comments/1675813`

(15) cdrcprds pour une réponse à une question d'algèbre, démonstration à l'appui à propos de pgcd. Également pour sa relecture sans pitié de la partie sur la cardinalité (en particulier $A \approx A \backslash B$) et pour avoir pointé l'utilité du théorème de comparabilité cardinale.

(16) Antoine Bensalah pour avoir répondu à une question sur Lax-Milgram tout en même temps que pointé une erreur dans la démonstration et fourni l'exemple 25.62 sur l'optimalité de l'inégalité.

(17) Guillaume Deschamps pour ses remarques à propos du fait que le chapitre « constructions des ensembles » est très ardu.

(18) Guillaume Barriere pour sa relecture attentive jusqu'aux corps.

(19) Samy Clementz pour avoir découvert une faute dans la définition de mesure positive sur un espace mesurable.

(20) Sylvain Rousseau pour avoir clarifié une construction dans le théorème de Bower version C^∞.

(21) Maxmax pour des typos dans l'index thématique.

(22) Laurent Choulette pour une typo dans les propriétés du neutre d'un groupe.

(23) Pierre Lairez pour la démonstration du théorème d'inversion de limite et de dérivée 12.379 sans passer par les intégrales (et les lemmes correspondants à propos du module de continuité).

(24) Gregory Berhuy pour des réponses d'algèbres dans les catégories facile, moyen et difficile.

(25) Benoît Tran pour avoir signalé un paquet de typos dans la démonstration de l'ellipsoïde de John-Loewner et ses dépendances.

(26) Provatiscus pour avoir signalé un paquet de choses pas claires, et surtout pour avoir trouvé une faute dans la démonstration du fait qu'une fonction continue sur $\mathbb{Q}$ se prolonge en une fonction continue sur $\mathbb{R}$. Et pour cause : cet énoncé est faux. `https://github.com/LaurentClaessens/mazhe/issues/124`

(27) William pour l'environnement `example` qui gère correctement le triangle.

(28) Colin Pitrat pour de nombreuses remarques, typos et relecture de théorèmes.

(29) Bruno Turgeon pour une très belle moisson de fautes de frappe (euphémisme pour dire « mon ignorance crasse de l'orthographe »).

(30) Sacha Dhénin pour une belle quantité de fautes de frappes et pour avoir soulevé quelques points pas clairs (par exemple la définition du sous-groupe engendré dans le cas de la partie vide).

(31) Patrice Goyer pour m'avoir signalé quelques fautes et pas mal de points pas clairs dans les polynômes et dans la théorie généraliste des ensembles. Et pour m'avoir fait remarquer (deux fois) que mon script de déploiement ne marchait pas.

(32) Alain Vigne pour une quantité (presque) indénombrable de fautes d'orthographe et de mauvaises tournure de phrase dans les espaces vectoriels, la construction de nombres, la théorie des groupes et les anneaux. Il m'a également pointé quelque fautes et points vraiment pas clairs dans des démonstrations dont la correction a permit de bien améliorer la qualité du texte.

0.1.2 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 le 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).

0.1.3 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 3.36.

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

J'ai souvent donné entre parenthèses 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 hypertextes vers des sites, des blogs, etc. Tous ces gens ont fait du bon boulot. Sans toute cette « communauté », l'internet serait mort [2].

0.2 Originalité

Ces notes ne sont pas originales par leur contenu : ce sont toutes des choses qu'on trouve facilement sur internet ; je crois que la bibliographie est éloquente à ce sujet. Ce cours se distingue des autres sur les points suivants.

La longueur J'ai décidé de ne pas me soucier de la taille du fichier. Il fera cinq mille pages s'il le faut, mais il restera en un bloc. Étant donné qu'il n'existe qu'une seule mathématique, il ne m'a pas semblé intéressant de produire une division artificielle entre l'analyse, la géométrie ou l'algèbre. Tous les résultats d'une branche peuvent (et sont) être utilisés dans toutes les autres branches.

Dans cette optique, je me suis évertué à ne créer que des références « vers le haut ». À moins d'oubli de ma part [3], il n'y a aucun endroit du texte qui dépend d'un lemme démontré plus bas. Le fait qu'un théorème B soit plus bas qu'un théorème A signifie qu'on peut démontrer A sans savoir B.

La licence Ce document est publié sous une licence libre. Elle vous donne explicitement le droit de copier, modifier et redistribuer.

Les mises à jour Ce document est régulièrement mis à jour. Des fautes d'orthographe sont corrigées (presque) chaque jour. Si vous me signalez une faute de mathématique, elle sera corrigée.

2. 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.

3. Par exemple pour les théorèmes pour lesquels je n'ai pas lu ni a fortiori écrit de preuves.

Transparence Je ne fais pas semblant que ces notes soient parfaites. Les points sur lesquels je ne suis pas sûr, les preuves que j'ai inventées moi-même sont clairement indiqués pour inciter le lecteur à redoubler de prudence. Une liste de questions à résoudre est inclue en la section 0.7. Voir 0.3 pour plus de détails.

0.3 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 prudence.

La référence [1] D'abord les références à [1] indiquent qu'une bonne partie de ce qui suit est de l'invention personnelle de l'auteur. Cela ne veut évidemment pas dire que c'est moi qui ai découvert le résultat. Ça veut dire que je n'ai pas trouvé le résultat ou certaines parties de la preuve.

Les notes en bas de page Certaines notes en bas de page sont écrites dans une fonte spéciale [4]. Elles indiquent des points sur lesquels je doute ou des étapes de calculs que je ne parviens pas à reproduire en suivant mes sources. Lorsque vous voyez une telle note, redoublez de prudence, allez voir la source, et écrivez-moi si vous pouvez résoudre le problème.

Les environnements dédiés Et enfin certains problèmes sont indiqués plus longuement dans un environnement dédié en petits caractères comme ceci :

¡¡ Avertissement/question au lecteur !! 0.1
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.

0.4 Quelques choix qui peuvent provoquer des quiproquos

Comme tout cours de mathématique, ce cours fait des choix qui sont parfois discutables. Voici quelques points sur lesquels les choix faits ici ne sont peut-être pas ceux fait par tout le monde. Ce sont donc des points sur lesquels vous devez faire attention pour éviter les quiproquos lors par exemple d'un oral dans un concours.

(1) Nous utilisons la définition « épointée » de limite d'une fonction en un point. Elle diffère de celle donnée par le ministère de l'enseignement en France. Si votre but est de passer un concours d'enseignement en France, vous devriez lire 12.2 ; dans tous les autres cas, la définition prise ici est celle qu'il vous faut.

(2) Un compact est une partie d'un espace topologique pour lequel tout recouvrement par des ouverts admet un sous-recouvrement fini. Le fait d'être séparable n'est pas inclus dans la définition de compact. De nombreux textes français incluent la séparabilité dans la compacité.

(3) Le logarithme sur $\mathbb{C}$ est une application $\ln\colon \mathbb{C}^* \to \mathbb{C}$ définie partout sauf en zéro. Elle n'est donc pas continue sur la fameuse demi-droite. À ne pas confondre avec une *détermination* du logarithme qui est par définition continue et donc non définie sur la demi-droite.

Cela est un choix très discutable. La raison de donner à la notation « ln » cette signification est simplement de suivre l'usage de Sage. Pour Sage, $\ln(-1)$ existe et vaut $i\pi$.

Voir les remarques 26.78.

(4) Le mot « corps » n'implique pas la commutativité bien que tous les corps du Frido soient commutatifs, et nous n'utilisons pas la terminologie « anneau à division ». Voir la section -2.3 et la discussion 6.1.

4. Les notes comme celle-ci signifient qu'il y a certaines choses dont je ne suis pas sûr.

0.5 Autres choix pas spécialement standards

Nous listons ici quelques choix qui n'induisent pas de différences ou d'incompatibilité avec les autres cours, mais qui doivent être compris et justifiés.

(1) Nous n'utilisons pas les notations $o(x)$ ou autres $O(N^2)$. D'abord parce que je n'ai jamais très bien compris comment elles fonctionnent, et ensuite (surtout) parce que ces notations induisent en erreur. Ce sont des notations qui cachent, sous des notations a peu près intuitive, l'utilisation de théorèmes pas simples.

Écrire

$$f(x) = P(x) + o(x^2), \tag{0.1}$$

c'est un peu comme quand on écrit (horreur !)

$$F(x) = \int f(x)dx + C. \tag{0.2}$$

Où est le x à droite ? Quel est le statut de C ?

Même chose pour la notation $f(x) = P(x) + o(x^2)$. Le x de $o(x^2)$ est-il le x qu'on a à gauche ? Si $g(x) = Q(x) + o(x^2)$, est-ce le même o que celui de f ?

(2) Nous allons être plus calme avec la notation $A[X]$ pour les polynômes sur l'anneau A, et encore moins $A[X_1, \ldots, X_n]$ pour les polynômes de n variables. Au lieu de cela nous utilisons $\mathcal{P}(A)$ et $\mathcal{P}_n(A)$.

Est-ce que vous diriez que $A[X] = A[Y]$? Quelle est exactement la nature de X dans $P = X^2 + 1$ ou dans $P(X) = X^2 + 1$? Si $P \in A[X]$ vaut $P(X) = X^2$ et si $Q \in A[Y]$ vaut $Q(Y) = Y^2$, est-ce que vous oseriez écrire $P = Q$?

0.5.1 Mathématique intéressante

Définition 0.2.
Une notion mathématique est ***intéressante*** *si elle permet de répondre à une question que l'on peut se poser sans connaitre la notion.*

Exemple 0.3.
Étant donné un segment dans le plan, quels sont les triangles rectangles dont ce segment est l'hypoténuse ?

Nous n'avons pas besoin de cercles pour poser cette question. Mais nous avons besoin de connaissances sur les cercles pour y répondre. Donc l'étude des cercles est intéressante. △

Exemple 0.4.
Comment fonctionne la gravitation ?

Cette question peut être posée sans connaitre de calcul tensoriel, d'équations différentielles ou d'intégrales. Et pourtant, tous ces concepts sont utiles pour y répondre. △

0.6 Sage est là pour vous aider

Il existe de nombreux logiciels de mathématique. Notre préféré est Sage pour une raison très précise : Sage est (en simplifiant) un module pour python. Donc quand on travaille en Sage, on dispose de tout Python. La syntaxe et la structure de Sage ne sont pas *ad hoc* pour faire des maths, 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.

0.6.1 Lancez-vous dans Sage

(1) Aller sur http://www.sagemath.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 [4].

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

0.6.2 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 43.1 et 43.2.

(2) Tracer des graphes de fonctions, exemple 43.2.

(3) Tracer des courbes en trois dimensions, voir exemple 12.207. 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 43.3.

(5) Résoudre des systèmes d'équations linéaires. Voir les exemples 43.4 et 43.5. Lire aussi la documentation.

(6) Tout savoir d'une forme quadratique, voir exemple 43.6.

(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 les extrémums de la fonctions (sous réserve d'être capable de résoudre certaines équations), voir les exemples 43.7 et 43.8.

(8) Indiquer une infinité de solutions à une équation en utilisant des paramètres. L'exemple 43.9 montre ça avec une équation algébrique. Un exemple concernant des 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)
[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 43.10.

(10) Calculer des approximations numériques comme dans l'exemple 43.11.

(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 19.68.

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 0.5.
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 15.137 pour le logarithme de nombres négatifs. Et aussi ceci :

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.

```

 SageMath version 7.3, Release Date: 2016-08-04
 Type "notebook()" for the browser-based notebook interface.
 Type "help()" for help.

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

tex/sage/sageSnip017.sage

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

$$\lim_{x\to 0} \frac{1}{x} = \infty \tag{0.3}$$

ainsi que

$$\lim_{x\to 0} \frac{1}{x^2} = +\infty. \tag{0.4}$$

Voir aussi la compactification en un point d'Alexandroff 7.88.

0.7 Comment contribuer et aider ?

0.7.1 Des preuves qui manquent

Vous trouverez un peu partout des énoncés sans preuves. Certaines sont surement très faciles, et d'autres probablement assez compliquées. N'hésitez pas à rédiger une preuve et me l'envoyer.

Vous pouvez m'envoyer vos preuves sous forme de « c'est bien fait dans tel cours », avec une URL.

Ne me dites juste pas « c'est bien fait dans tel *livre* ». Je ne travaille pas à l'université, et je n'ai pas accès à une bibliothèque universitaire ; je n'ai donc pas réellement accès à ces fameux « livres » dont tout le monde parle.

0.7.2 Du texte qui manque

Vous remarquerez que de nombreuses pages du Frido sont des enchainements de théorèmes et démonstrations sans articulations. Autrement dit, il manque ce qu'à l'agrégation on dirait à l'oral quand on présente le plan. Si vous avez des idées de choses à ajouter ici où là, faites le moi savoir.

0.7.3 Des exemples qui manquent

Si vous connaissez de bons exemples, faites-le moi savoir.

0.7.4 Trucs de programmation et de LaTeX

(1) Comment faire en sorte que les mots commençant par « é » soient avec les « e » dans l'index, et non avant les « a » ? Il me faudrait un mécanisme plus automatique que faire `machin@truc`.

Une fonction en python qui prend en entrée le fichier bbl sous forme de string et qui ressort sous forme de string le fichier bbl modifié me convient.

(2) Il y a des problèmes dans la table des matières. « Table des matières », « Index », et « Liste des notations » ne pointent pas vers la bonne page.

(3) 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`.

0.8 Les politiques éditoriales

Certaines parties de ce texte ne respectent pas les politiques éditoriales. Ce sont des erreurs de jeunesse, et j'en suis le premier triste.

0.8.1 Licence libre

Je crois que c'est clair.

0.8.2 pdflatex

Tout est compilable avec pdfLaTeX. Pas de pstricks, de psfrag ou de ps<quoiquecesoit>.

0.8.3 utf8

Je crois que c'est clair.

0.8.4 Notations

On essaie d'être cohérent dans les notations et les conventions. Pour la transformée de Fourier par exemple, je crois que la définition du produit scalaire dans L^2, des coefficients de Fourier, de la transformation et de la transformation inverse sont cohérents. Cela demande, lorsqu'on suit un livre qui ne suit pas les conventions utilisées ici, de convertir parfois massivement.

0.8.5 De la bibliographie

On évite d'écrire en haut de chapitre « les références pour ce chapitre sont ... ». Il est mieux d'écrire au niveau des théorèmes, entre parenthèses, les références.

Lorsqu'on écrit l'énoncé d'un théorème sans retranscrire la démonstration, il faut mettre une référence vers un document *en ligne* qui en contient la preuve. Il est vraiment fastidieux de chercher une preuve sur internet et de tomber sur des dizaines de documents qui donnent l'énoncé mais pas la preuve.

0.8.6 Faire des références à tout

Lorsqu'un utilise le théorème des accroissements finis, il ne faut pas écrire « d'après le théorème des accroissements finis, blablabla ». Il faut écrire un `\ref` explicite vers le résultat. Cela alourdit un peu le texte, mais lorsqu'on joue avec un texte de plus de 2000 pages, il est parfois laborieux de trouver le résultat qu'on cherche (surtout si il existe plusieurs versions d'un résultat et que l'on veut faire référence à une version particulière).

0.8.7 Des listes de liens internes

Le début du Frido contient une espèce d'index thématique. Il serait bon de l'étoffer.

0.8.8 Pas de références vers le futur

Dans le Frido, *aucune* preuve ne peut faire une référence vers un résultat prouvé plus bas. On n'utilise pas le théorème 10 dans la démonstration du théorème 7. Cela est une contrainte forte sur le découpage en chapitres et sur l'ordre de présentation des matières.

Il est bien entendu accepté et même encouragé de mettre des notes du type « Nous verrons plus loin un théorème qui ... ». Tant que ce théorème n'est pas *utilisé*, ça va.

Chapitre 1

Construction des ensembles de nombres

1.1 Quelques éléments sur les ensembles

1.1.1 Petit mot d'introduction

Le Frido n'est pas supposé être lu dans l'ordre de la première à la dernière page ; les matières y sont présentées dans l'ordre logique mathématique, et non dans l'ordre logique pédagogique, et encore moins par ordre de difficulté croissante.

1.1 (On saute la théorie des ensembles)**.**
En mathématique, si on lit une démonstration et que l'on veut vraiment tout justifier, et justifier toutes les étapes de tous les résultats utilisés, on tombe forcément un jour sur les axiomes.

Or l'axiomatique est un sujet particulièrement difficile. Nous n'allons donc pas « tout justifier » jusque là. Nous n'allons même pas préciser quel système d'axiome est utilisé. En particulier nous n'allons pas donner l'axiomatique des ensembles : nous allons supposer connus les ensembles et leurs principales propriétés.

Bref. Nous supposons avoir une théorie des ensembles qui tient la route. En particulier nous supposons connues les notions suivantes :

(1) ensemble vide,

(2) ensemble, appartenance, intersection, union,

(3) application entre deux ensembles, notation $f(x)$ pour désigner l'image de x par f,

(4) produit cartésien de plusieurs ensembles.

Ce sont toutes des choses dont la construction à partir des axiomes n'est en aucun cas évidente. En particulier, des « définitions » comme « l'intersection de deux ensembles est l'ensemble contenant exactement les éléments communs aux deux ensembles » ne sont pas correctes parce qu'elles passent à côté de l'existence et de l'unicité d'un tel ensemble.

1.2 (On saute la grammaire)**.**
Nous n'allons pas non plus formaliser la grammaire des expressions mathématiques[1]. Nous supposons que vous êtes capables de lire des expressions comme

$$x \in \mathbb{N} \Rightarrow \{a \geqslant x\} \text{ est infini.} \tag{1.1}$$

Tout cela pour dire que le Frido ne traitera que de la partie facile de la mathématique.

Définition 1.3.
Deux ensembles A et B sont ***disjoints*** *si leur intersection est vide*[2] *; en d'autres termes, si il n'existe aucun élément commun à A et B.*

1. J'utilise ici des mots que je ne comprends pas, juste pour me donner l'air malin.
2. Remarquez que les mots « intersection » et « vide » sont de ceux que nous avons décidé de ne pas définir.

1.4.
Remarquez par exemple que la première phrase de l'article de Wikipédia sur la construction de $\mathbb{N}$ est « Partant de la théorie des ensembles, on identifie 0 à l'ensemble vide, puis on construit ... ». Il est bien précisé que l'on part d'une théorie des ensembles.

1.5.
La suite de ce chapitre sera essentiellement sans exemple parce qu'avant d'avoir construit les ensembles de nombres, je ne sais pas très bien quels exemples on peut donner de quoi que ce soit.

1.1.2 Injection, surjection, bijection

Définition 1.6.
Soient deux ensembles E et F. Une application $f\colon E \to F$ est

(1) ***surjective*** *si pour tout $y \in F$, il existe $x \in E$ tel que $y = f(x)$;*

(2) ***injective*** *si pour tout $y \in F$, il existe au plus un $x \in E$ tel que $y = f(x)$;*

(3) ***bijective*** *si elle est à la fois injective et surjective.*

La méthode la plus courante pour démontrer qu'une application $f\colon E \to F$ est injective est de considérer $x, y \in E$ tels que $f(x) = f(y)$, et de prouver à partir de là que $x = y$. Ou alors de supposer $x \neq y$ et d'obtenir une contradiction.

La technique de la contradiction est évidemment la plus courante lorsque l'égalité $f(x) = g(x)$ implique une équation faisant intervenir $1/(x - y)$.

Lemme 1.7.
Soient deux ensembles A et B ainsi qu'une application $f\colon A \to B$. Nous supposons qu'il existe une application $g\colon B \to A$ telle que $f \circ g = \mathrm{Id}_B$ et $g \circ f = \mathrm{Id}_A$.

Alors f est une bijection.

Démonstration. En deux parties.

(i) **Injection** Supposons que $f(a) = f(b)$. Alors en appliquant g des deux côtés, et en utilisant le fait que $g \circ f = \mathrm{Id}_A$, nous trouvons $a = b$.

(ii) **Surjection** Soit $x \in B$. Posons $a = g(x)$. Alors, en utilisant le fait que $f \circ g = \mathrm{Id}_B$ nous avons
$$f(a) = (f \circ g)(x) = x. \tag{1.2}$$
Donc x est dans l'image de f et f est surjective.

□

1.1.3 Ensemble ordonné

1.8.
L'**axiome du choix** que nous acceptons peut s'énoncer comme ceci[5] : 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 1.9 ([6]).
Une ***relation binaire*** *d'un ensemble E vers un ensemble F est une partie de $E \times F$.*

Si G est une relation binaire entre E et F, nous notons $x\mathcal{R}_G y$ et nous disons que x est en relation avec y.

Définition 1.10.
Une ***relation d'ordre*** *sur un ensemble E est une relation binaire*[3] *(notée $\leqslant$) sur E telle que pour tous $x, y, z \in E$,*

réflexivité *: $x \leqslant x$*

3. Définition 1.9.

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

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

Pour suivre les notations de la définition 1.9, la partie G de $E \times E$ est une relation d'ordre lorsque

(1) $(x, x) \in G$ pour tout $x \in G$,

(2) Si $(x, y) \in G$ et $(y, x) \in G$, alors $x = y$,

(3) Si $(x, y) \in G$ et $(y, z) \in G$, alors $(x, z) \in G$.

Dans la suite nous n'allons plus écrire de relations binaires en détaillant l'ensemble sous-jacent.

*Lorsque nous avons un ensemble E et une relation d'ordre $\leqslant$ sur E, nous disons que le couple $(E, \leqslant)$ est un **ensemble ordonné**.*

Définition 1.11.
*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 1.12 ([7])**.**
*Soit un ensemble ordonné $(E, \leqslant)$. Un élément $M \in E$ est un **élément maximal** de E si pour tout $x \in E$ tel que $M \leqslant x$, nous avons $M \leqslant x \Rightarrow x = M$.*

*Nous disons que $m \in E$ est un **élément minimal** si pour tout $x \in E$, nous avons $x \leqslant m \Rightarrow x = m$.*

Définition 1.13.
*Soit un ensemble ordonné $(E, \leqslant)$ et une partie A de E. Nous disons que $m \in A$ est un **minimum** de A si pour tout $x \in A$, l'élément m est comparable à x et $m \leqslant x$.*

*Un élément $p \in E$ est un **minorant** de A si pour tout $a \in A$, les éléments p et a sont comparables et $p \leqslant a$.*

*Les notions de **maximum** et de **majorant** sont définies de façon analogue.*

Lemme 1.14.
Si E est un ensemble ordonné et si A est une partie finie totalement ordonnée de E, alors A possède un unique minimum et un unique maximum.

1.15.
Notons qu'il n'est pas demandé à un élément maximal[4] d'être comparable à tous les autres éléments. Si $(E, \leqslant)$ n'est pas totalement ordonné, un élément maximal peut ne majorer qu'une partie de E.

Un élément maximal est plus grand que tous les éléments avec lesquels il est comparable.

Dans un ensemble totalement ordonné, les notions d'élément maximal de E et de maximum[5] de E coïncident ; dans le cas d'un ordre partiel, ces notions sont distinctes.

Il se peut que nous parlions d'un « maximum » ou « élément maximum » au lieu d'un « élément maximal », en particulier en utilisant le lemme de Zorn. Si vous voyez de telles choses, n'hésitez pas à me le dire.

Par exemple, quand on applique le lemme de Zorn pour démontrer l'existence d'une base dans un espace vectoriel de dimension infinie, on obtient une famille libre qui est maximale, c'est-à-dire qui est un élément maximal dans l'ensemble, ordonné par inclusion, des familles libres de l'espace vectoriel. C'est un élément maximal, et non un maximum, car l'ensemble des familles libres d'un espace vectoriel non réduit à $\{0\}$ n'admet pas de maximum (pour l'inclusion). Voir 4.23 et 4.24.

Lorsqu'une partie possède un minimum, ce dernier est nommé le « plus petit élément » de la partie. Attention : il n'en existe pas toujours. D'innombrables exemples pourront être vus lorsque nous aurons construit $\mathbb{Q}$ et $\mathbb{R}$. Typiquement les intervalles du type $]a, b[$.

4. Définition 1.12
5. Définition 1.13.

Exemple 1.16 ([1]).
Soit un ensemble E ainsi que $a \neq b$ dans E. Nous considérons les parties

$$A = \{P \subset E \text{ tel que } a \in P, b \notin P\} \tag{1.3a}$$
$$B = \{P \subset E \text{ tel que } b \in P, a \notin P\}. \tag{1.3b}$$

Et enfin nous considérons l'ensemble $F = A \cup B$, c'est à dire l'ensemble des parties de E qui contiennent soit a soit b mais pas les deux. Nous ordonnons partiellement F par l'inclusion.

Dans $(F, \subset)$, l'élément $E \backslash \{b\}$ est un élément maximal, mais pas un majorant. △

Définition 1.17.
*Un ensemble ordonné est **bien ordonné** si toute partie non vide possède un plus petit élément.*

Autrement dit, l'ensemble ordonné E est bien ordonné si pour toute partie non vide A, il existe $x \in A$ tel que $x \leqslant y$ pour tout $y \in A$.

1.18.
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 1.376 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 $y \leqslant x$.

Exemple 1.19.
Si E est un ensemble, l'inclusion est un ordre sur l'ensemble des parties de E, mais pas un ordre total parce que si X, Y sont des parties de E, alors nous n'avons pas automatiquement soit $X \subset Y$ soit $Y \subset X$. △

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

Définition 1.20.
*Soit un ensemble totalement ordonné $(E, \leqslant)$. Un **intervalle** de E est une partie I telle que tout élément compris entre deux éléments de I soit dans I. En language mathématique, la partie I de E est un intervalle si*

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

1.1.4 Lemme de Zorn

Nous admettons l'axiome du choix qui s'énonce de la façon suivante[8] :

> Pour tout ensemble X d'ensembles non vides, il existe une fonction définie sur X, appelée fonction de choix, qui à chaque ensemble A appartenant à X associe un élément de cet ensemble A.

Définition 1.21 (Ensemble inductif[9]).
*Un ensemble est **inductif** si toute partie totalement ordonnée admet un majorant.*

Lemme 1.22 (Lemme de Zorn[10]).
Tout ensemble ordonné inductif non vide admet au moins un élément maximal.

Proposition 1.23 ([11]).
Soient un ensemble inductif $(E, \leqslant)$ et $b \in E$. Il existe un élément maximal[6] *$m \in E$ tel que $b \leqslant m$.*

Démonstration. En plusieurs parties.

6. Définition 1.12.

(i) **Un ensemble** Nous considérons

$$E_b = \{x \in E \text{ tel que } b \leqslant x\}. \tag{1.4}$$

(ii) **E_b est inductif** Soit une partie non vide totalement ordonnée A de E_b. Puisque E est inductif, la partie A admet un majorant $m \in E$.

Comme A est non vide, nous pouvons considérer $x \in A$. Nous avons $b \leqslant x$ parce que $x \in A \subset E_b$. Mais comme m est un majorant de A, $x \leqslant m$. Bref, nous avons les inégalités

$$b \leqslant x \leqslant m. \tag{1.5}$$

Donc $m \in E_b$. Nous avons prouvé que m est un majorant de A contenu dans E_b. Donc E_b est inductif.

(iii) **Zorn** Puisque $(E_b, \leqslant)$ est inductif, le lemme de Zorn 1.22 nous indique que E_b a un élément maximal. Nous le notons m.

(iv) **m est maximal dans E** Supposons avoir un élément $a \in E$ tel que $m \leqslant a$. Nous avons

$$b \leqslant m \leqslant a, \tag{1.6}$$

et donc $a \in E_b$. Mais m est maximal dans E_b, donc $a = m$.

□

1.1.5 Complémentaire

Définition 1.24.
Soit E, un ensemble et A, une partie de E (c'est-à-dire un sous-ensemble de E). Le ***complémentaire*** *de l'ensemble A, dans E, noté $\complement A$ est l'ensemble des éléments de E qui ne font pas partie de A :*

$$\complement A = E \backslash A = \{x \in E \text{ tel que } x \notin A\}. \tag{1.7}$$

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

Lemme 1.25 (Quelques relations ensemblistes)**.**
Soient $A, B, C \subset X$. Nous avons

(1) $X \backslash (A \cap B) = (X \backslash A) \cup (X \backslash B)$.
(2) $X \backslash (A \cup B) = (X \backslash A) \cap (X \backslash B)$.
(3) $A \cap (B \backslash C) = (A \cap B) \backslash C$.

Lemme 1.26.
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) $A \backslash B = A \cap \complement B$.
(3) $(A \backslash B)^c = A^c \cup B$.
(4) $A^c \backslash B^c = B \backslash A$.

Démonstration. Plusieurs points.

(i) **Pour (2)**

(ii) **Pour (3)** Il faut le faire en deux inclusions.

(i) $(A \backslash B)^c \subset A^c \cup B$ Supposons que $x \in (A \backslash B)^c$. Si $x \in A^c$, c'est bon. Supposons que x n'est pas dans A^c, et montrons que $x \in B$. Le fait que x ne soit pas dans A^c signifie que $x \in A$. Si x n'était pas dans B, alors x serait dans $A \backslash B$, ce qui est contraire à l'hypothèse. Donc $x \in B$.

(ii) $\underline{A^c \cup B \subset (A\backslash B)^c}$ Supposons d'abord que $x \in A^c$. Comme $A\backslash B \subset A$, si $x \in A^c$, alors $x \in (A\backslash B)^c$.

Si $x \in B$, alors x n'est pas dans $A\backslash B$ et donc x est dans $(A\backslash B)^c$.

(iii) **Pour (4)** Pour cette égalité, nous séparons 4 cas suivant que x est dans A ou B ou non. Bref, nous écrivons la table de vérité :

A	1	1	0	0
B	1	0	1	0
$A^c\backslash B^c$	0	0	1	0
$B\backslash A$	0	0	1	0

(1.8)

Les deux dernières lignes étant égales, nous avons l'égalité d'ensembles annoncée.

□

Définition 1.27 (différence symétrique)**.**
Si A et B sont des ensembles, l'ensemble $A\Delta B$ est la ***différence symétrique*** *d'ensembles :*

$$A\Delta B = (A \cup B)\backslash(A \cap B). \tag{1.9}$$

C'est l'ensemble des éléments étant soit dans A soit dans B mais pas dans les deux, ni dans aucun des deux. La table de vérité de $A\Delta B$ est intéressante :

A	1	1	0	0
B	1	0	1	0
$A\Delta B$	0	1	1	0

(1.10)

La deuxième colonne signifie que si $x \in A$ et $x \in B^c$, alors $x \in A\Delta B$.

Lemme 1.28 ([12])**.**
Si A et B sont des parties d'un ensemble, nous avons

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

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

(3) $(A\Delta B)^c = (A^c \cap B^c) \cup (A \cap B)$.

(4) Associativité : $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

Démonstration. En plusieurs points.

(i) **Pour (1)** Nous rappelons l'égalité $X^c\backslash Y^c = Y\backslash X$ du lemme 1.26(4). De là nous écrivons

$$A^c\Delta B^c = (A^c \cup B^c)\backslash(A^c \cap B^c) = (A^c \cap B^c)^c\backslash(A^c \cup B^c)^c = (A \cup B)\backslash(A \cap B) = A\Delta B. \tag{1.11}$$

(ii) **Pour (2)** Ici, il faut remarquer que $(A\Delta B) \cup B = A \cup B$ et que $(A\Delta B) \cap B = B\backslash A$, donc

$$(A\Delta B)\Delta B = (A \cup B)\backslash(B\backslash A) = A. \tag{1.12}$$

(iii) **Pour (3)** Il s'agit d'utiliser le lemme 1.26(3) :

$$(A\Delta B)^c = \Big((A \cup B)\backslash(A \cap B)\Big)^c \tag{1.13a}$$
$$= (A \cup B)^c \cup (A \cap B) \tag{1.13b}$$
$$= (A^c \cap B^c) \cup (A \cap B). \tag{1.13c}$$

(iv) **Pour l'associativité (4)** Nous écrivons les tables de vérités selon que x est dans A, B, C ou non. D'abord

A	1	1	1	1	0	0	0	0
B	1	1	0	0	1	1	0	0
C	1	0	1	0	1	0	1	0
$B\Delta C$	0	1	1	0	0	1	1	0
$A\Delta(B\Delta C)$	1	0	0	1	0	1	1	0

(1.14)

La quatrième ligne s'écrit sur le modèle de (1.10) en regardant les deuxièmes et troisièmes lignes. La dernière ligne se fait avec la première et la quatrième.

L'autre table de vérité se fait de la même manière :

A	1	1	1	1	0	0	0	0
B	1	1	0	0	1	1	0	0
C	1	0	1	0	1	0	1	0
$A\Delta B$	0	0	1	1	1	1	0	0
$(A\Delta B)\Delta C$	1	0	0	1	0	1	1	0

(1.15)

Puisque les lignes pour $A\Delta(B\Delta C)$ et pour $(A\Delta B)\Delta C$ sont identiques, nous avons égalité.

□

1.1.6 Relations d'équivalence

Définition 1.29.
Si E est un ensemble, une ***relation d'équivalence*** *sur E est une relation binaire*[7] *$\sim$ 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 $x \sim z$.*

Définition 1.30.
*Si E est un ensemble et si $\sim$ est une relation d'équivalence sur E, alors nous notons $E/\sim$ l'****ensemble quotient****, c'est-à-dire l'ensemble des classes d'équivalence dans E. Un élément de $E/\sim$ est de la forme*

$$[a] = \{x \in E \text{ tel que } x \sim a\}. \tag{1.16}$$

Lemme 1.31.
Soit un ensemble E et une relation d'équivalence $\sim$. Pour $a, b \in E$, nous avons $[a] = [b]$ si et seulement si $a \sim b$.

Démonstration. En deux parties.

(i) $\Rightarrow$ Nous supposons que $[a] = [b]$. Par réflexivité, $a \sim a$ et nous avons $a \in [a] = [b]$. Mais $a \in [b]$ signifie $a \sim b$, ce qu'il fallait.

(ii) $\Leftarrow$ Nous supposons que $a \sim b$, et nous démontrons que $[a] \subset [b]$ (pour l'inclusion inverse, vous devriez vous en sortir tout seul). Si $x \in [a]$, alors $x \sim a$. Mais $a \sim b$. Donc $x \sim a \sim b$, ce qui implique $x \sim b$ par transitivité. Or dire $x \sim b$ implique $x \in [b]$.

□

Exemple 1.32.
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. C'est une relation d'équivalence :

— un polygone P a toujours le même nombre de côtés que lui-même : $P \sim P$;

7. Définition 1.9.

— 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 ($P \sim R$).

△

Exemple 1.33.
Soit f une application entre deux ensembles E et F. Nous définissons une relation d'équivalence sur E par

$$x \sim y \Leftrightarrow f(x) = f(y). \tag{1.17}$$

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

$$\begin{aligned} g\colon E/\sim &\to F \\ [x] &\mapsto f(x) \end{aligned} \tag{1.18}$$

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

$$f = g \circ \pi. \tag{1.19}$$

△

1.2 Quelques structures algébriques

Nous collectons ici les définitions des principales structures algébriques.

Définition 1.34 (Groupe)**.**
*Un **groupe** est un ensemble G muni d'une opération interne $\cdot : G \times G \to G$ telle que*

(1) pour tous g, h, $k \in G$, $g \cdot (h \cdot k) = (g \cdot h) \cdot k$,

(2) il existe un élément $e \in G$ tel que $e \cdot g = g \cdot e = g$ pour tout $g \in G$,

(3) pour tout $g \in G$, il existe un élément $h \in G$ tel que $g \cdot h = h \cdot g = e$.

*Un groupe est **commutatif** ou **abélien** si $g \cdot h = h \cdot g$ pour tout $g, h \in G$.*

Notons que nous avons écrit $g \cdot h$ et non $\cdot(g, h)$ comme une notation purement fonctionnelle nous l'aurait suggéré. Dans les exemples concrets, selon les cas, la loi de groupe appliquée à g et h sera notée tantôt $g + h$, tantôt $g \cdot h$ ou, le plus souvent pour un groupe générique, simplement gh.

Définition 1.35 (morphisme, automorphisme)**.**
*Soient deux groupes G et H. Un **morphisme** entre G et H est une application $\alpha\colon G \to H$ telle que pour tout $g, h \in G$ nous ayons $\alpha(gh) = \alpha(g)\alpha(h)$.*

*Comme d'habitude, un isomorphisme est un morphisme bijectif. Un **automorphisme** de G est un isomorphisme de G vers G lui-même.*

Lemme 1.36.
Si G est un groupe, alors $\mathrm{Aut}(G)$ est un groupe pour la composition.

Démonstration. Soient α et β des automorphismes de G. Alors nous prouvons que $\alpha \circ \beta$ est un automorphisme de G :

$$(\alpha \circ \beta)(gh) = \alpha\big(\beta(g)\beta(h)\big) = \alpha\big(\beta(g)\big)\alpha\big(\beta(h)\big) = (\alpha \circ \beta)(g)(\alpha \circ \beta)(h). \tag{1.20}$$

□

Lemme 1.37.
Si G et H sont des groupes isomorphes, alors les groupes $\mathrm{Aut}(G)$ et $\mathrm{Aut}(H)$ sont isomorphes.

Démonstration. Soit un isomorphisme $f\colon G \to H$. D'abord pour tout $\alpha \in \mathrm{Aut}(G)$, l'application $f \circ \alpha \circ f^{-1}$ est un automorphisme de H. Cela est rapidement vérifié parce que f, α et f^{-1} sont des bijections et des morphismes.

Nous pouvons donc considérer l'application

$$\begin{aligned} \psi\colon \mathrm{Aut}(G) &\to \mathrm{Aut}(H) \\ \alpha &\mapsto f \circ \alpha \circ f^{-1}. \end{aligned} \tag{1.21}$$

(i) <u>ψ **est un morphisme**</u> Soient $\alpha, \beta \in \mathrm{Aut}(G)$. Vu que f est une bijection, nous pouvons introduire $f^{-1} \circ f$ partout où ça nous plaît :

$$\psi(\alpha\beta) = f \circ \alpha \circ \beta \circ f^{-1} = f \circ \alpha \circ f^{-1} \circ f \circ \beta \circ f^{-1} = \psi(\alpha) \circ \psi(\beta). \tag{1.22}$$

(ii) <u>ψ **est injective**</u> Si $\alpha\beta \in \mathrm{Aut}(G)$ sont tels que $\psi(\alpha) = \psi(\beta)$, alors

$$f \circ \alpha \circ f^{-1} = f \circ \beta \circ f^{-1}. \tag{1.23}$$

Comme f est une bijection, cela implique que $\alpha = \beta$.

(iii) <u>ψ **est surjective**</u> Si $\alpha\colon H \to H$ est un automorphisme, alors $\sigma = \psi(f^{-1} \circ \sigma \circ f)$ où il est facile de vérifier que $f^{-1} \circ \sigma \circ f \in \mathrm{Aut}(G)$.

□

Définition 1.38 (Anneau[13])**.**
*Un **anneau**[8] est un triplet $(A, +, \cdot)$ avec les conditions*

(1) $(A, +)$ est un groupe[9] commutatif. 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.

*L'anneau $(A, +, \cdot)$ est **commutatif** si pour tout $a, b \in A$ nous avons $a \cdot b = b \cdot a$.*

Définition 1.39 (Morphisme d'anneaux[14])**.**
*Si $(A, +, \cdot)$ et $(B, +, \cdot)$ sont des anneaux, un **morphisme d'anneaux** est une application $f\colon A \to B$ 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.

1.3 Les naturels

Définition 1.40 ([15])**.**
*Un **triplet naturel** est un triplet $(\mathcal{N}, o, s)$ où $\mathcal{N}$ est un ensemble, o est un élément de $\mathcal{N}$ et s est une application $s\colon \mathcal{N} \to \mathcal{N}$ satisfaisant les propriétés suivantes :*

(1) s est injective,

(2) $s(\mathcal{N}) = \mathcal{N} \backslash \{o\}$

(3) Si $A \subset \mathcal{N}$ est tel que $o \in A$ et $s(A) \subset A$, alors $A = \mathcal{N}$.

Le théorème suivant est typiquement de ceux qui vont demander de gratter la théorie axiomatique des ensembles avec une certaine précision[10].

8. Nous faisons le choix qu'un anneau admet toujours un neutre pour la multiplication. Certains ouvrages parlent dans ce cas d'anneau unitaire.

9. Groupe, définition 1.34.

10. Ou alors il y a quelque chose qui m'échappe. Écrivez-moi si vous connaissez une construction « simple ».

Théorème 1.41.
Il existe un[11] *triplet naturel.*

1.42 (Définition de $\mathbb{N}$)**.**
Pour la suite, nous considérons un triplet naturel $(\mathcal{N}, o, s)$ et nous notons $\mathbb{N} = \mathcal{N}$. Donc la nature de tous les objets que nous allons considérer à partir de maintenant dépend du choix de triplet naturel que nous faisons à présent. Le théorème 1.79 nous assurera que peu de choses devraient réellement dépendre de ce choix.

Nous notons 0 l'élément o et 1 l'élément $s(0)$. C'est tout ce dont nous avons besoin dans l'immédiat.

1.3.1 Applications définies par récurrence

Proposition 1.43 (Récurrence[15])**.**
Soit un triplet naturel $(\mathcal{N}, o, s)$ et une application $P\colon \mathcal{N} \to \{0,1\}$ vérifiant[12]

(1) $P(o) = 1$,

(2) pour tout $a \in \mathcal{N}$, si $P(a) = 1$, alors $P\big(s(a)\big) = 1$.

Alors $P(x) = 1$ pour tout $x \in \mathcal{N}$.

Démonstration. Nous posons

$$A = \{x \in \mathcal{N} \text{ tel que } P(x) = 1\}. \tag{1.24}$$

Cet ensemble vérifie la propriété 1.40(3). Donc $A = \mathcal{N}$. □

Théorème 1.44 ([11, 16])**.**
Soient E un ensemble, g une application de E dans E et b un élément de E. Alors il existe une unique application $f\colon \mathbb{N} \to E$ telle que :

(1) $f(0) = b$

(2) $f\big(s(n)\big) = g\big(f(n)\big)$ *pour tout* $n \in \mathbb{N}\backslash\{0\}$.

Démonstration. Nous commençons par l'unicité. Soient f_1 et f_2 deux telles applications. Nous posons

$$A = \{n \in \mathbb{N} \text{ tel que } f_1(n) = f_2(n)\}. \tag{1.25}$$

Nous avons $0 \in A$ parce que $f_1(0) = f_2(0) = b$.

Supposons que $f_1(k) = f_2(k)$. Alors nous avons

$$f_1\big(s(k)\big) = g\big(f_1(k)\big) = g\big(f_2(k)\big) = f_2\big(s(k)\big). \tag{1.26}$$

Nous en déduisons que $s(k) \in A$. Autrement dit $s(A) \subset A$. La définition 1.40(3) nous indique alors que $A = \mathbb{N}$, c'est-à-dire que $f_1 = f_2$.

Nous montrons à présent l'existence en plusieurs étapes.

(i) **L'ensemble est assez grand** Nous considérons l'ensemble $\mathcal{A}$ des parties $A \subset \mathbb{N} \times E$ telles que

(1) $(0, b) \in A$

(2) $(n, x) \in A \Rightarrow \big(s(n), g(x)\big) \in A$.

L'ensemble $\mathcal{A}$ est non vide parce que $\mathbb{N} \times E \in \mathcal{A}$.

11. Nous verrons plus tard que toute partie infinie d'un triplet naturel fournit un nouveau triplet naturel ; il en existe donc plusieurs.

12. Les plus pointilleux diront que 1 n'est pas encore défini. Bon j'avoue. Ce qui est important est que P prenne ses valeurs dans un ensemble contenant deux éléments distincts. Si maintenant vous râlez parce que « deux » est encore moins défini, prenez un ensemble quelconque A et dites que P prend ses valeurs dans $\{A, \mathcal{P}(A)\}$. Mais êtes-vous bien certain que $\mathcal{P}(A) \neq A$?

(ii) **Le plus petit** Nous posons

$$G = \bigcap_{A \in \mathcal{A}} A, \tag{1.27}$$

et nous prouvons que $G \in \mathcal{A}$. D'abord $(0, b) \in G$ parce que cet élément est dans chacun des $A \in G$. Ensuite si $(n, x) \in G$, alors pour tout $A \in \mathcal{A}$ nous avons $(n, x) \in A$ et donc $\big(s(n), g(x)\big) \in A$. Par conséquent $\big(s(n), g(x)\big) \in \bigcup_{A \in \mathcal{A}} A = G$.

Pour $n \in \mathbb{N}$ nous posons

$$G_n = \{x \in E \text{ tel que } (n, x) \in G\}. \tag{1.28}$$

Nous avons en particulier que $b \in G_0$ parce que $(0, b) \in G$.

(iii) **G contient un (n, x) pour tout n** Nous prouvons que pour tout $n \in \mathbb{N}$, il existe $x \in E$ tel que $(n, x) \in G$. Nous faisons ça avec la proposition 1.43 en posant

$$\begin{aligned} P\colon \mathbb{N} &\to \{0, 1\} \\ n &\mapsto \begin{cases} 1 & \text{si } G_n \neq \varnothing \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.29}$$

Puisque $(0, b) \in G$ nous avons $P(0) = 1$. Supposons que $P(k) = 1$ et montrons que $P\big(s(k)\big) = 1$. Comme $P(k) = 1$, il existe $x \in E$ tel que $(k, x) \in G$. De ce fait, $\big(s(k), g(x)\big) \in G$, ce qui donne $G_{s(k)} \neq \varnothing$ et $P\big(s(k)\big) = 1$.

(iv) **G_n est un singleton** Nous avons vu que G_n n'est jamais vide. Nous allons montrer que G_n est un singleton pour tout n. Pour cela nous posons

$$\begin{aligned} P\colon \mathbb{N} &\to \{0, 1\} \\ n &\mapsto \begin{cases} 1 & \text{si } G_n \text{ est un singleton} \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.30}$$

Nous prouvons par récurrence que $P(n) = 1$ pour tout n.

(i) **$P(0) = 1$** Nous commençons par prouver que $P(0) = 1$. Nous savons que $(0, b) \in G_0$. Supposons $a \neq b$ tel que $(0, a) \in G_0$. Alors en posant $G' = G \backslash \{(0, a)\}$ nous avons $G' \in \mathcal{A}$.

En effet $(0, b) \in G'$ parce que $(0, b) \in G$ et $(0, b) \neq (0, a)$. De plus si $(n, x) \in G'$, alors $\big(s(n), g(x)\big) \in G$. Mais comme $s(n) \neq 0$ nous avons $\big(s(n), g(x)\big) \neq (0, a)$ et donc $\big(s(n), g(x)\big) \in G'$.

L'ensemble G' serait un élément de $\mathcal{A}$ strictement inclus dans G. Impossible. Donc G_0 est un singleton.

(ii) **Récurrence** Supposons que $P(k) = 1$, c'est-à-dire que G_k est un singleton. Soit e l'unique élément de G_k : $(k, e) \in G$. Nous avons alors aussi que $\big(s(k), g(e)\big) \in G$. Nous devons prouver que si $y \in G_{s(k)}$, alors $y = g(e)$.

Supposons donc $y \neq g(e)$ soit dans $G_{s(k)}$. Nous posons

$$G' = G \backslash \{\big(s(k), y\big)\}. \tag{1.31}$$

Nous prouvons que $G' \in \mathcal{A}$. D'abord $(0, b) \in G'$ parce que $s(k) \neq 0$. Soit ensuite $(m, z) \in G'$. Si $m = k$, alors $z = e$ (parce que par hypothèse G_k est un singleton) et nous savons que $\big(s(m), g(e)\big) \in G'$. Si par contre $m \neq k$, comme s est injective, nous avons aussi $s(m) \neq s(k)$. Donc $\big(s(m), g(z)\big) \neq \big(s(k), y\big)$ et $\big(s(m), g(z)\big) \in G'$. Donc $G' \in \mathcal{A}$ et est strictement plus petit que G. Contradiction.

Nous concluons que $G_{s(k)}$ est un singleton, c'est-à-dire que $P\big(s(k)\big) = 1$.

(iii) **Conclusion** Nous avons prouvé que G_n est un singleton pour tout n.

(v) **Et enfin** Nous définissons $f(n)$ comme étant l'unique élément de G_n. Puisque $(0, b) \in G$ nous avons $G_0 = \{b\}$ et donc $g(0) = b$.

Par définition de f, nous avons $\big(n, f(n)\big) \in G$. Parce que $G \in \mathcal{A}$ nous avons alors

$$\big(s(n), g\big(f(n)\big)\big) \in G. \tag{1.32}$$

Autrement dit, $G_{s(n)} = \{g\big(f(n)\big)\}$. Cela montre que

$$f\big(s(n)\big) = g\big(f(n)\big), \tag{1.33}$$

et donc que f vérifie les propriétés demandées.

□

Remarque 1.45.
Pour faire une récurrence dont chaque élément dépend de tous les précédents (et non seulement du dernier), il faut un peu adapter. Voir 1.95 pour un exemple dans $\mathbb{N}$.

Corolaire 1.46 ([11])**.**
Soient deux ensembles X, Y, une application $\alpha\colon X \to Y$ et une application $\beta\colon Y \to Y$. Alors il existe une unique application $H\colon X \times \mathbb{N} \to Y$ telle que

(1) $H(x, 0) = \alpha(x)$ pour tout élément $x \in X$;

(2) $H(x, n+1) = \beta(H(x, n))$ pour tout élément $x \in X$ et pour tout $n \in \mathbb{N}$.

Démonstration. Pour faire le lien avec les notations du théorème 1.44, nous notons $E = \mathrm{Fun}(X, Y)$, $b = \alpha \in E$ et

$$\begin{aligned} g\colon E &\to E \\ s &\mapsto \beta \circ s. \end{aligned} \tag{1.34}$$

Le théorème 1.44 donne alors l'existence d'une application $f\colon \mathbb{N} \to E$ telle que

(1) $f(0) = b$

(2) $f(n+1) = g\big(f(n)\big)$.

Nous définissons alors

$$\begin{aligned} H\colon X \times \mathbb{N} &\to Y \\ (x, n) &\mapsto f(n)x, \end{aligned} \tag{1.35}$$

et nous vérifions qu'elle satisfait aux exigences.

(1) D'abord nous avons

$$H(x, 0) = f(0)x = b(x) = \alpha(x). \tag{1.36}$$

(2) Ensuite, pour $x \in X$ et $n \in \mathbb{N}$ nous avons :

$$\begin{aligned} H(x, n+1) &= f(x+1)x && (1.37a) \\ &= g\big(f(n)\big)x && (1.37b) \\ &= g\big(f(n)\big)x && (1.37c) \\ &= \big(\beta \circ f(n)\big)x && (1.37d) \\ &= \beta\big(f(n)x\big) && (1.37e) \\ &= \beta\big(H(x, n)\big). && (1.37f) \end{aligned}$$

Et voilà. □

1.3.2 Addition sur les naturels

Définition 1.47 (élément régulier[17])**.**
*Soit un ensemble E muni d'une opération $*\colon E \times E \to E$. Un élément $s \in E$ est **régulier à gauche** si pour tout $x, y \in E$ nous avons*

$$s * x = s * y \Rightarrow x = y. \tag{1.38}$$

L'élément s est régulier à droite si pour tout $x, y \in E$ nous avons

$$x * s = y * s \Rightarrow x = y. \tag{1.39}$$

*Il est **régulier** si il est régulier à gauche et à droite.*

Proposition-Définition 1.48 ([15, 1]).
Il existe une unique fonction $f\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N}$ vérifiant

(1) $f(a,0)=a$ *pour tout* $a\in\mathbb{N}$

(2) $f\big(a,s(b)\big)=s\big(f(a,b)\big)$ *pour tout* $a,b\in\mathbb{N}$.

Pour $a,b\in\mathbb{N}$ *nous notons* $f(a,b)=a+b$.

Lemme 1.49 ([15]).
Pour tout $a\in\mathbb{N}$ *nous avons* $s(a)=a+1$.

Démonstration. Nous avons :

$$\begin{aligned} s(a) &= s(a+0) && (1.40a)\\ &= a+s(0) && (1.40b)\\ &= a+1. && (1.40c) \end{aligned}$$

Justifications.

(1) Pour (1.40a) c'est dans la définition 1.48(1) de la somme.

(2) Pour (1.40b), c'est dans la définition 1.48(2) de la somme.

(3) Pour (1.40c). Le symbole « 1 » désigne l'élément $s(0)$ dans $\mathbb{N}$.

□

Proposition 1.50 ([15]).
En ce qui concerne la somme dans $\mathbb{N}$.

(1) La somme est associative et commutative.

(2) L'élément 0 *est neutre.*

(3) Tous les éléments de $\mathbb{N}$ *sont réguliers*[13] *par rapport à l'addition.*

Démonstration. En plusieurs parties.

(i) **Associative** Nous devons prouver que $(a+b)+c=a+(b+c)$ pour tout $a,b,c\in\mathbb{N}$. Pour ce faire, nous fixons $a,b\in\mathbb{N}$ et nous prouvons l'égalité demandée par récurrence sur c.
Pour $c=0$, nous avons $(a+b)+c=a+b$ et $a+(b+c)=a+b$. Donc nous sommes d'accord[14].
Nous vérifions avec $s(c)$:

$$\begin{aligned} (a+b)+s(c) &= s\big((a+b)+c\big) && (1.41a)\\ &= s\big(a+(b+c)\big) && (1.41b)\\ &= a+s(b+c) && (1.41c)\\ &= a+\big(b+s(c)\big). && (1.41d) \end{aligned}$$

Justifications.

— Pour (1.41b). C'est l'hypothèse de récurrence. À ce stade, je vous conseille d'être capable de rédiger complètement la récurrence et l'appel au théorème 1.44.

(ii) **Neutre** La définition de l'addition contient déjà $a+0=a$. Nous prouvons par récurrence que $0+a=a$ pour tout $a\in\mathbb{N}$.
Pour $a=0$, l'égalité demandé est correcte : $0+0=0$ parce que pour tout x dans $\mathbb{N}$, $0+x=x$.
Pour $s(a)$ nous avons

$$0+s(a)=s(0+a)=s(a). \tag{1.42}$$

La dernière égalité est l'hypothèse de récurrence.

13. Élément régulier pour une opération, définition 1.47.

14. Notez que nous n'avons pas utilisé le fait que 0 était neutre des deux côtés – chose que nous n'avons pas encore démontré. Nous avons seulement utilisé $a+0=a$, qui est dans la définition de la somme.

Nous posons

$$A = \{a \in \mathbb{N} \text{ tel que } 0 + a = a\}. \tag{1.43}$$

Nous avons prouvé que $0 \in A$ et que $s(A) \subset A$. Le théorème 1.44 nous assure alors que $A = \mathbb{N}$.

(iii) **Commutativité** Nous fixons $a \in \mathbb{N}$ et nous prouvons par récurrence sur b que $a+b = b+a$ pour tout $b \in \mathbb{N}$. Cela va être décomposé en plusieurs étapes.

(iv) $a + 0 = 0 + a$ Pour $b = 0$ c'est correct, car $b + 0 = 0 + b = b$ parce que 0 est neutre.

(v) $a + 1 = 1 + a$ Nous démontrons par récurrence sur a que $a+1 = 1+a$. Avec $a = 0$ c'est déjà fait. Pour les autres,

$$\begin{aligned}
s(a) + 1 &= (a + 1) + 1 && \text{lemme 1.49} && (1.44a)\\
&= (1 + a) + 1 && \text{hypothèse récurrence} && (1.44b)\\
&= 1 + (a + 1) && \text{associativité} && (1.44c)\\
&= 1 + s(a). && && (1.44d)
\end{aligned}$$

(vi) $a + b = b + a$ Nous y voici. Nous fixons a et nous prouvons par récurrence que $a+b = b+a$. Pour $b = 0$ c'est déjà fait. Pour les autres,

$$\begin{aligned}
a + s(b) &= a + (b + 1) && && (1.45a)\\
&= (a + b) + 1 && && (1.45b)\\
&= (b + a) + 1 && \text{hypothèse récurrence} && (1.45c)\\
&= b + (a + 1) && && (1.45d)\\
&= b + (1 + a) && \text{commutativité avec 1} && (1.45e)\\
&= (b + 1) + a && \text{associativité} && (1.45f)\\
&= s(b) + a. && && (1.45g)
\end{aligned}$$

Récurrence terminée.

(vii) **Régularité** Nous devons prouver que, pour tout $a, x, y \in \mathbb{N}$, si $a + x = a + y$ alors $x = y$. Nous allons procéder par récurrence en posant

$$A = \{a \in \mathbb{N} \text{ tel que } \forall x, y \in \mathbb{N}, a + x = a + y \Rightarrow x = y\}. \tag{1.46}$$

Puisque $0+x = x$ et $0+y = y$, nous avons $0 \in A$. Supposons à présent que $a \in A$ et montrons que $s(a) \in A$. Soient $x, y \in \mathbb{N}$ tels que $a + x = a + y$. Nous avons :

$$\begin{aligned}
& s(a) + x = s(a) + y && (1.47a)\\
\Rightarrow\ & s(a + x) = s(a + y) && (1.47b)\\
\Rightarrow\ & a + x = a + y && (1.47c)\\
\Rightarrow\ & x = y && (1.47d)
\end{aligned}$$

Justifications.

— Pour (1.47b). En utilisant la définition de l'addition et la commutativité, nous avons $s(a) + x = s(a + x)$.

— Pour (1.47c). Parce que s est injective ; c'est dans la définition 1.40 d'un triplet naturel.

— Pour (1.47d). Parce que $a \in A$.

Nous avons prouvé que $s(A) \subset A$, et donc que $A = \mathbb{N}$.

□

Lemme 1.51.
Nous avons $0 \neq 1$.

Démonstration. Par définition $1 = s(0)$. Comme s est à valeurs dans $\mathbb{N}\backslash\{0\}$, nous ne pouvons pas avoir $s(0) = 0$. □

Lemme 1.52 ([1]).
Si $a + b = 0$, alors $a = b = 0$.

Démonstration. Soient $a, b \in \mathbb{N}$ tels que $a+b = 0$, et supposons que $b \neq 0$. Par la définition 1.40(2), nous avons $b = s(c)$ pour un certain $c \in \mathbb{N}$.

Dans ce cas nous avons $a + b = a + s(c) = s(a + c) \neq 0$ parce que l'image de s ne contient pas 0. Hélas, par hypothèse nous avons $a + b = 0$. Nous avons obtenu une contradiction, et nous déduisons que $b = 0$.

Maintenant que nous savons que $b = 0$, il reste $0 = a + b = a + 0 = a$. □

1.3.3 Ordre sur les naturels

Définition 1.53 ([15]).
Pour $a, b \in \mathbb{N}$, nous notons $a \leqslant b$ si il existe $x \in \mathbb{N}$ tel que $a + x = b$.

Nous notons également $a < b$ si $a \leqslant b$ et $a \neq b$.

Lemme 1.54.
Nous avons $a \leqslant s(a)$ pour tout $a \in \mathbb{N}$.

Démonstration. Cela est une conséquence du lemme 1.49 : $s(a) = a + 1$. □

Proposition 1.55.
La relation $\leqslant$ est une relation d'ordre compatible avec l'addition.

Démonstration. Plusieurs choses à vérifier.

(i) **Réflexive** Nous avons $a \leqslant a$ parce que $a + 0 = a$.

(ii) **Antisymétrique** Soient $a, b \in \mathbb{N}$ tels que $a \leqslant b$ et $b \leqslant a$. Il existe $x, y \in \mathbb{N}$ tels que

$$b = a + x \tag{1.48a}$$
$$a = b + y. \tag{1.48b}$$

En substituant la seconde équation dans la première, $b = (b + y) + x$ que nous récrivons, en utilisant l'associativité [15],

$$0 + b = b + (x + y). \tag{1.49}$$

En utilisant la régularité, $0 = x + y$ et donc $x = y = 0$ par le lemme 1.52. Cela donne alors $a = b$.

(iii) **Transitive** Si $a \leqslant b$ et $b \leqslant c$, nous avons $n, p \in \mathbb{N}$ tels que $b = a + n$ et $c = b + p$. Donc

$$c = (a + n) + p = a + (n + p), \tag{1.50}$$

ce qui signifie que $a \leqslant c$. Notez l'utilisation de l'associativité de la somme, démontrée en la proposition 1.50(1).

(iv) **Compatibilité** Soient $a, b, n \in \mathbb{N}$ tels que $a \leqslant b$. Nous devons montrer que $a + n \leqslant b + n$. Puisque $a \leqslant b$, il existe $x \in \mathbb{N}$ tel que $b = a + x$. Par conséquent,

$$b + n = a + x + n = (a + n) + x, \tag{1.51}$$

qui signifie bien que $a + n \leqslant b + n$.

□

Lemme 1.56.
À propos d'ordre et de stricte inégalité.

15. Proposition 1.50(1).

(1) Si $x \leqslant a$ et $b \neq 0$, alors $x < a + b$.

(2) Si $x \leqslant a$, alors $x < s(a)$.

(3) Si $x < a$, alors $s(x) \leqslant a$.

Démonstration. En plusieurs parties.

(i) **Pour (1)** Si $x \leqslant a$, il existe $d \in \mathbb{N}$ tel que $x + d = a$. Nous avons alors aussi

$$x + d + b = a + b, \tag{1.52}$$

ce qui signifie que $x \leqslant a + b$. Mais si x était égal à $a + b$, nous aurions $d + b = 0$, ce qui impliquerait [16] $d = b = 0$, alors que l'hypothèse stipule que $b \neq 0$. Donc $x \neq a + b$.

(ii) **Pour (2)** Il s'agit seulement d'utiliser la point (1) avec $b = 1$ et le fait que $s(a) = a + 1$ par le lemme 1.49.

(iii) **Pour (3)** Par hypothèse, il existe $b \neq 0$ tel que $x + b = a$. Puisque $b \neq 0$, il existe $c \in \mathbb{N}$ tel que $b = s(c)$ et donc, tel que

$$x + s(c) = a. \tag{1.53}$$

En utilisant le fait que $s(c) = c + 1$ ainsi que l'associativité et la commutativité de l'addition (proposition 1.50(1)) nous avons

$$a = x + s(c) = s(x) + c, \tag{1.54}$$

ce qui prouve que $s(x) \leqslant a$.

□

Lemme 1.57.
L'élément 0 est l'unique plus petit élément de $\mathbb{N}$.

Démonstration. Puisque 0 est neutre pour l'addition [17], nous avons $a + 0 = a$ pour tout $a \in \mathbb{N}$ et donc $0 \leqslant a$ pour tout a. Cela veut dire que 0 est plus petit que tout élément de $\mathbb{N}$.

En ce qui concerne l'unicité, soit $z \in \mathbb{N}$ tel que $z \leqslant a$ pour tout $a \in \mathbb{N}$. Si $z \neq 0$, il existe $x \in \mathbb{N}$ tel que $z = s(x)$. Nous avons donc $z \geqslant x$ en même temps que $x \leqslant z$. Cela implique $z = x$ (parce qu'une relation d'ordre est symétrique) et donc $z = z + 1$. En utilisant la régularité de z pour l'addition nous en déduisons que $0 = 1$, ce qui est impossible par le lemme 1.51. □

Lemme 1.58.
Si $a \leqslant b$ et $b \leqslant a$, alors $a = b$.

Démonstration. L'inégalité $a \leqslant b$ dit qu'il existe $x \in \mathbb{N}$ tel que $a + x = b$. En mettant cela dans l'inégalité $b \leqslant a$ nous trouvons $a + x \leqslant a$ qui donne, via la proposition 1.55 : $x \leqslant 0$. Nous en déduisons que $x = 0$ parce que zéro est l'unique minimum de $\mathbb{N}$ par le lemme 1.57. □

Proposition 1.59.
Le couple $(\mathbb{N}, \leqslant)$ est totalement ordonné [18].

Démonstration. Soit $a \in \mathbb{N}$. Nous devons prouver que pour tout $x \in \mathbb{N}$ nous avons $x \leqslant a$ ou $a \leqslant x$ (non exclusifs). Nous posons

$$A = \{x \in \mathbb{N} \text{ tel que } x \leqslant a\} \tag{1.55a}$$

$$B = \{x \in \mathbb{N} \text{ tel que } a \leqslant x\}, \tag{1.55b}$$

et nous prouvons que $A \cup B = \mathbb{N}$ en montrant que $0 \in A \cup B$ et que $s(A \cup B) \subset A \cup B$.

Nous avons $0 \in A \subset A \cup B$ par le lemme 1.57.

Pour étudier $s(A \cup B)$, nous considérons $x \in A \cup B$ et nous subdivisons en deux cas selon que $x \in A$ ou $x \in B$.

16. Par le lemme 1.52.
17. Proposition 1.50(2).
18. Définition 1.11.

(i) **Si $x \in B$** Si $x \in B$, alors $a \leqslant x \leqslant s(x)$ parce que $x \leqslant s(x)$ par le lemme 1.54. Donc $s(x) \in B \subset A \cup B$.

(ii) **Si $x \in A$** Si $x \in A$, il y a deux possibilités : $x = a$ et $x \neq a$. Si $x = a$, alors $a \leqslant s(x)$ et donc $s(x) \in A \subset A \cup B$.

Si $x \neq a$, alors le lemme 1.56(3) nous indique que $s(x) \leqslant a$ et donc $s(x) \in A \subset A \cup B$.

Nous avons donc prouvé que $s(A \cup B) \subset A \cup B$, et donc que $A \cup B = \mathbb{N}$. □

Proposition 1.60 ([15, 1]).
L'ensemble ordonné $(\mathbb{N}, \leqslant)$ vérifie les propriétés suivantes.

(1) L'élément 0 est l'unique minimum de $\mathbb{N}$.

(2) Toute partie non vide a un unique plus petit élément.

(3) L'ensemble $\mathbb{N}$ n'a pas de plus grand élément.

(4) Toute partie non vide majorée a un unique plus grand élément.

Démonstration. Point par point.

(i) **Pour (1)** C'est le lemme 1.57.

(ii) **Pour (2)** Soit une partie A non vide dans $\mathbb{N}$. Si $0 \in A$, nous avons fini.

(i) **L'ensemble B** Nous supposons donc que A ne contient pas zéro et nous définissons

$$B = \{n \in \mathbb{N} \backslash A \text{ tel que } n \leqslant a, \forall a \in A\}. \tag{1.56}$$

(ii) **Un élément particulier dans B** L'ensemble B vérifie :

— $0 \in B$

— $B \neq \mathbb{N}$ parce que A est non vide.

La contraposée de la condition (3) de la définition 1.40 d'un triplet naturel implique que $s(B) \not\subset B$. Autrement dit, il existe $b \in B$ tel que $s(b) \notin B$.

(iii) **Deux fonctions sur A** Puisque $b \in B$, nous avons une application $c\colon A \to \mathbb{N}$ telle que $b + c(a) = a$. Nous avons $c(a) \neq 0$ parce que $c(a) = 0$ signifierait $b = a$, ce qui est impossible parce que $a \in A$ et $b \in B$.

Comme pour tout $a \in A$, l'élément $c(a)$ est non nul, il existe une fonction $d\colon A \to \mathbb{N}$ telle que $c(a) = d(a) + 1$.

(iv) $s(b) \in A$ Supposons que $s(b) \notin A$. Alors il existe $a \in A$ tel que $s(b) \leqslant a$ est faux. Puisque l'ordre est total (proposition 1.59), nous avons

$$a \leqslant s(b). \tag{1.57}$$

Comme $b \in B$ nous avons aussi

$$b \leqslant a. \tag{1.58}$$

Et enfin nous avons

$$b \neq a \tag{1.59}$$

parce que $a \in A$ et $b \in B$.

Les conditions (1.58) et (1.59) se résument en $b < a$. Le lemme 1.56(3) nous indique alors que $s(b) \leqslant a$. Cela mis à côté de (1.57) conclut que $a = s(b)$, et donc que $s(b)$ est dans A. Contradiction. Nous en concluons que $s(b) \in A$.

(v) **$s(b)$ est un minimum de A** En utilisant la commutativité et l'associativité de la somme nous avons, pour tout $a \in A$:

$$a = b + c(a) = b + \big(d(a) + 1\big) = (b + 1) + d(a) = s(b) + d(a). \tag{1.60}$$

Donc $s(b) \leqslant a$ pour tout $a \in A$. Mais comme $s(b) \in A$, l'élément $s(b)$ est bien un minimum de A.

(vi) **Unicité** Si a et a' sont des minimums de A, alors $a \leqslant a'$ et $a' \leqslant a$. Nous en déduisons que $a = a'$.

(iii) **Pour (3)** Si $M \in \mathbb{N}$ majore tous les éléments de $\mathbb{N}$, alors en particulier $M \geqslant s(M)$. Mais le lemme 1.54 nous indique que $M \leqslant s(M)$. Nous avons donc $s(M) = M$, c'est-à-dire $M = M + 1$. En utilisant la régularité de M [19], nous trouvons $0 = 1$, ce qui est impossible par le lemme 1.51.

(iv) **Pour (4)** Soit A une partie non vide et majorée de $\mathbb{N}$.

(i) **L'ensemble B** Nous posons

$$B = \{n \in \mathbb{N} \backslash A \text{ tel que } a \leqslant n, \forall a \in A\}. \tag{1.61}$$

(ii) **B est non vide** Soit un majorant M de A : pour tout $a \in A$ nous avons $a \leqslant M$. Nous avons $s(M) \notin A$, parce que si $s(M)$ était dans A, ce serait un élément de A strictement plus grand que tout $a \in A$. Donc B est non vide parce qu'il contient $s(M)$.

(v) **Minimum** Puisque B est non vide, il possède un plus petit élément que nous notons b. Nous savons que $b \neq 0$ parce que sinon A serait vide. Il existe donc $c \in \mathbb{N}$ tel que $b = s(c)$.

(vi) **$a \leqslant c$ pour tout $a \in A$** Comme $s(c) \in B$ nous avons $a < s(c)$ pour tout $a \in A$. Donc, par le lemme 1.56(3) nous avons $s(a) \leqslant s(c)$, c'est-à-dire $a + 1 \leqslant c + 1$. Par régularité nous avons $a \leqslant c$.

Nous avons prouvé que $a \leqslant c$ pour tout $a \in A$.

(vii) **$c \in A$** Si c n'est pas dans A, alors il est dans B et il contredit la minimalité de b. Donc c est dans A.

(viii) **Conclusion** L'élément c est dans A tout en étant plus petit que tout élément de A.

(ix) **Unicité** Si x est un élément minium de A, alors nous avons $x \leqslant c$ parce que x est minimum et $c \leqslant x$ parce que c est minimum, et donc $x = c$.

□

Lemme 1.61.
Toute partie finie non vide de $\mathbb{N}$ est majorée et minorée.

Lemme-Définition 1.62.
Si A est une partie de $\mathbb{N}$, il existe un unique élément $m \in \mathbb{N}$ tel que

$$\begin{cases} m \in A & (1.62a) \\ m \leqslant a \, \forall a \in A. & (1.62b) \end{cases}$$

Cet élément est noté $\min(A)$ *et nommé* ***minimum de*** *A.*

Si A est majoré, il existe un unique élément $M \in \mathbb{N}$ tel que

$$\begin{cases} M \in A & (1.63a) \\ M \geqslant a \, \forall a \in A. & (1.63b) \end{cases}$$

Cet élément est noté $\max(A)$ *et nommé* ***maximum de*** *A.*

Nous verrons dans le lemme 1.70 qu'une partie de $\mathbb{N}$ admet un maximum si et seulement si elle est finie.

Lemme 1.63 ([1])**.**
Quelques affirmations sur l'ordre dans $\mathbb{N}$.

(1) Il n'existe pas de $n \in \mathbb{N}$ tel que $n < 0$.

(2) Si $a, b \in \mathbb{N}$ vérifient $a > b$, alors il n'existe pas de x dans $\mathbb{N}$ tel que $a + x = b$.

19. Dit plus simplement : en simplifiant par M.

Démonstration. En deux parties.

(i) **Pour (1)** Nous savons par la proposition 1.60(1) que 0 est l'unique minimum de $\mathbb{N}$. Nous avons donc forcément $0 \leqslant n$. Si n vérifie de plus $n \leqslant 0$ alors nous avons $n = 0$ par symétrie de la relation d'ordre $\leqslant$. Il n'est donc pas possible d'avoir $n \neq 0$.

(ii) **Pour (2)** Si $b \leqslant a$ il existe $c \in \mathbb{N}$ tel que $b + c = a$. Et comme $a \neq b$, c n'est pas nul et il existe $y \in \mathbb{N}$ tel que $c = s(y)$. Bref, nous avons

$$b + s(y) = a. \tag{1.64}$$

Si de plus il existe $x \in \mathbb{N}$ tel que $a + x = b$ nous aurions

$$a + x + s(y) = a. \tag{1.65}$$

Comme a est régulier pour l'addition [20], nous avons

$$x + s(y) = 0, \tag{1.66}$$

ce qui signifie, par le lemme 1.52 que $x = s(y) = 0$. Puisque s prend ses valeurs dans $\mathbb{N}\backslash\{0\}$, cela est impossible.

□

Définition 1.64.
Soient $a, b \in \mathbb{N}$ tels que $a \leqslant b$. Nous notons par $\{a, \dots, b\}$ l'ensemble

$$\{x \in \mathbb{N} \text{ tel que } a \leqslant x \leqslant b\}. \tag{1.67}$$

Proposition 1.65.
Toute application $\mathbb{N} \to \mathbb{N}$ strictement croissante est injective.

Démonstration. Soit une application strictement croissante $f\colon \mathbb{N} \to \mathbb{N}$. Soient $a, b \in \mathbb{N}$ tels que $f(a) = f(b)$. Puisque l'ordre est total [21], nous supposons que $a \leqslant b$. Si $a = b$ nous avons terminé. Nous supposons donc que $a \neq b$, c'est-à-dire que $a < b$. Par stricte croissance nous avons alors $f(a) < f(b)$ qui signifie $f(a) \leqslant f(b)$ et $f(a) \neq f(b)$. Contradiction. Il n'existe donc pas de $a \neq b$ tels que $f(a) = f(b)$. L'application f est donc injective. □

Lemme 1.66 ([1])**.**
Si S n'est pas majoré dans $\mathbb{N}$, alors il existe une bijection $\mathbb{N} \to S$.

Démonstration. Nous considérons l'application suivante :

$$\begin{aligned} g\colon S &\to S \\ n &\mapsto \min\{x \in S \text{ tel que } x > n\}. \end{aligned} \tag{1.68}$$

Cette application est bien définie parce que tout partie non vide de $\mathbb{N}$ a un plus petit élément [22]. Maintenant nous définissons $f\colon \mathbb{N} \to S$ par

$$\begin{cases} f(0) = \min(S) & \text{(1.69a)} \\ f(n+1) = g\big(f(n)\big). & \text{(1.69b)} \end{cases}$$

C'est le théorème 1.44 qui nous permet de le faire. Nous montrons que f est bijective.

(i) **Injective** Nous avons

$$f(n+1) \in \{x \in S \text{ tel que } x > f(n)\}. \tag{1.70}$$

Donc f est strictement croissante. Elle est donc injective.

20. Proposition 1.50(3).
21. Proposition 1.59.
22. Proposition 1.60(2).

(ii) **Surjective** Soit $a \in S$. Nous allons voir que a est dans l'image de f. Pour cela nous posons

$$A = \{x \in \mathbb{N} \text{ tel que } f(x) < a\}. \tag{1.71}$$

Cet ensemble est majoré par a. En effet si $x \in A$ nous avons $x \leqslant f(x) < a$. La partie A de $\mathbb{N}$ possède un maximum. Nous notons $M = \max(A)$. Ce M a deux propriétés intéressantes.

(i) **D'abord** Puisque $M \in A$, nous avons $f(M) < a$. Une autre façon de dire cela est de dire que

$$a \in \{x \in S \text{ tel que } x > f(M)\}. \tag{1.72}$$

Or $f(M+1) = \min\{x \in S \text{ tel que } x > f(M)\}$. Donc $f(M+1) \leqslant a$.

(ii) **Ensuite** Puisque M est le maximum de A, $M+1$ majore A, c'est-à-dire que $f(M+1) \geqslant a$.

(iii) **Les deux ensemble** Nous avons prouvé que $f(M+1) \leqslant a$ et $f(M+1) \geqslant a$. Nous en déduisons, par le lemme 1.58, que $f(M+1) = a$.

□

1.67.
Durant la preuve du lemme 1.66, nous n'avons pas été loin de prouver que

$$\big(\min(S), S, g\big) \tag{1.73}$$

est un triplet naturel.

Toute partie non bornée de $\mathbb{N}$ donne lieu à un triplet naturel.

Définition 1.68 ([1])**.**
Soit un ensemble muni d'une loi de composition interne $(A, +)$. Soit $n \in \mathbb{N}$ et $a \in A$. Nous définissons $n \times A$ par

$$\begin{cases} 0 \times a = 0 & (1.74a) \\ (n+1) \times a = n \times a + a. & (1.74b) \end{cases}$$

Définition 1.69.
*Un ensemble totalement ordonné muni d'une loi de composition interne $(A, +, \leqslant)$ est **archimédien** si pour tout $x, y \in A$ avec $x > 0$, il existe $n \in \mathbb{N}$ tel que $n \times x \geqslant y$ (voir la définition 1.68).*

Lemme 1.70.
Une partie de $\mathbb{N}$ admet un maximum si et seulement si elle est finie.

1.3.4 Multiplication dans les naturels

Proposition-Définition 1.71.
Il existe une unique fonction $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ telle que

(1) $f(a, 0) = 0$ pour tout $a \in \mathbb{N}$

(2) $f\big(a, s(b)\big) = f(a, b) + a$ pour tout $a, b \in \mathbb{N}$.

*Cette fonction est la **multiplication** et nous notons $f(a, b) = a \times b$, voire ab quand il n'y a pas d'ambigüité. Le nombre $a \times b$ est nommé le **produit** de a par b.*

Démonstration. En deux parties.

(i) **Fonctions définies par récurrence** Soit $a \in \mathbb{N}$. Par le théorème 1.44, il existe une unique application $f_a\colon \mathbb{N} \to \mathbb{N}$ telle que

$$\begin{cases} f_a(0) = 0 & (1.75a) \\ f_a\big(s(b)\big) = f_a(b) + a & (1.75b) \end{cases}$$

(ii) **Existence** Nous considérons, pour chaque $a \in \mathbb{N}$ la fonction f_a définie par les conditions (1.75). En posant $f(a,b) = f_a(b)$, nous avons une application qui vérifie toutes les conditions.

(iii) **Unicité** Soient des applications f et g vérifiant les propriétés demandées. Soit $a \in \mathbb{N}$. Nous pouvons définir $f_a \colon \mathbb{N} \to \mathbb{N}$ et $g_a \colon \mathbb{N} \to \mathbb{N}$ par $f_a(n) = f(a,b)$ et $g_a(b) = g(a,b)$.
Les applications f_a et g_a vérifient toutes deux les conditions (1.75), et sont donc égales : $f_a = g_a$ pour tout a. Donc $f = g$.

□

1.72.
Nous supposons que le lecteur connait déjà la priorité des opérations. Il saura donc interpréter des expressions comme $a \times b + c$ comme voulant dire $(a \times b) + c$ sans que nous ayons à ajouter des parenthèses.

Proposition 1.73 ([15, 1]).
La multiplication a les propriétés suivantes.

(1) $n \times 1 = n$ *pour tout* $n \in \mathbb{N}$.

(2) $1 \times n = n$ *pour tout* $n \in \mathbb{N}$.

(3) La multiplication est commutative.

(4) $0 \times n = 0$ *pour tout* $n \in \mathbb{N}$

(5) L'élément 1 *est neutre pour la multiplication.*

(6) La multiplication est distributive par rapport à l'addition.

(7) La multiplication est associative.

Démonstration. Point par point.

(i) **Pour (1)** Nous avons $n \times 1 = n \times s(0) = (n \times 0) + n = 0 + n = n$. Donc $n \times 1 = n$.

(ii) **Pour (2)** Nous le faisons par récurrence. Par définition c'est vrai pour $n = 0$. En ce qui concerne la récurrence, nous supposons que $1 \times n = n$, et nous prouvons que $1 \times s(n) = s(n)$:

$$1 \times s(n) = (1 \times n) + 1 = n + 1 = s(n). \tag{1.76}$$

(iii) **Pour (3)** Soit $a \in \mathbb{N}$. Nous prouvons par récurrence sur $b \in \mathbb{N}$ que $a \times b = b \times a$. Pour $b = 0$ c'est bon. Pour la récurrence, nous supposons que $a \times b = b \times a$ et nous prouvons que $a \times s(b) = s(b) \times a$:

$$\begin{aligned}
a \times s(b) &= a \times b + a & & \text{(1.77a)}\\
&= b \times a + a & \text{récurrence} & \text{(1.77b)}\\
&= (b \times a) + (1 \times a) & \text{par (2)} & \text{(1.77c)}\\
&= (b + 1) \times a & \text{distributivité} & \text{(1.77d)}\\
&= s(b) \times a. & & \text{(1.77e)}
\end{aligned}$$

(iv) **Pour (4)** C'est vrai pour $n = 0$ par la définition 1.71(1). En ce qui concerne $s(n)$, nous avons

$$0 \times s(n) = (0 \times n) + 0 = 0. \tag{1.78}$$

(v) **Pour (5)** C'est la combinaison de (1) et (2).

(vi) **Pour (6)** Soient $a, b \in \mathbb{N}$. Nous prouvons par récurrence sur $c \in \mathbb{N}$ que[23]

$$(a+b) \times c = a \times c + b \times c. \tag{1.79}$$

Pour $c = 0$, nous avons $(a+b) \times 0 = 0$ ainsi que $a \times 0 = b \times 0 = 0$ en vertu des points précédents sur la multiplication par zéro. Pour la récurrence nous utilisons associativité et commutativité de la somme :

$$\begin{aligned}
(a+b) \times s(c) &= (a+b) \times c + (a+b) && (1.80a)\\
&= a \times c + b \times c + a + b && (1.80b)\\
&= (a \times c + a) + (b \times c + b) && (1.80c)\\
&= \big(a \times s(c)\big) + \big(b \times s(c)\big). && (1.80d)
\end{aligned}$$

(vii) **Pour (7)** Soient $a, b \in \mathbb{N}$. Nous démontrons par récurrence sur $c \in \mathbb{N}$ que $(a \times b) \times c = a \times (b \times c)$. Pour $c = 0$ l'égalité est triviale. Nous supposons que l'égalité est correcte pour c, et nous la prouvons pour $s(c)$:

$$\begin{aligned}
(a \times b) \times s(c) &= \big((a \times b) \times c\big) + a \times b && & (1.81a)\\
&= \big(a \times (b \times c)\big) + a \times b && \text{récurrence} & (1.81b)\\
&= \big((b \times c) \times a\big) + b \times a && \text{commutativité} & (1.81c)\\
&= (b \times c + b) \times a && \text{distributivité} & (1.81d)\\
&= \big(b \times s(c)\big) \times a && \text{définition 1.71(2)} & (1.81e)\\
&= a \times \big(b \times s(c)\big) && \text{commutativité.} & (1.81f)
\end{aligned}$$

□

Lemme 1.74 ([1])**.**
La multiplication est compatible avec l'ordre :

$$a \leqslant b \Rightarrow a \times n \leqslant b \times n \tag{1.82}$$

pour tout $n \in \mathbb{N}$.

Démonstration. Par la définition 1.53 de l'ordre, si $a \leqslant b$, il existe $c \in \mathbb{N}$ tel que $b = a + c$. En utilisant la distributivité[24], nous avons

$$b \times n = (a+c) \times n = a \times n + c \times n. \tag{1.83}$$

Nous en déduisons que $a \times n \leqslant b \times n$ parce que $c \times n \in \mathbb{N}$. □

Lemme 1.75.
Si $a \times b = 0$*, alors* $a = 0$ *ou* $b = 0$ *(ou les deux).*

Démonstration. Supposons que $a \neq 0$. Alors il existe $c \in \mathbb{N}$ tel que $a = s(c)$. Nous avons

$$0 = a \times b = s(c) \times b = c \times b + b. \tag{1.84}$$

Le lemme 1.52 nous dit alors que $c \times b = b = 0$. □

Lemme 1.76 ([1])**.**
Si $a < b$ *et si* $n \neq 0$*, alors*

$$a \times n < b \times n. \tag{1.85}$$

23. Nous n'écrivons pas toutes les parenthèses parce que les règles de priorité des opérations sont supposées connues. J'invite cependant la lectrice à remarquer qu'une formalisation de ces règles n'est probablement pas facile. Pour que tout soit rigoureux, il faudrait un algorithme qui parcourt une suite de caractères et l'interprète en ajoutant correctement les parenthèses.

24. Proposition 1.73(6).

Démonstration. L'hypothèse $a \leqslant b$ implique qu'il existe $c \in \mathbb{N}$ tel que $a + c = b$. De plus $c \neq 0$ parce que $a \neq b$. En utilisant la distributivité [25], nous avons

$$b \times n = (a + c) \times n = (a \times n) + (c \times n). \tag{1.86}$$

Cela prouve que $a \times n \leqslant b \times n$. Et comme c et n ne sont pas nuls, nous avons même [26] $c \times n \neq 0$ et donc $a \times n < b \times n$. □

Une version dans $\mathbb{Z}$ sera le lemme 1.105.

Lemme 1.77 ([1]).
Soient $a \neq 0$ et $b > 1$ dans $\mathbb{N}$. Alors

$$ab > a. \tag{1.87}$$

Démonstration. Il s'agit d'une application du lemme 1.76 en partant de l'inégalité $1 < b$ et en la « multipliant » par a. □

Proposition 1.78 ([1]).
Tous les naturels non nuls sont réguliers par rapport à la multiplication. Autrement dit, si $a \neq 0$, alors nous avons

$$a \times x = a \times y \Rightarrow x = y. \tag{1.88}$$

Démonstration. Soit $a \neq 0$ dans $\mathbb{N}$. Nous supposons que $a \times x = a \times y$. Puisque l'ordre sur $\mathbb{N}$ est total (proposition 1.59), nous pouvons supposer que $y \geqslant x$; sinon il suffit de permuter les rôles de x et y dans tout ce qui suit.

Il existe $d \in \mathbb{N}$ tel que $y = x + d$. En utilisant l'hypothèse $a \times y = a \times x$ et la distributivité [27],

$$a \times x = a \times y = a \times (x + d) = (a \times x) + (a \times d). \tag{1.89}$$

Puisque $(a \times x)$ est régulier pour la somme [28] nous en déduisons que

$$0 = a \times d. \tag{1.90}$$

Le lemme 1.75 dit alors que $a = 0$ ou que $d = 0$. Étant donné que $a \neq 0$ par hypothèse, nous déduisons que $d = 0$, c'est-à-dire que $x = y$. □

1.3.5 Presque unicité des triplets naturels

Il existe de nombreux triplets naturels ; l'existence d'un triplet naturel est un théorème de la théorie des ensembles que nous avons accepté. Nous avons déjà à peu près montré que toute partie non bornée de $\mathbb{N}$ donne lieu à un nouveau triplet naturel. Voir 1.67.

Nous voyons maintenant que tous les triplets naturels sont équivalents au moins pour l'ordre.

Théorème 1.79 ([1]).
Soient des triplets naturels $(\mathcal{N}_1, o_1, s_1)$ et $(\mathcal{N}_2, o_2, s_2)$. Alors

(1) il existe une unique application $f : \mathcal{N}_1 \to \mathcal{N}_2$ telle que

(1a) $f(o_1) = o_2$

(1b) $f \circ s_1 = s_2 \circ f$.

(2) Une telle application est une bijection croissante.

Démonstration. En plusieurs points.

25. Proposition 1.73(6).
26. Lemme 1.75.
27. Proposition 1.73(6).
28. Proposition 1.50(3).

(i) **Existence** Nous voyons $(\mathcal{N}_1, o_1, s_1)$ comme un triplet naturel, et $\mathcal{N}_2$ comme un simple ensemble. Nous pouvons appliquer le théorème 1.44 à $(\mathcal{N}_1, o_1, s_1)$. L'élément o_1 va jouer le rôle de 0 alors que o_2 va jouer le rôle de b. L'application g est s_2. Bref, il existe une unique application $f\colon \mathcal{N}_1 \to \mathcal{N}_2$ telle que

(1) $f(o_1) = o_2$

(2) $f\big(s_1(n)\big) = s_2\big(f(n)\big)$

pour tout $n \in \mathcal{N}_1$.

(ii) **Unicité** Le théorème 1.44 donne déjà l'unicité. Nous la faisons quand même, juste pour vous faire plaisir. Soit g, une autre application vérifiant les mêmes conditions. Pour faire la récurrence de façon très explicite, nous posons

$$\begin{aligned} P\colon \mathcal{N}_1 &\to \{0,1\} \\ x &\mapsto \begin{cases} 1 & \text{si } g(x) = f(x) \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.91}$$

Notre but est de prouver que $P(x) = 1$ pour tout $x \in \mathcal{N}_1$, en utilisant la récurrence telle que décrite dans la proposition 1.43.

Nous avons $f(o_1) = o_2 = g(o_1)$. Donc $P(o_1) = 1$. Nous supposons que, pour un certain $a \in \mathcal{N}_1$, nous ayons $P(a) = 1$, et nous prouvons que $P\big(s_1(a)\big) = 1$.

Nous avons $g(a) = f(a)$, et nous prenons s_2 des deux côtés, nous avons succéssivement

$$(s_2 \circ g)(a) = (s_2 \circ f)(a) \tag{1.92a}$$
$$(g \circ s_1)(a) = (f \circ s_1)(a) \tag{1.92b}$$
$$g\big(s_1(a)\big) = f\big(s_1(a)\big). \tag{1.92c}$$

La dernière égalité signifie que $P\big(s_1(a)\big) = 1$. La proposition 1.43 implique que $P(x) = 1$ pour tout $x \in \mathcal{N}_1$.

(iii) **Bijection, définir l'inverse** Nous allons trouver un inverse et le lemme 1.7 nous dit que c'est suffisant. La partie « existence », en inversant les rôles de $\mathcal{N}_1$ et $\mathcal{N}_2$ nous donne une application $g\colon \mathcal{N}_2 \to \mathcal{N}_1$ telle que

(1) $g(o_2) = o_1$

(2) $g \circ s_2 = s_1 \circ g$.

Nous allons prouver que g est un inverse de f.

(iv) $f \circ g = \mathrm{Id}$ Nous posons $A = \{x \in \mathcal{N}_2 \text{ tel que } (f \circ g)(x) = x\}$. Nous avons

$$f\big(g(o_2)\big) = f(o_1) = o_2, \tag{1.93}$$

et donc $o_2 \in A$.

Supposons que $x \in A$. Alors

$$(f \circ g)\big(s_2(x)\big) = (f \circ \underbrace{g \circ s_2}_{s_1 \circ g})(x) \tag{1.94a}$$
$$= (\underbrace{f \circ s_1}_{= s_2 \circ f} \circ g)(x) \tag{1.94b}$$
$$= (s_2 \circ f \circ g)(x) \tag{1.94c}$$
$$= s_2\big((f \circ g)(x)\big) \tag{1.94d}$$
$$= s_2(x) \tag{1.94e}$$

Donc $s_2(x) \in A$. Nous en déduisons que $A = \mathcal{N}_2$ par le point (3) de la définition 1.40 d'un triplet naturel.

(v) $\underline{g \circ f = \mathrm{Id}}$ J'imagine que c'est la même chose que dans l'autre sens (ci-dessus)[29].

□

Proposition 1.80.
L'ensemble structuré $(\mathbb{N}, +, \times, \leqslant)$ *est archimédien*[30]*. En d'autres termes, pour tout* $a, b \in \mathbb{N}\backslash\{0\}$*, il existe* $n \in \mathbb{N}$ *tel que*

$$b < n \times a. \tag{1.95}$$

Démonstration. Soient $a, b \in \mathbb{N}$ avec $a \neq 0$.

Si $a > b$, nous avons le résultat avec $n = 1$.

Si $a = b$, en prenant $n = s(1)$ nous avons le résultat. En effet $s(1) \times a = a + a$. Puisque $a \neq 0$, nous avons $a + a \geqslant a$ et $a + a \neq a$, donc $s(1) \times a > a$.

La vraie vie est avec $a < b$. Nous posons

$$X = \{x \in \mathbb{N} \text{ tel que } 1 \leqslant x \times a \leqslant b\} \tag{1.96}$$

et

$$B = \{x \times a \text{ tel que } x \in X\}. \tag{1.97}$$

L'ensemble X est non vide parce que $1 \in X$. L'ensemble B est alors également non vide, et majoré par b. La proposition 1.60(4) nous indique alors que B possède un plus grand élément que nous allons noter $x_0 \times a$ ($x_0 \in X$).

Nous posons $n = s(x_0)$, et nous avons

$$x_0 \times a < x_0 \times a + a = s(x_0) \times a = n \times a. \tag{1.98}$$

Nous en déduisons que $n \times a$ n'est pas dans B parce que $x_0 \times a$ est le plus grand élément de B. Donc x_0 n'est pas dans X ; nous n'avons donc pas les inégalités

$$1 \leqslant n \times a \leqslant b. \tag{1.99}$$

Laquelle des deux inégalités est fausse ? Puisque $n = s(x_0) \geqslant 1$ et que $a \geqslant 1$, nous avons $1 \leqslant n \times a$. Donc c'est la seconde inégalité qui est fausse. Nous avons donc $n \times a > b$. □

Définition 1.81.
Soit A *un ensemble muni d'une loi de composition interne*[31] *notée* $+$*. Si nous avons une application* $\alpha\colon \mathbb{N} \to A$*, alors nous définissons la notation* $\sum_{i=0}^{N} \alpha(i)$ *par récurrence de la façon suivante :*

(1) $\sum_{i=0}^{0} \alpha(i) = \alpha(0)$,

(2) $\sum_{i=0}^{k} \alpha(i) = \sum_{i=0}^{k-1} \alpha(i) + \alpha(k)$.

1.82.
Si vous êtes attentif, vous remarquerez que la définition 1.81 a besoin du théorème 1.44 pour s'assurer que $\sum_{i=0}^{N}$ est bien définie pour tout $N \in \mathbb{N}$.

Proposition 1.83 (La multiplication est une somme itérée[15])**.**
Pour tout $a, b \in \mathbb{N}$*, nous avons*

$$\sum_{i=1}^{n} a = a \times n. \tag{1.100}$$

Démonstration. Nous le faisons par récurrence en partant de $n = 1$. Avec $n = 1$ nous avons $\sum_{i=1}^{1} a = a$, et $a \times 1 = a$. Donc c'est bon.

29. Je n'ai pas fait les calculs ; écrivez-moi si ça pose un problème.
30. Définition 1.69.
31. Peut-être un anneau, mais comme nous avons l'intention, dans les propositions 1.83 et suivantes, de faire des sommes vers $(\mathbb{N}, +)$, plutôt un monoïde.

Pour la récurrence nous avons :

$$a \times s(n) = a \times n + a = \sum_{i=1}^{n} a + a = \sum_{i=1}^{n+1} a = \sum_{i=1}^{s(n)} a. \tag{1.101}$$

□

Lemme 1.84.
Soit $a > 1$ dans $\mathbb{N}$. Pour tout $n \geqslant 1$ nous avons $na \leqslant a^n$.

Démonstration. Par récurrence. Avec $n = 1$ nous avons bien $a \leqslant a$; pas de problème. Supposons que $na \leqslant a^n$, et montrons le pas de récurrence. Nous avons :

$$\begin{aligned}
(n+1)a &= na + a && (1.102a)\\
&\leqslant na + na && \text{parce que } a \leqslant na \quad (1.102b)\\
&= 2na && (1.102c)\\
&\leqslant 2a^n && \text{récurrence} \quad (1.102d)\\
&\leqslant aa^n && \text{parce que } a \geqslant 2 \quad (1.102e)\\
&= a^{n+1}. && (1.102f)
\end{aligned}$$

□

Proposition 1.85 ([15])**.**
Soit $a > 1$. Alors

(1) l'application $n \mapsto a^n$ est strictement croissante ;

(2) l'ensemble $\{a^n$ tel que $n \in \mathbb{N}\}$ n'est pas majoré.

Démonstration. Nous avons

$$\begin{aligned}
a^{n+1} &= a^n \times a && (1.103a)\\
&> a^n \times 1 && \text{lemme 1.77} \quad (1.103b)\\
&= a^n. && (1.103c)
\end{aligned}$$

Cela prouve le premier point.

Pour le second point, soit $m \in \mathbb{N}$. Nous devons trouver $N \in \mathbb{N}$ tel que $a^N \geqslant m$. Puisque $\mathbb{N}$ est archimédien [32], nous pouvons considérer N tel que $Na > m$. Le lemme 1.52 nous assure alors que

$$m < Na \leqslant a^N. \tag{1.104}$$

□

Théorème 1.86 (division euclidienne [15])**.**
Pour tout $a \in \mathbb{N}$, pour tout $b \in \mathbb{N}\backslash\{0\}$, il existe un unique couple $(q, r) \in \mathbb{N}^2$ tel que $a = bq + r$ avec $0 \leqslant r < b$.

*Si $r = 0$, nous disons que a est **divisible** par b.*

Démonstration. Existence puis unicité.

(i) **Existence** Nous posons

$$A = \{bx \text{ tel que } x \in \mathbb{N}, bx \leqslant a\}. \tag{1.105}$$

L'ensemble A contient 0 (avec $x = 0$) et est majoré par a. Donc il possède un plus grand élément que nous notons bq. Puisque $bq \in A$, nous avons $bq \leqslant a$ et donc il existe $r \in \mathbb{N}$ tel que

$$bq + r = a. \tag{1.106}$$

32. Proposition 1.80.

Il reste à montrer que $r < b$. Supposons que $r \geqslant b$. Il existerait alors un x tel que $b + x = r$. En mettant ça dans (1.106),

$$bq + b + x = a, \tag{1.107}$$

c'est-à-dire $b(q+1) + x = a$, qui signifierait $b(q+1) \leqslant a$, ce qui est faux parce que bq est le plus grand élément de A.

(ii) **Unicité** Supposons que nous ayons

$$a = bq + r = bq' + r' \tag{1.108}$$

avec $0 \leqslant r < b$ et $0 \leqslant r' < b$. Il y a trois possibilités : $q' < q$, $q' = q$ et $q' > q$.

(i) **Si $q' < q$** Alors il existe $x \in \mathbb{N}$ tel que $q' + x = q$, et nous avons

$$b(q' + x) + r = bq' + bx + r, \tag{1.109}$$

ce qui, après distribution et simplification, donne $r' = bx + r$. Puisque nous avons $x \geqslant 1$, il vient

$$r' = bx + r \geqslant b + r \geqslant b. \tag{1.110}$$

Cela n'est pas possible parce que $r' < b$. Le cas $q' < q$ n'est pas possible.

(ii) **Si $q' = q$** Nous avons alors immédiatement $bq + r = bq + r'$ et donc $r = r'$. Unicité.

(iii) **Si $q' > q$** En posant $q + x = q'$ nous trouvons la même impossibilité que dans le cas $q' < q$.

□

1.3.6 Écriture d'un naturel dans une base

1.87.
Nous avons déjà donné la notation $1 = s(0)$. Nous continuons avec $2 = s(1)$, $3 = s(2)$, $4 = s(3)$, $5 = s(4)$, $6 = s(5)$, $7 = s(6)$, $8 = s(7)$ et $9 = s(8)$.

Nous allons maintenant voir comment écrire des nombres plus grands.

Si $b > 1$ et $N \in \mathbb{N}$ sont donnés, nous notons

$$C_{b,N} = \big\{u \in \{0, \dots, b-1\}^{N+1} \text{ tel que } u_N \neq 0\big\}. \tag{1.111}$$

où les u_i sont numérotés à partir de 0 ; donc dire $u_N \neq 0$ revient à dire que le *dernier* est non nul, et non l'avant dernier. Nous définissons [33]

$$\begin{aligned} \varphi_{b,N} \colon C_{b,N} &\to \mathbb{N} \\ u &\mapsto \sum_{i=0}^{N} u_i b^i. \end{aligned} \tag{1.112}$$

Cette application $\varphi_{b,N}$ sera encore bien étudiée pour la partie décimale d'un réel. Voir la définition 11.327.

Lemme 1.88 ([15]).
Soient $b > 1$, $N \geqslant 0$ ainsi que $u \in C_{b,N}$. Alors

$$b^N \leqslant \varphi_{b,N}(u) < b^{N+1}. \tag{1.113}$$

Démonstration. En séparant la somme nous avons

$$\varphi_{b,N}(u) = u_N b^N + \sum_{i=0}^{N-1} u_i b^i. \tag{1.114}$$

Puisque $u_N \geqslant 1$ nous avons $b^N \leqslant u_N b^N$, et donc

$$b^N \leqslant u_N b^N \leqslant \varphi_{b,N}(u). \tag{1.115}$$

Voilà qui prouve la première inégalité de (1.113).

Pour prouver que $\varphi_{b,N}(u) < b^{N+1}$, nous faisons une récurrence sur N.

33. Le symbole de sommation est défini par 1.81.

(i) **Pour $N = 0$** Nous devons prouver que $\varphi_{b,0}(u) < b$. Par définition $\varphi_{b,N}(u) = u_0 b^0$. Puisque $u \in \{0, \dots, b-1\}^{N+1}$, nous avons $u_0 \leqslant b - 1 < b$.

(ii) **Récurrence** Nous supposons que pour tout $u \in C_{b,N}$ nous avons $\varphi_{b,N}(u) < b^{N+1}$. Et nous devons montrer que pour tout $v \in C_{b,N+1}$ nous avons $\varphi_{b,N+1}(v) < b^{N+2}$.

Nous posons $u = (v_0, \dots, v_N)$; nous avons alors

$$\begin{aligned}
\varphi_{b,N+1}(v) &= v_{N+1} b^{N+1} + \sum_{i=0}^{N} v_i b^i && (1.116a)\\
&= v_{N+1} b^{N+1} + \varphi_{b,N}(u) && (1.116b)\\
&< v_{N+1} b^{N+1} + b^{N+1} && \text{récurrence} \quad (1.116c)\\
&= (v_{N+1} + 1) b^{N+1} && (1.116d)\\
&\leqslant b b^{N+1} && \text{parce que } v_{N+1} \leqslant b - 1 \quad (1.116e)\\
&= b^{N+2}. && (1.116f)
\end{aligned}$$

□

Lemme 1.89 ([1]).
Soient $x \in \mathbb{N}$ ainsi que $b \geqslant 2$. Nous posons

$$N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.117}$$

Alors

(1) Si $n > N$ alors $\varphi_{b,n}(u) > x$ pour tout $u \in C_{b,n}$.

(2) Si $n < N$ alors $\varphi_{b,n}(u) < x$ pour tout $u \in C_{b,n}$.

Démonstration. En deux parties.

(i) **Si $n > N$** Nous avons, par définition de $C_{b,n}$ que $u_n \neq 0$, de telle sorte que

$$\varphi_{b,n}(u) \geqslant u_n b^n \geqslant b^n > x. \tag{1.118}$$

La dernière inégalité est due au fait que $n \notin \{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}$.

(ii) **Si $n < N$** Nous avons

$$x \geqslant b^N > \varphi_{b,n}(u). \tag{1.119}$$

Le seconde inégalité est une conséquence du lemme 1.88.

□

Théorème 1.90 ([15]).
Soit $b \geqslant 2$. Si $x \in \mathbb{N}$, alors il existe un unique $N \in \mathbb{N}$ et un unique $u \in C_{b,N}$ tels que

$$x = \varphi_{b,N}(u). \tag{1.120}$$

Démonstration. Nous commençons par $x < b$. Dans ce cas, $N = 0$ parce que si $u_k \neq 0$ avec $k \neq 0$, nous avons

$$\sum_{i=0}^{N} u_i b^i \geqslant u_k b^k \geqslant b > x. \tag{1.121}$$

Donc $x = x_0 b^0 = u_0$. Bref, dans le cas $x < b$ nous avons obligatoirement $N = 0$ et $u_0 = x$.

Nous étudions à présent le cas $x \geqslant b$ que nous subdivisons en plusieurs étapes.

(i) $N \geqslant 1$ Si $N = 0$, alors $\varphi_{b,0}(u) = u_0 < b \leqslant x$. Donc $N \geqslant 1$.

Notons incidemment que nous pouvons parler de $N - 1$ à partir de maintenant.

(ii) **Unicité, préambule** Le lemme 1.89 nous indique que si $x = \varphi_{b,N}(u)$ pour un certain $N \in \mathbb{N}$ et un certain $u \in C_{b,N}$, alors

$$N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.122}$$

Nous posons

$$X_k = \sum_{i=k}^{N} u_i b^{i-k}, \tag{1.123}$$

et nous allons montrer que le couple (X_{k+1}, u_k) est le résultat de la division euclidienne[34] de X_k par b.

D'abord, $u_k < b$, donc ça a bien la tête d'un reste. Ensuite, pour le quotient,

$$\begin{aligned} bX_{k+1} + u_k &= b \sum_{i=k+1}^{N} u_i b^{i-(k+1)} + u_k && (1.124a)\\ &= \sum_{i=k+1}^{N} u_i b^{i-k} + u_k && (1.124b)\\ &= \sum_{i=k}^{N} u_i b^{i-k} && (1.124c)\\ &= X_k. && (1.124d) \end{aligned}$$

(iii) **Unicité** En quoi cela fait-il avancer la choucroute ? Supposons que $\varphi_{b,N}(u) = \varphi_{b,M}(v)$. Alors nous avons déjà prouvé que

$$M = N = \max\{k \in \mathbb{N} \text{ tel que } b^k \leqslant x\}. \tag{1.125}$$

Ensuite nous devons montrer que $u = v$. Nous posons $X_k = \sum_{i=k}^{N} u_i b^{i-k}$ et $Y_k = \sum_{i=k}^{N} v_i b^{i-k}$. Notez que

$$X_0 = Y_0 = x. \tag{1.126}$$

Si $X_k = Y_k$, alors par unicité de la division euclidienne nous avons $X_{k+1} = Y_{k+1}$ et $u_k = v_k$. Par récurrence nous avons $X_k = Y_k$ et $u_k = v_k$ pour tout k.

(iv) **Existence** Soit $x \in \mathbb{N}$. Nous posons $y_0 = x$ et

$$y_k = by_{k+1} + u_k \tag{1.127}$$

avec $u_k < b$. Vus l'unicité dans la division euclidienne et le théorème[35] 1.44 permettant la définition par récurrence, ces conditions définissent deux suites (u_k) et (y_k) dans $\mathbb{N}$.

Montrons qu'il existe un $N \in \mathbb{N}$ tel que $y_n = 0$ pour tout $n \geqslant N + 1$. Nous avons :

$$\begin{aligned} 2y_{k+1} &\leqslant by_{k+1} && \text{parce que } b \geqslant 2 && (1.128a)\\ &\leqslant y_k && \text{pcq } by_{k+1} + u_k = y_k. && (1.128b) \end{aligned}$$

Bref : $2y_{k+1} \leqslant y_k$. Par récurrence[36] nous trouvons que

$$2^k y_k \leqslant x \tag{1.129}$$

parce que $y_0 = x$. Par le lemme 1.84, si k est assez grand,

$$2ky_k \leqslant 2^k y_k \leqslant x. \tag{1.130}$$

34. Théorème 1.86.
35. Nous ne citerons pas toujours ce théorème à chaque fois que nous définissons quelque chose par récurrence.
36. Faut-il citer la proposition 1.43 et donner explicitement la fonction P ?

Puisque $\mathbb{N}$ est archimédien[37], nous pouvons considérer $s \in \mathbb{N}$ tel que $2s > x$. À ce moment nous avons

$$y_n = 0 \tag{1.131}$$

pour tout $n \geqslant s$. Nous posons

$$N = \max\{k \text{ tel que } y_k \neq 0\}. \tag{1.132}$$

Prouvons par récurrence sur l que

$$y_{N-l} = \sum_{i=N-l}^{N} u_i b^{(i+l)-N}. \tag{1.133}$$

Notez que $i + l \geqslant N - l + l = N$, donc $(i + l) - N$ a un sens.

(i) **Pour $l = 0$** Avec $l = 0$ nous avons $\sum_{i=N-l}^{N} u_i b^{(i+l)-N} = u_N$. Il faut donc voir que $y_N = u_N$. Nous avons

$$y_N = by_{N+1} + u_N. \tag{1.134}$$

En se rappelant que $y_{N+1} = 0$, nous avons le résultat.

(ii) **Pour $l + 1$** Pour la récurrence nous avons le calcul suivant :

$$\begin{aligned}
y_{N-l-1} &= by_{N-l} + u_{N-l-1} && (1.135a)\\
&= b \sum_{i=N-l}^{N} u_i b^{(i+l)-N} + u_{N-l-1} && (1.135b)\\
&= \sum_{i=N-l}^{N} u_i b^{(i+l)-N+1} + u_{N-l-1} && (1.135c)\\
&= \sum_{i=N-l-1}^{N} u_i b^{(i+l+1)-N}. && (1.135d)
\end{aligned}$$

La récurrence est prouvée. L'égalité (1.133) est validée pour tout l.

En posant $l = N$ dans (1.133) nous trouvons

$$y_0 = \sum_{i=0}^{N} u_i b^i. \tag{1.136}$$

Mais la définition de la suite (y_k) contient $y_0 = x$. Donc nous avons prouvé que

$$x = \sum_{i=0}^{N} u_i b^i = \varphi_{b,N}(u). \tag{1.137}$$

□

Exemple 1.91.
Comment écrire le nombre b en base b ? Nous devons trouver un N et une suite (u_i) tels que

$$b = \sum_{i=0}^{N} u_i b^i. \tag{1.138}$$

Il est facile de voir que le choix $N = 1$ et $u = (0, 1)$ fonctionne bien : $b = 1 \times b^1 + 0$. Nous avons donc

$$b = \varphi_{b,1}(1, 0). \tag{1.139}$$

Nous écrivons cela plus sobrement $b = 10$. △

37. Proposition 1.80.

1.92.
À part des cas très exceptionnels, nous utilisons toujours la base $b = s(9) = s^9(0)$. Nous nous permettons donc d'écrire « 64 » le nombre $\varphi_{s(9),2}(6,4)$. Vous saviez que tout groupe simple d'ordre $\varphi_{s(9),2}(6,0)$ est isomorphe au groupe alterné $A_{\varphi_{s(9),0}(5)}$? C'est la proposition 5.40.

La proposition suivante dit que le nombre qui a le plus de chiffres est le plus grand.

Proposition 1.93 ([15])**.**
Si $u \in C_{b,N}$ et $v \in C_{b,M}$ avec $M > N$ alors $\varphi_{b,N}(u) < \varphi_{b,M}(v)$.

Démonstration. Le lemme 1.88 nous dit que

$$b^N \leqslant \varphi_{b,N}(u) < b^{N+1} \tag{1.140}$$

et

$$b^M \leqslant \varphi_{b,M}(v) < b^{M+1}. \tag{1.141}$$

Puisque $M > N$ nous avons $b^{N+1} \leqslant b^M$ et donc

$$\varphi_{b,N}(u) < b^{N+1} \leqslant b^M \leqslant \varphi_{b,M}(v). \tag{1.142}$$

□

La proposition suivante dit que si deux nombres s'écrivent avec le même nombre de chiffres, le plus grand est celui dont le premier chiffre différent est le plus grand. Autrement dit, les nombres en écriture de position se classent par ordre lexicographique.

Proposition 1.94.
Soient $u, v \in C_{b,N}$ tels que $u_i = v_i$ pour $i = r+1, \dots, N$. Si $u_r > v_r$ alors $\varphi_{b,N}(u) > \varphi_{b,N}(v)$.

Démonstration. En découpant les sommes nous avons

$$\varphi_{b,N}(u) = \sum_{i=r+1}^{N} u_i b^i + u_r b^r + \sum_{i=0}^{r-1} u_i b^i \tag{1.143}$$

et

$$\varphi_{b,N}(v) = \sum_{i=r+1}^{N} u_i b^i + v_r b^r + \sum_{i=0}^{r-1} v_i b^i. \tag{1.144}$$

Puisque $b^r > \sum_{i=0}^{r-1} u_i b^i$ (lemme 1.88), nous avons aussi

$$b^r + \sum_{i=0}^{r-1} u_i b^i > \sum_{i=0}^{r-1} v_i b^i. \tag{1.145}$$

Et le calcul final :

$$\begin{aligned}
\varphi_{b,N}(v) &< \sum_{i=r+1}^{N} v_i b^i + v_r b^r + b^r + \sum_{i=0}^{r-1} u_i b^i && \text{pcq (1.145)} && (1.146a)\\
&= \sum_{i=r+1}^{N} u_i b^i + (v_r+1) b^r + \sum_{i=0}^{r-1} u_i b^i && && (1.146b)\\
&\leqslant \sum_{i=r+1}^{N} u_i b^i + u_r b^r + \sum_{i=0}^{r-1} u_i b^i && \text{pcq } u_r \geqslant v_r + 1 && (1.146c)\\
&= \sum_{i=0}^{N} u_i b^i && && (1.146d)\\
&= \varphi_{b,N}(u). && && (1.146e)
\end{aligned}$$

Et voilà. □

1.95.
Il est aussi possible de définir des choses par récurrence de telle sorte que l'élément x_n soit définit en fonction de tous les x_i $(i < n)$. Voyons comment définir la suite

$$\begin{cases} x_1 = 1 & (1.147a) \\ x_n = \sum_{k=1}^{n-1} x_n & (1.147b) \end{cases}$$

en utilisant le théorème 1.44. Il faut prendre $E = \mathbb{N} \cup \mathbb{N} \times \mathbb{N} \cup \dots$ et ensuite $b = 1$ et

$$g(x_1, \dots, x_{n-1}) = (x_1, \dots, x_{n-1}, \sum_{k=1}^{n-1} x_k). \tag{1.148}$$

L'élément x_n de la suite est la projection sur la n^{e}composante de l'élément $f(n)$.

1.4 Les entiers

Proposition-Définition 1.96 ([15]).
Soient $a, b, a', b' \in \mathbb{N}$. Nous disons que $(a, b) \sim (a', b')$ si et seulement si

$$a + b' = b + a' \tag{1.149}$$

(1) $\sim$ est une relation d'équivalence sur $\mathbb{N}^2$.
(2) Si $(a, b) \sim (a', b')$ et $(x, y) \sim (x', y')$ alors

$$(a + x, b + y) \sim (a' + x', b' + y'). \tag{1.150}$$

*L'ensemble des **entiers** est*

$$\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim, \tag{1.151}$$

et nous notons $\overline{a, b} \in \mathbb{Z}$ la classe de $(a, b) \in \mathbb{N} \times \mathbb{N}$.

Démonstration. En plusieurs points.

(i) **Symétrie** C'est la commutativité de la somme dans $\mathbb{N}$, proposition 1.50(1).

(ii) **Réflexive** Immédiat.

(iii) **Transitive** Nous supposons que $(a, b) \sim (u, v)$ et que $(u, v) \sim (x, y)$. Alors nous avons

$$a + v = u + b \tag{1.152a}$$
$$u + y = v + x. \tag{1.152b}$$

En additionnant membre à membre,

$$a + v + u + y = u + b + v + x. \tag{1.153}$$

La commutativité nous permet de mettre u et v à droite dans chacun des deux membres. Ensuite la proposition 1.50(3) nous permet de simplifier par $u + v$. Il reste $a + y = b + x$, qui signifie $(a, b) \sim (x, y)$.

(iv) **Pour (2)** L'hypothèse donne les égalités

$$a + b' = b + a' \tag{1.154a}$$
$$x + y' = y + x' \tag{1.154b}$$

En sommant, et en utilisant l'associativité,

$$(a + x) + (b' + y') = (b + y) + (a' + x'). \tag{1.155}$$

Cela signifie bien que $(a + x, b + y) \sim (a' + x', b' + y')$.

□

Lemme 1.97.
Soient $a, b \in \mathbb{N}$. Nous avons $(a, b) \sim (0, 0)$ si et seulement si $a = b$.

Démonstration. Dire que $(a, b) \sim (0, 0)$ est équivalent à dire que $a + 0 = b + 0$, ou encore que $a = b$. □

Proposition-Définition 1.98 ([15])**.**
Soient $a, b, x, y \in \mathbb{N}$. L'application

$$\begin{aligned} f\colon \overline{(a,b) \times \overline{(x,y)}} &\to \mathbb{Z} \\ \big((a',b'),(x',y')\big) &\mapsto (a'+x', b'+y') \end{aligned} \tag{1.156}$$

est constante.

Nous nommons sa valeur $\overline{(a,b)} + \overline{(x,y)}$.

Démonstration. Cela est une conséquence de la proposition 1.96(2). □

Proposition 1.99.
La paire $(\mathbb{Z}, +)$ est un groupe commutatif.

Démonstration. En plusieurs points.

(i) **Neutre** Le neutre est $e = \overline{(0,0)}$. En effet,

$$\overline{(a,b)} + \overline{(0,0)} = \overline{(a+0, b+0)} = \overline{(a,b)}. \tag{1.157}$$

De même $e + \overline{(a,b)} = \overline{(a,b)}$ par commutativité de la somme dans $\mathbb{N}$.

(ii) **Inverse** Il est facile de vérifier que $\overline{(b,a)}$ est l'inverse de $\overline{(a,b)}$.

(iii) **Associativité** Calcul direct en utilisant l'associativité dans $\mathbb{N}$.

□

Proposition 1.100.
L'application

$$\begin{aligned} \iota\colon \mathbb{N} &\to \mathbb{Z} \\ n &\mapsto \overline{(n,0)} \end{aligned} \tag{1.158}$$

est un morphisme[38] *injectif.*

Démonstration. Le fait que ce soit un morphisme est le calcul

$$\iota(a+b) = \overline{(a+b,0)} = \overline{(a,0)} + \overline{(b,0)} = \iota(a) + \iota(b). \tag{1.159}$$

Pour l'injectivité, supposons que $\iota(a) = \iota(b)$. Alors $\overline{(a,0)} = \overline{(b,0)}$, c'est-à-dire $a + 0 = b + 0$. Donc $a = b$. □

1.4.1 Opposé

Lemme 1.101.
Tout élément de $\mathbb{Z}$ a un représentant de la forme $(a, 0)$ ou $(0, b)$.

Démonstration. Soient $a, b \in \mathbb{N}$. Si $b \leqslant a$, alors nous avons

$$(a,b) \sim (a-b, 0) \tag{1.160}$$

où la différence est calculée dans $\mathbb{N}$ et a un sens parce que nous avons supposé $b \leqslant a$. Si par contre $a \leqslant b$ alors

$$(a,b) \sim (0, b-a). \tag{1.161}$$

Puisque l'ordre sur $\mathbb{N}$ est total[39], tous les cas sont couverts. □

38. Certes $\mathbb{N}$ n'est pas un groupe, donc le mot « morphisme » est un peu abusé, mais vous voyez ce que je veux dire.

39. Proposition 1.59.

Lemme-Définition 1.102.
Soit $z \in \mathbb{Z}$. L'application[40]

$$\begin{aligned} f\colon z &\to \mathbb{Z} \\ (a,b) &\mapsto \overline{(b,a)} \end{aligned} \tag{1.162}$$

est constante.

Nous nommons $-z$ sa valeur.

Démonstration. Soient (a,b) et (x,y) dans z. Nous avons successivement :

— $(a,b) \sim (x,y)$.
— $a+y = b+x$.
— $(b,a) \sim (y,x)$
— $\overline{(b,a)} = \overline{(y,x)}$
— $f(a,b) = f(x,y)$.

D'où la constance de f. □

Lemme 1.103.
Nous avons

(1) $\mathbb{Z} = \iota(\mathbb{N}) \cup -\iota(\mathbb{N})$
(2) $\iota(\mathbb{N}) \cap -\iota(\mathbb{N}) = \{0\}$.

Démonstration. Nous avons

$$\iota(\mathbb{N}) = \{\overline{(n,0)} \text{ tel que } n \in \mathbb{N}\} \tag{1.163a}$$
$$-\iota(\mathbb{N}) = \{\overline{(0,n)} \text{ tel que } n \in \mathbb{N}\}. \tag{1.163b}$$

Montrons à présent les deux points.

(i) **Pour (1)** Nous savons par le lemme 1.101 que tous les éléments de $\mathbb{Z}$ sont de la forme (1.163a) ou (1.163b).

(ii) **Pour (2)** Si $z \in \iota(\mathbb{N}) \cap -\iota(\mathbb{N})$, il existe $n,m \in \mathbb{N}$ tels que $\overline{(n,0)} = \overline{(0,m)}$, ce qui signifie en particulier que $(n,0) \sim (0,m)$ ou encore que $n+m=0$. Le lemme 1.52 dit alors que $n=m=0$.
Nous avons donc $z = \overline{(0,0)} = 0$.

□

1.4.2 Ordre sur $\mathbb{Z}$

Si $z \in \mathbb{Z}$, nous disons que $z \in \mathbb{N}$ lorsque $z \in \iota(\mathbb{N})$. C'est un abus de notation qu'il est difficile de ne pas faire.

Proposition-Définition 1.104 (Relation d'ordre [15])**.**
Nous disons que $x \leqslant y$ si et seulement si $y - x \in \mathbb{N}$.

L'ensemble $(\mathbb{Z}, \leqslant)$ est totalement ordonné.

Une version dans $\mathbb{R}$ sera le lemme 1.369.

Lemme 1.105.
Soient $a > 0$ et $b > 1$ dans $\mathbb{Z}$. Nous avons

$$ab > a. \tag{1.164}$$

40. Pour rappel, z est une classe d'équivalence dans $\mathbb{N} \times \mathbb{N}$, c'est-à-dire une partie de $\mathbb{N} \times \mathbb{N}$. Ça a un sens de prendre z comme ensemble sur lequel on définit une fonction.

Lemme 1.106.
Toute partie bornée de $\mathbb{Z}$ possède un plus grand élément.

Proposition 1.107.
Soit $a, b \in \mathbb{Z}$ tels que a divise b. Alors $|a| \leqslant |b|$.

Lemme 1.108.
L'ensemble $\mathbb{Z}$ est infini dénombrable.

1.5 Quelques résultats de cardinalité

1.5.1 Équipotence, surpotence, subpotence

Les notions d'équipotence, surpotence et de subpotence permettent de comparer les « tailles » des ensembles sans avoir besoin de la théorie des ordinaux. Tout ceci ne sera pas très souvent utile par la suite. Un exemple d'utilisation de ces notions est le théorème de Steinitz 6.128 qui démontre l'existence de clôture algébrique pour tout corps.

Définition 1.109 ([18, 19])**.**
Soient deux ensembles A et B.

*(1) Les ensembles A et B sont **équipotents** si il existe une bijection entre A et B. Nous notons $A \approx B$.*

*(2) L'ensemble A est **surpotent** à B si il existe une surjection de A vers B. Nous notons $A \succeq B$.*

*(3) L'ensemble A est **subpotent** à B si il existe une injection de A vers B. Nous notons $A \preceq B$.*

Nous disons également « strictement » surpotent quand il y a surpotence mais pas équipotence, et de même pour la subpotence. Les symboles $\succ$ et $\prec$ sont alors utilisés.

Proposition 1.110 ([1, 11])**.**
L'ensemble A est subpotent à B si et seulement si B est surpotent à A.

Démonstration. En deux parties.

(i) $\Rightarrow$ Nous supposons que A est subpotent à B. Il existe une injection $\varphi\colon A \to B$. Nous définissons $f\colon B \to A$ par

$$f(x) = \begin{cases} \varphi^{-1}(x) & \text{si } x \in \varphi(A) \\ a & \text{sinon} \end{cases} \tag{1.165}$$

où a est un élément quelconque de A. Cette application est bien définie parce que φ est injective, de telle sorte que φ^{-1} est bien définie. Puisque φ est définie sur tout a, l'application f est une surjection.

(ii) $\Leftarrow$ Nous supposons que B est surpotent à A. Il existe donc une surjection $\varphi\colon B \to A$. Pour chaque $x \in A$ nous considérons un élément $b_x \in \varphi^{-1}(x)$, qui existe parce que φ est surjective. Nous considérons ensuite l'application

$$\begin{aligned} f\colon A &\to B \\ x &\mapsto b_x. \end{aligned} \tag{1.166}$$

Nous prouvons que f est une injection. Supposons que $x, y \in A$ soient tels que $f(x) = f(y)$. Nous avons $b_x = b_y$. Donc

$$x = \varphi(b_x) = \varphi(b_y) = y. \tag{1.167}$$

Nous avons prouvé que $x = y$, et donc que f est injective.

□

Vu que l'ensemble des ensembles n'existe pas [41], nous n'allons pas énoncer le fait que ces notions donnent une relation d'ordre sur les ensembles ; il faudrait parler de classes et nous ne nous en sortirions pas. Nous allons toutefois énoncer quelques résultats qui vont dans ce sens. Pour en savoir plus, vous pouvez lire les différentes pages de Wikipédia sur les nombres cardinaux.

1.5.2 Un peu d'infinité

Définition 1.111 (ensemble Dedekind infini)**.**
*Un ensemble est **infini** si il peut être mis en bijection avec un de ses sous-ensembles propres (c'est-à-dire différent de lui-même).*

*Un ensemble est **fini** si il n'est pas infini.*

1.112.
Nous adoptons les notions d'ensembles finis et infinis au sens de Dedekind. De nombreuses sources (dont wikipédia [20, 21]) définissent un ensemble fini comme étant un ensemble en bijection avec une partie de $\mathbb{N}$ de la forme $\{0, \dots, N\}$. Alors un ensemble est infini si il n'est pas fini.

Cependant, d'une part *les deux définitions d'ensembles infinis ne sont pas équivalentes*, mais d'autre part, elles sont équivalentes si on accepte l'axiome du choix [42]. Or le Frido accepte l'axiome du choix sans vergogne et sous toutes ses formes. Nous démontrerons donc, en utilisant le lemme de Zorn, qu'un ensemble A est fini (définition 1.111) si et seulement si il existe une bijection $\{0, \dots, N\} \to A$ pour un certain $N \in \mathbb{N}$. Ce sera le théorème 1.120.

Lemme 1.113.
Toute partie d'un ensemble fini est finie.

Démonstration. Nous allons prouver la contraposée : si un ensemble contient une partie infinie, alors il est infini. Soit $A \subset B$ où A est infini. Nous allons prouver que B est infini. En vertu de la définition 1.111, il existe une partie $A' \subsetneq A$ et une bijection $\sigma\colon A' \to A$.

Nous considérons la partie $B' = A' \cup (B \backslash A)$, qui est une partie stricte de B. Puisque $A' \cap (B \backslash A) = \varnothing$, nous pouvons définir

$$\begin{aligned} \varphi\colon B' &\to B \\ x &\mapsto \begin{cases} \sigma(x) & \text{si } x \in A' \\ x & \text{si } x \in B \backslash A. \end{cases} \end{aligned} \tag{1.168}$$

Montrons que φ est une bijection.

(i) **Surjectif** Nous avons $\varphi(A') = A$ et $\varphi(B \backslash A) = B \backslash A$. Donc

$$\varphi(B') = \varphi(A') \cup \varphi(B \backslash A) = A \cup (B \backslash A) = B. \tag{1.169}$$

(ii) **Injectif** Si $\varphi(x) = \varphi(y)$, nous avons 4 possibilités suivant que x et y sont dans A' ou $B \backslash A$.
Si $x, y \in A'$, alors $\varphi(x) = \varphi(y)$ implique $\sigma(x) = \sigma(y)$ et donc $x = y$ parce que σ est injective.
Si $x \in A'$ et $y \in B \backslash A$ alors $\varphi(x) = \sigma(x) \in A$ et $\varphi(y) = y \in B \backslash A$. Ce cas n'est pas possible. Le cas $x \in B \backslash A$ et $y \in A'$ n'est pas possible non plus.
Si $x, y \in B \backslash A$, alors $\varphi(x) = \varphi(y)$ implique immédiatement $x = y$.

Nous avons une bijection entre B' et B alors que B' est un sous-ensemble strict de B. Donc B est infini. □

Proposition 1.114 ([1])**.**
Si A est fini et si $\omega \notin A$, alors $A \cup \{\omega\}$ est fini.

41. Voir le corolaire 1.144.
42. Et même seulement l'axiome du choix dénombrable ; si vous voulez en savoir plus, lisez la page wikipédia [22].

Démonstration. Supposons que $A \cup \{\omega\}$ est infini. Il existe un sous-ensemble strict de $A \cup \{\omega\}$ en bijection avec $A \cup \{\omega\}$. Soient donc $B \subsetneq A \cup \{\omega\}$ et $\sigma\colon B \to A \cup \{\omega\}$ une bijection.

Il y a deux possibilités : soit ω est dans B, soit non.

(i) $\underline{\omega \notin B}$ Alors $B \subset A$, et il existe $x \in B$ tel que $\sigma(x) = \omega$. Considérons $B' = B \backslash \{x\}$; cela est une partie propre de A. Ensuite nous définissons

$$\begin{aligned} \varphi\colon B\backslash\{x\} &\to A \\ a &\mapsto \sigma(a). \end{aligned} \tag{1.170}$$

C'est injectif parce que σ est injective, et c'est surjectif parce que

$$\varphi\big(B\backslash\{x\}\big) = \sigma(B)\backslash\sigma(\{\omega\}) = \sigma(B)\backslash\{\omega\} = (A \cup \{\omega\})\backslash\{\omega\} = A. \tag{1.171}$$

Pour la dernière égalité nous avons utilisé le fait que ω n'est pas dans A.

(ii) **Si $\omega \in B$** Puisque B est une partie propre de $A \cup \{\omega\}$, il existe $x \in A \backslash B$. Nous considérons $B' = \big(B\backslash\{\omega\}\big) \cup \{x\}$ et nous définissons

$$\begin{aligned} \varphi\colon B' &\to A \\ b &\mapsto \begin{cases} \sigma(b) & \text{si } b \neq x \\ \sigma(\omega) & \text{si } b = x. \end{cases} \end{aligned} \tag{1.172}$$

Nous montrons à présent que φ est une bijection.

(i) **Injectif** Soient u, v tels que $\varphi(u) = \varphi(v)$. Il y a 4 possibilités suivant que u ou v est égal à x.

Si $u = v = x$ on est bon.

Si $u = x$ et $v \neq x$, alors $\varphi(u) = \sigma(\omega)$ et $\varphi(v) = \sigma(v)$. Mais $x \neq \omega$ parce que $x \in A$, donc cette situation n'est pas possible parce que σ est injective.

Si $u \neq x$ et $v \neq x$, alors $\varphi(u) = \sigma(u)$ et $\varphi(v) = \sigma(v)$. Dans ce cas l'injectivité de σ fait que $x = y$.

(ii) **Surjective** Soit $y \in A$. Vu que $\sigma\colon B \to A \cup \{\omega\}$ est surjective, il existe $b \in B$ tel que $\sigma(b) = y$. Si $b \neq \omega$ alors $\varphi(b) = \sigma(b) = y$. Si au contraire $b = \omega$, alors $\varphi(x) = \sigma(\omega) = y$. Dans les deux cas, y est dans l'image de φ.

Dans tous les cas nous avons construit une bijection entre une partie propre $B \subsetneq A \cup \{\omega\}$ et $A \cup \{\omega\}$. □

La proposition suivante est à peu près prise comme définition d'un ensemble fini dans [23] qui donne également une preuve de l'équivalence avec notre définition.

Proposition 1.115.
L'ensemble $\mathbb{N}$ est infini[43].

Démonstration. Nous considérons la partie propre[44]

$$A = \{n \in \mathbb{N} \text{ tel que } n \geqslant 1\}. \tag{1.173}$$

Ensuite nous posons

$$\begin{aligned} \sigma\colon \mathbb{N} &\to A \\ x &\mapsto s(x). \end{aligned} \tag{1.174}$$

Le fait que σ prenne ses valeurs dans A est parce que s prend ses valeurs dans $\mathbb{N}\backslash\{0\} = A$.

(i) **σ est injective** Parce que s l'est.

43. Définition 1.111.

44. Le fait que ce soit une partie propre est dû au fait que 0 n'est pas dedans d'une part parce que le lemme 1.54 dit que $0 \leqslant 1$, et d'autre part parce que le lemme 1.52 donne $0 \neq 1$.

(ii) **σ est surjective** C'est dans la définition 1.40(2)

Nous avons donc une bijection entre $\mathbb{N}$ et un sous-ensemble strict. □

Lemme 1.116.
Soit $N \in \mathbb{N}$. La partie $\{0, \dots, N\}$ est finie[45].

Démonstration. Par récurrence sur N. Avec $N = 0$, la partie $\{0\}$ est finie parce que son seul sous-ensemble propre est $\varnothing$ qui n'est pas en bijection avec $\{0\}$.

Supposons que $\{0, \dots, N\}$ est fini. Alors $\{0, \dots, N\} \cup \{N + 1\}$ est fini par le lemme 1.114. □

Lemme 1.117.
Si $p \neq q$, alors il n'existe pas de bijection entre $\{0, \dots, p\}$ et $\{0, \dots, q\}$.

Démonstration. Supposons pour fixer les idées que $p \leqslant q$. Dans ce cas $\{0, \dots, p\}$ est une partie stricte de $\{0, \dots, q\}$. Vu que $\{0, \dots, q\}$ est fini (lemme 1.116), il n'y a pas de bijection avec ses parties strictes. □

Proposition 1.118 ([1])**.**
Si A est infini et si $\sigma\colon A \to B$ est injective, alors B est infini.

Démonstration. Nous allons prouver que $\sigma(A)$ est une partie infinie de B. Puisque A est infini, nous pouvons considérer une partie $A' \subsetneq A$ et une bijection $\varphi_A\colon A' \to A$. Nous définissons

$$\begin{aligned} \varphi_B\colon \sigma(A') &\to \sigma(A) \\ y &\mapsto \sigma\Big(\varphi\big(\sigma^{-1}(y)\big)\Big). \end{aligned} \tag{1.175}$$

Cette définition a un sens parce que si $y \in \sigma(A')$, alors il existe un unique $x \in A'$ tel que $\sigma(x) = y$ parce que σ est injective. De là, $\varphi_A(x) \in A$ et nous pouvons lui appliquer σ.

Nous montrons que φ_B est une bijection.

(i) **Injective** Supposons que $\varphi_B(a) = \varphi_B(b)$, c'est-à-dire que

$$(\sigma \circ \varphi_A \circ \sigma^{-1})(a) = (\sigma \circ \varphi_A \circ \sigma^{-1})(b). \tag{1.176}$$

Étant donné que σ et φ_A sont injectives, nous avons $\sigma^{-1}(a) = \sigma^{-1}(b)$. En appliquant σ des deux côtés, nous trouvons $a = b$.

(ii) **surjective** Soit $y \in \sigma(A)$. En prenant $x \in (\sigma \circ \varphi_A^{-1} \circ \sigma^{-1})(y)$ nous avons $\varphi_B(x) = y$.

Donc B contient une partie infinie ($\sigma(A)$). Le lemme 1.113 conclut que B est infini. □

Lemme 1.119.
À propos d'applications entre ensembles finis.

(1) Si $\sigma\colon A \to B$ est une application quelconque et si A est fini, alors $\sigma(A)$ est une partie finie de B.

(2) Si $\sigma\colon A \to B$ est surjective et si A est fini, alors B est fini.

Démonstration. Nous allons utiliser le lemme de Zorn. Nous considérons l'ensemble

$$\mathcal{A} = \{X \subset A \text{ tel que } \sigma\colon X \to \sigma(A) \text{ est injective}\} \tag{1.177}$$

que nous ordonnons (partiellement) par l'inclusion.

(i) **$\mathcal{A}$ est inductif** Soit une partie totalement ordonnée $\mathcal{F}$ de $\mathcal{A}$. Nous considérons $Y = \bigcup_{X \in \mathcal{F}} X$, et nous prouvons que Y est un majorant de $\mathcal{F}$.

Pour cela nous commençons par prouver que $Y \in \mathcal{A}$. Soient $a, b \in Y$ tels que $\sigma(a) = \sigma(b)$. Il existe $X_1, X_2 \in \mathcal{F}$ tels que $a \in X_1$ et $b \in X_2$. Supposons pour fixer les idées que $X_1 \leqslant X_2$ ($\mathcal{F}$ étant totalement ordonné nous avons toujours $X_1 \leqslant X_2$ ou $X_2 \leqslant X_1$). Puisque l'ordre est

45. Pour rappel, la définition de $\{0, \dots, N\}$ est 1.64.

l'inclusion, cela signifie que $X_1 \subset X_2$. Nous avons donc $a, b \in X_2$, alors que σ est injective sur X_2. Donc $\sigma(a) = \sigma(b)$ implique $a = b$, et σ est injective sur Y. Nous avons donc prouvé que $Y \in \mathcal{A}$.

Puisque pour tout $X \in \mathcal{F}$ nous avons $X \subset Y$, nous avons $X \leqslant Y$ (dans $\mathcal{A}$) pour tout $X \in \mathcal{F}$. Bref, Y est un majorant de $\mathcal{F}$ dans $\mathcal{A}$.

Toute partie totalement ordonnée de $\mathcal{A}$ est majorée. Cela signifie que $\mathcal{A}$ est inductif[46].

(ii) **Zorn** L'ensemble $\mathcal{A}$ étant inductif et non vide (les singletons dans A sont dans $\mathcal{A}$), il possède un élément maximal[47] par le lemme de Zorn 1.22. Nous nommons A' un élément maximal dans $\mathcal{A}$.

(iii) **Bijective** L'application $\sigma\colon A' \to \sigma(A)$ est injective parce que $A' \in \mathcal{A}$. Nous devons prouver qu'elle est surjective.

Supposons que $y \in \sigma(A)\backslash\sigma(A')$. Alors il existe $a \in A\backslash A'$ tel que $\sigma(a) = y$. Dans ce cas, la partie $A' \cup \{a\}$ est un majorant de A' dans $\mathcal{A}$, ce qui est impossible.

Donc $\sigma\colon A' \to \sigma(A)$ est bijective.

(iv) **Conclusion** L'ensemble A' est fini en tant que partie de l'ensemble fini A (lemme 1.113). L'application σ étant injective, la proposition 1.118 conclut que $\sigma(A')$ est fini. Et comme $\sigma(A')$ n'est autre que $\sigma(A)$ nous avons fini.

La partie (1) est prouvée. La partie (2) est maintenant facile. La partie (1) dit que $\sigma(A)$ est une partie finie de B, mais si σ est surjective, alors $\sigma(A) = B$. □

Proposition-Définition 1.120 (Cardinal d'un ensemble fini[1])**.**
Soit un ensemble non vide I.

(1) L'ensemble I est fini[48] *si et seulement si il existe une bijection entre I et $\{0, \dots, N\}$ pour un certain $N \in \mathbb{N}$.*

(2) Si I est fini, il existe un unique $N \in \mathbb{N}$ tel que I soit en bijection avec $\{0, \dots, N\}$.

Dans ce cas, le nombre $N + 1$ est le ***cardinal*** *de I, et est noté $\mathrm{Card}(I)$. Pour l'ensemble vide, nous définissons $\mathrm{Card}(\varnothing) = 0$.*

Démonstration. En plusieurs parties.

(i) **(1)⇐** Soit un ensemble A en bijection avec $\{0, \dots, N\}$. Nous avons vu que $\{0, \dots, N\}$ est fini dans le lemme 1.116. Le lemme 1.119(2) conclut que A est fini.

(ii) **(1)⇒** Le vrai sport est de faire l'implication inverse. Nous supposons que A est un ensemble fini, et nous allons prouver qu'il est en bijection avec $\{0, \dots, N\}$ pour un N bien choisi. Nous allons utiliser le lemme de Zorn. Soit

$$\mathcal{A} = \{(N, \varphi) \text{ tel que } \varphi\colon \{0, \dots, N\} \to A \text{ est injective}\}. \tag{1.178}$$

Nous mettons sur $\mathcal{A}$ la relation d'ordre donnée par $(N_1, \varphi_1) \leqslant (N_2, \varphi_2)$ lorsque

(1) $N_1 \leqslant N_2$

(2) φ_2 étend φ_1, c'est-à-dire que $\varphi_2 = \varphi_1$ sur $\{0, \dots, N\}$.

(i) **$\mathcal{A}$ est inductif** Soit une partie $\mathcal{F}$ totalement ordonnée de $\mathcal{A}$. Nous considérons la partie suivante de $\mathbb{N}$:

$$S = \{n \in \mathbb{N} \text{ tel que } \exists (n, \varphi) \in \mathcal{F}\}. \tag{1.179}$$

Si S est majoré, alors il a un maximum (proposition 1.60(4)). Si M est le maximum de S, le (M, φ) de $\mathcal{F}$ qui correspond à ce maximum est un majorant de $\mathcal{F}$.

46. Plus précisément c'est l'ensemble ordonné $(\mathcal{A}, \subset)$ qui est inductif.
47. Définition 1.12 ; voir aussi 1.15.
48. Ensemble fini, définition 1.111.

Supposons –pour l'absurde– que $\mathcal{F}$ n'est pas majoré ; en particulier S n'est pas majoré. Pour chaque $n \in S$, il existe une application φ_n telle que $(n, \varphi_n) \in \mathcal{F}$. Cela nous permet de définir

$$\begin{aligned} \phi\colon S &\to A \\ n &\mapsto \varphi_n(n). \end{aligned} \tag{1.180}$$

Montrons que ϕ est injective. Si $\phi(m) = \phi(n)$, alors $\varphi_m(m) = \varphi_n(n)$. Supposons pour fixer les idées que $n \leqslant m$. Vu que (n, φ_n) et (m, φ_m) sont dans $\mathcal{F}$ qui est ordonné, φ_m prolonge φ_n ; en particulier $\varphi_n(n) = \varphi_m(n)$. Mais comme φ_m est injective, $m = n$.

Nous avons donc une injection $\phi\colon S \to A$. Mais S est non borné et donc infini[49]. La proposition 1.118 conclut que A est infini, ce qui est contraire aux hypothèses.

Donc $\mathcal{F}$ a un majorant et $\mathcal{A}$ est inductif.

(ii) **Lemme de Zorn** Le lemme de Zorn dit que $\mathcal{A}$ a un élément maximal.

(iii) **Conclusion** Soit (N, φ) un élément maximal de $\mathcal{A}$. Nous allons prouver que $\varphi\colon \{0, \dots, N\} \to A$ est une bijection. Que φ soit injective est une conséquence du fait qu'elle est dans $\mathcal{A}$.

Pour prouver que φ est surjective, nous supposons qu'elle ne l'est pas. Soit $a \in A$ qui n'est pas dans l'image de φ. En posant

$$\begin{aligned} \phi\colon \{0, \dots, N+1\} &\to A \\ x &\mapsto \begin{cases} \varphi(x) & \text{si } x \neq N+1 \\ a & \text{si } x = N+1, \end{cases} \end{aligned} \tag{1.181}$$

le couple $(N+1, \phi)$ majore strictement (N, φ). Ce qui est une contradiction.

(iii) **(2) existence** L'existence est ce que nous venons de montrer ci-dessus.

(iv) **(2) unicité** Supposons que I soit en bijection avec $\{0, \dots, M\}$ et avec $\{0, \dots, N\}$. Il existe donc une bijection entre $\{0, \dots, M\}$ et $\{0, \dots, N\}$. Par le lemme 1.117, cela implique que $M = N$.

□

Nous ne définissons pas ce qu'est le cardinal d'un ensemble infini ; c'est très compliqué et ça ne nous servira pas.

Lemme 1.121.
Si A est une partie infinie de $\mathbb{N}$, alors pour tout n, la partie $A \backslash \{0, \dots, n\}$ est non vide.

Démonstration. Si $A \backslash \{0, \dots, n\}$ était vide, cela signifierait que A est une partie de $\{0, \dots, n\}$. Or nous savons que $\{0, \dots, n\}$ est fini (lemme 1.116), et que toute partie d'un ensemble fini est finie (lemme 1.113). Donc nous aurions que A est fini, ce qui est contraire à l'hypothèse. □

Lemme 1.122 ([1]).
Union d'ensembles finis.

(1) Si A et B sont des ensembles finis disjoints, alors $A \cup B$ est fini et

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B). \tag{1.182}$$

(2) Si A et B sont des ensembles finis, alors $A \cup B$ est fini.

(3) Si A est fini et si $B \subset A$ alors

$$\mathrm{Card}(A \backslash B) = \mathrm{Card}(A) - \mathrm{Card}(B). \tag{1.183}$$

(4) Si A et B sont des ensembles quelconques, alors

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B). \tag{1.184}$$

49. Le lemme 1.66 dit qu'il existe une bijection entre S et $\mathbb{N}$. De là nous concluons que S est infini parce qu'un ensemble en bijection avec un ensemble infini est infini par la proposition 1.118.

(5) Si les $\{A_i\}_{i=1,\dots,n}$ sont des ensembles disjoints, alors

$$\mathrm{Card}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mathrm{Card}(A_i). \tag{1.185}$$

(6) Si I ou J est infini, alors $I \cup J$ est infini.

Démonstration. Point par point.

(i) **Pour (1)** Puisque A et B sont finis, la proposition 1.120 nous dit qu'il existe des naturels N et M ainsi que des bijections $\varphi_A\colon \{0,\dots,N\} \to A$ et $\varphi_B\colon \{0,\dots,M\} \to B$. Maintenant l'application

$$\begin{aligned} \varphi\colon \{0,\dots,M+N+1\} &\to A \cup B \\ n &\mapsto \begin{cases} \varphi_A(n) & \text{si } n \leqslant N \\ \varphi_B(n-N-1) & \text{si } n > N \end{cases} \end{aligned} \tag{1.186}$$

est une bijection. Le fait que A et B soient disjoints est important pour l'injectivité. La proposition 1.120 nous dit qu'alors $A \cup B$ est fini. De plus, par définition le cardinal de $A \cup B$ est $N + M$.

(ii) **Pour (2)** Nous ne supposons plus que A et B sont disjoints. Nous posons $I = A$ et $J = B \backslash A$. Avec ça, I et J sont disjoints et finis (comme parties des ensembles finis, lemme 1.113), et vérifient $I \cup J = A \cup B$. Le point (1) indique que $I \cup J$ est fini.

(iii) **Pour (3)** L'ensemble A peut être écrit sous la forme d'une union disjointe : $A = B \cup (A \backslash B)$. Les ensembles B et $A \backslash B$ étant disjoints, nous avons

$$\mathrm{Card}(A) = \mathrm{Card}(B) + \mathrm{Card}(A \backslash B). \tag{1.187}$$

(iv) **Pour (4)** Nous utilisons quelques égalités d'ensembles pour ramener $A \cup B$ à des cas déjà traités :

$$A \cup B = A \cup (B \backslash A) = A \cup \big(B \backslash (A \cap B)\big). \tag{1.188}$$

Nous avons en particulier utilisé $B \backslash A = B \backslash (A \cap B)$. La chose intéressante dans (1.188) est que l'union est disjointe et que $A \cap B \subset B$. Nous pouvons donc écrire

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}\big(B \backslash (A \cap B)\big) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B). \tag{1.189}$$

(v) **Pour (5)** Récurrence en utilisant le point (4).

(vi) **Pour (6)** Toute partie d'un ensemble fini est finie (lemme 1.113). Donc si $I \cup J$ était fini, I et J devraient l'être.

□

Définition 1.123.
*Un ensemble est **dénombrable** si il peut être mis en bijection avec $\mathbb{N}$. Il est **non dénombrable** si 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 1.111, c'est qu'un ensemble fini n'est ni dénombrable ni non dénombrable [50].

Lemme 1.124.
Si A est dénombrable et si il existe une surjection $f\colon A \to B$, alors B est fini ou dénombrable.

Lemme 1.125.
Si A est un ensemble fini ou dénombrable, alors il existe une surjection $\mathbb{N} \to A$.

50. Beaucoup de sources disent qu'un ensemble est dénombrable lorsqu'il est en bijection avec une partie de $\mathbb{N}$. Cela laisse la porte ouverte aux ensembles finis. Par exemple Wikipédia[24].

1.5.3 Dénombrabilité et ensemble des naturels

Proposition 1.126 ([1, 11]).
Toute partie infinie de $\mathbb{N}$ est dénombrable.

Démonstration. Soit A, une partie infinie de $\mathbb{N}$.

(i) **Définition de σ** Nous voulons construire une application $\sigma\colon \mathbb{N} \to A$ telle que

$$\begin{cases} \sigma(0) = \min(A) & (1.190a) \\ \sigma(k+1) = \min\Big(A\backslash\sigma\big(\{0,\dots,k\}\big)\Big) & (1.190b) \end{cases}$$

Les lâches, par prudence, diront juste que c'est défini par récurrence et n'insisteront pas. Nous, nous insistons.

Nous allons définir $\sigma(n)$ à l'aide du théorème 1.44. Pour cela nous posons $E = \mathcal{P}(A)$, $b = \varnothing$ et

$$\begin{aligned} g\colon E &\to E \\ Z &\mapsto \begin{cases} A & \text{si } Z = A \\ Z \cup \{\min(A\backslash Z)\} & \text{sinon.} \end{cases} \end{aligned} \tag{1.191}$$

Notons que la proposition 1.60(2) nous indique que toute partie non vide de $\mathbb{N}$ possède un minimum ; la définition de g a donc un sens. Le théorème 1.44 donne alors une application $f\colon \mathbb{N} \to E$ telle que

(1) $f(0) = b = \varnothing$

(2) $f(n+1) = g\big(f(n)\big)$ pour tout $n \geqslant 0$.

Prouvons par récurrence que $f(n)$ est un ensemble fini pour tout n. D'abord $f(0) = \varnothing$. Ensuite, si $n \geqslant 0$ est tel que $f(n)$ est fini, alors en particulier $f(n) \neq A$ et nous avons

$$f(n+1) = g\big(f(n)\big) = f(n) \cup \{\min\big(A\backslash f(n)\big)\}. \tag{1.192}$$

Dans ce cas, $f(n+1)$ est également fini comme union de deux ensembles finis.

Nous posons

$$\sigma(n) = \min\big(A\backslash f(n)\big). \tag{1.193}$$

Avec $n = 0$, nous avons $\sigma(0) = \min\big(A\backslash\varnothing\big) = \min(A)$. La condition (1.190a) est donc déjà satisfaite.

Nous devons encore prouver (1.190b). Pour tout n, la relation entre $f(n)$ et $\sigma(n)$ est donnée par

$$\begin{cases} f(0) = \varnothing & (1.194a) \\ f(n+1) = f(n) \cup \sigma(n). & (1.194b) \end{cases}$$

Par récurrence nous avons alors

$$f(n) = \bigcup_{k=0}^{n-1} \{\sigma(k)\} = \sigma\big(\{0,\dots,n-1\}\big) \tag{1.195}$$

pour tout $n \geqslant 1$. Nous avons alors la condition 1.190b en substituant (1.195) dans la définition (1.193) écrite avec $n+1$:

$$\sigma(n+1) = \min\big(A\backslash f(n+1)\big) = \min\Big(A\backslash\sigma\big(\{0,\dots,n\}\big)\Big). \tag{1.196}$$

(ii) **σ est strictement croissante** Vu que $A\backslash\sigma\{0,\dots,k\} \subset A\backslash\sigma\{0,\dots,k-1\}$, le minimum est plus grand ou égal : $\sigma(k+1) \geqslant \sigma(k)$. Mais $\sigma(k+1)$ est sélectionné dans l'ensemble $A\backslash\sigma\{0,\dots,k\}$, qui ne contient justement pas $\sigma(k)$. Donc $\sigma(k+1) \neq \sigma(k)$.

(iii) **σ est définie sur $\mathbb{N}$** Il faut montrer que pour tout k, l'ensemble $A\backslash\sigma\{0,\dots,k\}$ est non vide. Si il l'était, cela signifierait que $A\subset\sigma\{0,\dots,k\}$. Par le lemme 1.119(1), la partie $\sigma\{0,\dots,k\}$ est finie dans $\mathbb{N}$. Le lemme 1.113 dit alors qu'en tant que partie de $\sigma\{0,\dots,k\}$, l'ensemble A est fini. Mais comme les hypothèses disent que A est infini, nous avons une contradiction et nous concluons que σ est bien définie sur tout $\mathbb{N}$.

(iv) **σ est injective** Une application $\mathbb{N}\to\mathbb{N}$ strictement croissante est injective par la proposition 1.65.

(v) **σ est surjective** Soit $a\in A$. Vu que σ est strictement croissante et que $\sigma(0)\geqslant 0$, nous avons $\sigma(a)\geqslant a$. Si $\sigma(a)=a$ nous avons terminé. Supposons $\sigma(a)>a$. Alors

$$\min\big(A\backslash\sigma\{0,\dots,a\}\big)>a. \tag{1.197}$$

Si $\sigma\big(\{0,\dots,a\}\big)$ ne contenait pas a, alors $A\backslash\sigma(\{0,\dots,a\})$ le contiendrait et nous n'aurions pas l'inégalité (1.197). Donc $a\in\sigma\big(\{0,\dots,a\}\big)$ et a est bien dans l'image de σ.

□

1.127.
La proposition 1.126 pourrait être prouvée plus facilement en acceptant le théorème de Cantor-Schröder-Bernstein 1.139. Il existe une injection $A\to\mathbb{N}$ parce que A est une partie de $\mathbb{N}$. Mais puisque A est infini, il possède une partie dénombrable. Cela donne une surjection $A\to\mathbb{N}$ et donc une injection $\mathbb{N}\to A$. Le théorème de Cantor-Schröder-Bernstein conclut.

Cela dit, une telle preuve demanderait des outils plus complexes.

1.128.
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[51]. Son utilité est de pouvoir créer un langage de programmation pouvant traiter des paires d'entiers rien qu'en traitant des entiers.

Proposition 1.129 (Une bijection $\mathbb{N}\times\mathbb{N}\to\mathbb{N}$)**.**
La fonction

$$\begin{aligned} f\colon\mathbb{N}\times\mathbb{N}&\to\mathbb{N}\\ (x,y)&\mapsto\begin{cases} y^2+x & \text{si } x<y\\ x^2+x+y & \text{si } y\leqslant x.\end{cases}\end{aligned} \tag{1.198}$$

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.

(i) **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 $k=x'^2+x'+y'$ avec $y'\leqslant x'$.

On y va.

(1) Nous supposons $y^2+x=t^2+z$ avec $x<y$ et $z<t$. Pour fixer les idées, nous supposons $t>y$ et nous posons $t=y+s$ ($s\geqslant 1$). En substituant, et en isolant z,

$$\begin{aligned} z&=x-2sy-s^2 &(1.199a)\\ &<x-2sy &(1.199b)\\ &<x-2sx &(1.199c)\\ &=x(1-2s) &(1.199d)\\ &<0. &(1.199e)\end{aligned}$$

51. 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.

Impossible parce que $z \geqslant 0$.

(2) De même nous supposons $x^2 + x + y = z^2 + z + t$ avec $y \leqslant x$ et $t \leqslant z$. Nous posons $z = x + s$, et nous déballons le même genre de calculs en isolant t.

(3) Enfin nous supposons $y^2 + x = z^2 + z + t$ avec $x < y$ et $t \leqslant z$. Les plus courageux diviseront en trois cas : $y < z$, $y = z$ et $y > z$ et feront les calculs. Par exemple, pour le cas $y > z$ nous posons $y = z + s$ et nous substituons :

$$(y + s)^2 + x = z^2 + z + t \tag{1.200}$$

qui donne

$$x = z + t - 2zs - s^2 < 2z - 2zs - s^2 = 2z(1 - s) - s^2 \leqslant -s < 0 \tag{1.201}$$

parce que $s \geqslant 1$, donc $1 - s \leqslant 0$.

(ii) **f est surjective** Nous devons prouver que tous les éléments de $\mathbb{N}$ sont dans l'image de $\mathbb{N} \times \mathbb{N}$ par f. En premier lieu, $0 = f(0, 0)$. C'est un bon début. Soit $a \in \mathbb{N}$ non nul ; nous montrons que tous les nombres de a^2 à $(a + 1)^2$ sont des images de f. D'abord $a^2 = f(0, a)$, ensuite les nombres

$$f(1, a), f(2, a), \ldots, f(a - 1, a) \tag{1.202}$$

prennent les valeurs $a^2 + 1$, …, $a^2 + a - 1$. Enfin nous avons $f(a, 0) = a^2 + a$ 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).

Corolaire 1.130.
Il existe des parties $\{\mathbb{N}_i\}_{i\in\mathbb{N}}$ telles que $\bigcup_{i\in\mathbb{N}} \mathbb{N}_i = \mathbb{N}$ et que chaque $\mathbb{N}_i$ soit en bijection avec $\mathbb{N}$

Démonstration. Nous considérons la bijection $f\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ donnée par (l'inverse de celle donnée) par la proposition 1.129, et nous posons

$$\mathbb{N}_i = f^{-1}(i, \mathbb{N}). \tag{1.203}$$

L'application

$$f\colon \mathbb{N}_i \to \{(i, k)\}_{k\in\mathbb{N}} \tag{1.204}$$

est une bijection. Or l'ensemble $\{(i, k)\}_{k\in\mathbb{N}}$ est évidemment en bijection avec $\mathbb{N}$. Par composition nous avons le résultat. □

Lemme 1.131 ([1])**.**
Si il existe une surjection $\mathbb{N} \to A$, alors A est fini ou dénombrable.

Démonstration. Pour chaque $a \in A$, l'ensemble $f^{-1}(a)$ est une partie de $\mathbb{N}$.

(i) **Une application** La proposition 1.60(2) nous permet de poser

$$\begin{aligned} \sigma\colon A &\to \mathbb{N} \\ a &\mapsto \min\big(f^{-1}(a)\big). \end{aligned} \tag{1.205}$$

(ii) **σ est injective** Supposons que $\sigma(a) = \sigma(b)$. Nous appelons x ce nombre :

$$x = \min\big(f^{-1}(a)\big) = \min\big(f^{-1}(b)\big). \tag{1.206}$$

Nous avons $x \in f^{-1}(a) \cap f^{-1}(b)$, ce qui implique que $f(x) = a$ et que $f(x) = b$; donc $a = b$. Donc σ est une injection.

(iii) **A est infini** Si A est fini, le lemme est prouvé. Donc à partir de maintenant nous supposons que A est infini. Le but est de prouver qu'il est dénombrable, c'est-à-dire de construire une bijection $A \to \mathbb{N}$.

(iv) **$\sigma(A)$ est dénombrable** Puisque $\sigma\colon A \to \mathbb{N}$ est injective et que A est infini, la proposition 1.118 dit que $\sigma(A)$ est infini dans $\mathbb{N}$. La proposition 1.126 nous dit alors que $\sigma(A)$ est dénombrable.

Soit une bijection $\varphi\colon \sigma(A) \to \mathbb{N}$.

(v) **La candidate bijection** Nous posons

$$f = \varphi \circ \sigma\colon A \to \mathbb{N} \tag{1.207}$$

et nous allons prouver que c'est une bijection.

(vi) **Injective** Puisque φ et σ sont injectives, l'égalité $(\varphi\sigma)(a) = (\varphi\sigma)(b)$ implique immédiatement $a = b$.

(vii) **Surjective** Soit $k \in \mathbb{N}$. Puisque φ et σ sont des injections, nous pouvons poser $a = (\sigma^{-1}\varphi^{-1})(k)$. Il est alors immédiat que $f(a) = k$.

□

Proposition 1.132 ([1, 24]).
Une union dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable.

Démonstration. Soient A_i des ensembles finis ou dénombrables. Nous posons $A = \bigcup_{i\in\mathbb{N}} A_i$, et nous considérons les parties $\mathbb{N}_i$ du corolaire 1.130. Puisque A_i est dénombrable ou fini et que $\mathbb{N}_i$ est dénombrable, il existe une surjection $\varphi_i\colon \mathbb{N}_i \to A_i$.

Nous définissons $s\colon \mathbb{N} \to \mathbb{N}$ par $n \in \mathbb{N}_{s(n)}$, et nous posons enfin

$$\begin{aligned} \varphi\colon \mathbb{N} &\to A \\ n &\mapsto \varphi_{s(n)}(n). \end{aligned} \tag{1.208}$$

Nous prouvons que φ est surjective.

Soit $a \in A_i$. Il existe $n \in \mathbb{N}_i$ tel que $a = \varphi_i(n)$. Mais comme $n \in \mathbb{N}_i$, nous avons $s(n) = i$. Donc

$$a = \varphi_i(n) = \varphi_{s(n)}(n) = \varphi(n). \tag{1.209}$$

Donc $\varphi\colon \mathbb{N} \to A$ est surjective.

Le lemme 1.131 conclut que A est fini ou dénombrable. □

Lemme 1.133 ([1]).
Si N est un ensemble dénombrable, alors il existe une bijection $g\colon \{1,2\} \times N \to N$.

Démonstration. D'abord nous définissons une bijection $\varphi\colon \{0,1\} \times \mathbb{N} \to \mathbb{N}$ par

$$\begin{aligned} \varphi\colon \{0,1\} \times \mathbb{N} &\to \mathbb{N} \\ (n,k) &\mapsto 2k + n. \end{aligned} \tag{1.210}$$

Ensuite si $f\colon \mathbb{N} \to N$ est une bijection, il suffit de poser $g(n,k) = f\big(\varphi(n,k)\big)$. □

Proposition 1.134 ([24]).
Si N est un ensemble dénombrable, alors pour tout $n \in \mathbb{N}$, l'ensemble N^n est dénombrable.

Les ensembles dénombrables sont les plus petits ensembles infinis possibles, comme en témoigne la proposition suivante.

Proposition 1.135.
Tout ensemble infini contient une partie en bijection avec $\mathbb{N}$.

Démonstration. Soient un ensemble infini E_0 et une partie propre E_1 en bijection avec E_0. Nous notons $\varphi\colon E_0 \to E_1$ une bijection.

Soit $x_0 \in E_0 \backslash E_1$ (axiome du choix et tout ça). Nous définissons

$$\begin{aligned} \psi\colon \mathbb{N} &\to \{\varphi^k(x_0)\} \\ n &\mapsto \varphi^n(x_0) \end{aligned} \tag{1.211}$$

et nous allons prouver que c'est une bijection. Que ce soit surjectif est immédiat. Pour l'injectivité, soit $\varphi^k(x_0) = \varphi^l(x_0)$ avec $k \neq l$. Supposons pour fixer les notations que $k > l$. Alors, vu que φ est inversible nous pouvons écrire

$$x_0 = \varphi^{k-l}(x_0) = \varphi\big(\varphi^{k-l-1}(x_0)\big) \tag{1.212}$$

où il est entendu que $\varphi^0(x_0) = x_0$. Cela signifie que x_0 est dans l'image de φ, c'est-à-dire dans E_1, ce que nous avons exclu par choix de x_0 dans $E_0 \backslash E_1$. Donc en réalité $\varphi^k(x_0) \neq \varphi^l(x_0)$ dès que $k \neq l$. □

Proposition 1.136.
Toute partie d'un ensemble fini est finie, et toute partie d'un ensemble dénombrable est finie ou dénombrable.

Lemme 1.137.
Soit un ensemble E non dénombrable ainsi qu'une application $f\colon E \to F$ où F est un ensemble quelconque. Si $f(E)$ est dénombrable (ou fini), alors il existe $y \in f(E)$ tel que $f^{-1}(y)$ est indénombrable.

Démonstration. Nous avons

$$E = \bigcup_{y \in F} f^{-1}(y). \tag{1.213}$$

Si tous les $f^{-1}(y)$ sont dénombrables, alors E est une union dénombrable (F est dénombrable) d'ensembles dénombrables. Il serait donc dénombrable (proposition 1.132), ce qui est contraire à l'hypothèse. □

1.5.4 Théorème de Cantor-Schröder-Bernstein

Lemme 1.138 ([25])**.**
Soient un ensemble A et une partie B de A. Si il existe une injection $f\colon A \to B$ alors il existe une bijection $\alpha\colon A \to B$.

Nous donnons deux preuves de ce lemme.

Première preuve de 1.138. Nous posons $Y = A \backslash B$ et nous décomposons la preuve en étapes.

(i) **Les $f^k(Y)$ sont disjoints** Vu que f prend ses valeurs dans B, nous avons $f^k(Y) \subset B$ pour tout k. Et vu que $Y = A \backslash B$, nous avons

$$f^k(Y) \cap Y = \varnothing \tag{1.214}$$

pour tout k. L'application f étant injective, elle vérifie $f(C \cap D) = f(C) \cap f(D)$. Nous appliquons f^m des deux côtés de (1.214) :

$$f^{k+m}(Y) \cap f^m(Y) = \varnothing \tag{1.215}$$

pour tout $k, m \in \mathbb{N}$.

(ii) **Une décomposition** Nous posons

$$X = \bigcup_{k \in \mathbb{N}} f^k(Y) = Y \cup \bigcup_{k=1}^{\infty} f^k(Y). \tag{1.216}$$

Vu que $f(X) \subset B$ nous avons l'égalité

$$B = f(X) \cup \big(B \backslash f(X)\big). \tag{1.217}$$

(iii) $A\backslash X = B\backslash f(X)$ Supposons $x \in A\backslash X$. Vu que $Y = A\backslash B$ est dans X, l'élément x n'est pas dans $A\backslash B$ et donc est dans B parce qu'il est dans A. Mais x n'est pas dans X et en particulier pas dans $f(X)$ parce que $f(X) \subset X$. Donc x est dans $B\backslash f(X)$.

Dans l'autre sens, nous supposons que $x \in B\backslash f(X)$. Vu que $B \subset A$ nous avons $x \in A$. Comme x est hors de $f(X)$, il est hors des $f^k(Y)$ pour $k \geqslant 1$. Mais $x \in B$, donc x est hors de $A\backslash B = f^0(Y)$. Donc x est hors de $f^k(Y)$ pour tout $k \geqslant 0$. Donc x est hors de X.

(iv) **La bijection** Nous considérons l'application

$$\begin{aligned} \alpha\colon A &\to B \\ x &\mapsto \begin{cases} f(x) & \text{si } x \in X \\ x & \text{si } x \in A\backslash X. \end{cases} \end{aligned} \tag{1.218}$$

Nous démontrons dans les points suivants que α est bijective.

(v) **Injective** Nous supposons $\alpha(x) = \alpha(y)$. Il y a 4 possibilités suivant que x et y soient dans X ou $A\backslash X$.

Si $x, y \in X$ alors $f(x) = f(y)$ et donc $x = y$ parce que f est injective.

Si $x \in X$ et $y \in A\backslash X$, alors $f(x) = y$. Mais $f(x) \in f(X)$ et $y \in A\backslash X = B\backslash f(X)$. Donc l'élément $f(x) = y$ est dans $f(X) \cap \big(B\backslash f(X)\big) = \emptyset$. Il n'est donc pas possible d'avoir $\alpha(x) = \alpha(y)$ avec $x \in X$ et $y \in A\backslash X$.

Si $x \in A\backslash X$ et $y \in X$, c'est la même chose.

Si $x, y \in A\backslash X$, alors $x = \alpha(x) = \alpha(y) = y$.

(vi) **Surjective** Soit $y \in B$. Il y a deux possibilités : $y \in X$ et $y \in A\backslash X$. La première se divise en deux : $y \in Y$ et $y \in \bigcup_{k=1}^{\infty} f^k(Y)$. On y va.

(i) $y \in Y$ Ce cas n'est pas possible parce que $y \in B$ alors que $Y = A\backslash B$.

(ii) $y \in f^k(Y)$ **avec** $k \geqslant 1$ Nous avons

$$y \in f\big(f^{k-1}(Y)\big) \subset f(X) \subset \alpha(A). \tag{1.219}$$

(iii) $y \in A\backslash X$ Alors $y = \alpha(y)$.

□

Deuxième preuve de 1.138[11]. Nous posons $Y = A\backslash B$ et

$$\mathcal{M} = \{M \subset A \text{ tel que } Y \cup f(M) \subset M\}. \tag{1.220}$$

Nous allons dire de nombreuses choses à propos de ce $\mathcal{M}$.

(i) **$\mathcal{M}$ est non vide** Nous avons $A \in \mathcal{M}$ parce que Y et $f(A)$ sont dans A.

(ii) **Si $M \in \mathcal{M}$ alors $Y \subset M$** C'est dans la définition de $\mathcal{M}$.

(iii) **Encore un ensemble** Vu que $\mathcal{M}$ est non vide, nous pouvons poser

$$X = \bigcap_{M \in \mathcal{M}} M \tag{1.221}$$

sans nous poser trop de questions. Cela étant fait, nous pouvons passer aux choses sérieuses.

(iv) $\underline{f(X) \subset M \textbf{ pour tout } M \in \mathcal{M}}$ Soit $M \in \mathcal{M}$. Nous avons

$$\begin{aligned} f(X) &\subset f(M) && (1.222a)\\ &\subset Y \cup f(M) && (1.222b)\\ &\subset M && (1.222c) \end{aligned}$$

Justifications.

— Pour (1.222a). Parce que $X \subset \mathcal{M}$.

— Pour (1.222c). Parce que $M \in \mathcal{M}$.

(v) $\underline{X \in \mathcal{M}}$ Vu que $Y \subset M$ pour tout M dans $\mathcal{M}$, nous avons $Y \subset \bigcap_{M\in\mathcal{M}} M = X$. Nous devons prouver $f(X) \subset X$. Nous venons de prouver que $f(X) \subset M$ pour tout $M \in \mathcal{M}$, donc

$$f(X) \subset \bigcap_{M\in\mathcal{M}} M = X. \tag{1.223}$$

L'ensemble X est le plus petit élément de $\mathcal{M}$ pour l'inclusion.

(vi) $\underline{Y \cup f(X) \subset X}$ Ça fait partie de $X \in \mathcal{M}$. Mais c'est bien de le dire explicitement parce que nous allons l'utiliser quelques fois dans la suite.

(vii) $\underline{Y \cup f(X) \in \mathcal{M}}$ Vu que $X \in \mathcal{M}$, nous savons déjà que $Y \cup f(X) \subset X$. En appliquant f des deux côtés,

$$f\big(Y \cup f(X)\big) \subset f(X). \tag{1.224}$$

En ajoutant Y des deux côtés,

$$Y \cup f\big(Y \cup f(X)\big) \subset Y \cup f(X), \tag{1.225}$$

et donc $Y \cup f(X) \in \mathcal{M}$.

(viii) $\underline{Y \cup f(X) = X}$ Nous savons déjà que $Y \cup f(X) \in \mathcal{M}$. Vu que X est le plus petit élément de $\mathcal{M}$, nous avons

$$X \subset Y \cup f(X). \tag{1.226}$$

L'inclusion inverse étant déjà faite, nous avons l'égalité.

(ix) $\underline{B\backslash f(X) = A\backslash X}$ Nous avons :

$$\begin{aligned} A\backslash X &= A\backslash\big(Y \cup f(X)\big) && (1.227a)\\ &= (A\backslash Y) \cap \big(A\backslash f(X)\big) && (1.227b)\\ &= B \cap \big(A\backslash f(X)\big) && (1.227c)\\ &= (B \cap A)\backslash f(X) && (1.227d)\\ &= B\backslash f(X). && (1.227e) \end{aligned}$$

Justifications :

— Pour (1.227b). Complémentaire de réunion, lemme 1.25(2).

— Pour (1.227c). Parce que $A\backslash Y = B$ du fait que $Y = A\backslash B$.

— Pour (1.227d). Lemme 1.25(3).

— Pour (1.227e). Parce que $B \subset A$.

(x) **Notre bijection** Nous voulons définir

$$\begin{aligned} \alpha\colon A &\to B\\ x &\mapsto \begin{cases} f(x) & \text{si } x \in X\\ x & \text{si } x \in A\backslash X. \end{cases} \end{aligned} \tag{1.228}$$

Pour y parvenir, nous devons prouver que α prend effectivement ses valeurs dans B. Ensuite nous prouverons que α est une bijection.

(xi) **α prend ses valeurs dans B** Si $x \in X$ nous avons $\alpha(x) = f(x) \in B$. Si au contraire $x \in A\backslash X$ nous avons

$$\alpha(x) = x \in A\backslash X = B\backslash f(X) \subset B. \tag{1.229}$$

(xii) **α est injective** Soient $x, y \in A$ tels que $\alpha(x) = \alpha(y)$. Il y a quatre possibilités suivant que x et y sont dans X ou dans $A\backslash X$.

(i) **$x \in X$, $y \in X$** Alors $f(x) = \alpha(x) = \alpha(y) = f(y)$. Vu que f est injective, nous trouvons que $x = y$.

(ii) **$x \in X$, $y \in A\backslash X$** Nous avons $\alpha(x) = f(x) \in f(X)$ et $\alpha(y) = y \in A\backslash X = B\backslash f(X)$. Il n'est donc pas possible d'avoir $\alpha(x) = \alpha(y)$ dans ce cas.

(iii) **$x \in A\backslash X$, $y \in X$** Idem.

(iv) **$x \in A\backslash X$, $y \in A\backslash X$** Dans ce cas nous avons $\alpha(x) = x$ et $\alpha(y) = y$. Donc $x = y$.

(xiii) **α est surjective** Soit $b \in B$. Il y a deux possibilités : $b \in f(X)$ ou $b \notin f(X)$.

(i) **Si $b \in f(X)$** Soit $x \in X$ tel que $f(x) = b$. Alors $\alpha(x) = f(x) = b$.

(ii) **Si $b \notin f(X)$** Alors $b \in B\backslash f(X) = A\backslash X$, et donc $\alpha(b) = b$.

□

Théorème 1.139 (Cantor-Schröder-Bernstein).
Soient deux ensembles A et B pour lesquels il existe des injections $f\colon A \to B$ et $g\colon B \to A$. Alors il existe une bijection entre A et B.

Démonstration. La composée $g \circ f\colon A \to A$ est injective et prend ses valeurs dans $g\big(f(A)\big) \subset g(B) \subset A$. Bref, l'application $g \circ f\colon A \to g(B)$ est injective. Le lemme 1.138 donne alors une bijection $\varphi\colon A \to g(B)$.

Nous montrons que $g^{-1} \circ \varphi\colon A \to B$ est une bijection.

(i) **Injective** Nous supposons $x, y \in A$ tels que

$$g^{-1}\big(\varphi(x)\big) = g^{-1}\big(\varphi(y)\big). \tag{1.230}$$

Nous appliquons g des deux côtés : $\varphi(x) = \varphi(y)$. Puisque φ est une bijection, cela entraîne $x = y$.

(ii) **Surjective** Soit $b \in B$. En posant $a = \varphi^{-1}\big(g(b)\big)$ nous avons bien $(g^{-1} \circ \varphi)(a) = b$.

□

1.5.5 Comparabilité cardinale

Le théorème de comparabilité cardinale énonce que si A et B sont des ensemble, alors nous avons toujours $A \geq B$ ou $A \leq B$ (ou les deux ; dans ce cas $A \approx B$ par Cantor-Schröder-Bernstein).

Théorème 1.140 (Théorème de comparabilité cardinale[1, 26, 27]).
Entre deux ensembles, il existe forcément une injection de l'un dans l'autre.

Démonstration. Nous allons montrer que le graphe d'une injection de A dans B ou de B dans A est donné par un élément maximal (au sens de l'inclusion) de l'ensemble (inductif) des graphes d'injections d'une partie de A dans une partie de B.

Nous allons utiliser le lemme de Zorn 1.22 à l'ensemble [52]

$$\mathcal{A} = \left\{ (X, Y, \varphi) \text{ tel que } \begin{cases} X \subset A \\ Y \subset B \\ \varphi\colon X \to Y \text{ est injective.} \end{cases} \right\} \tag{1.231}$$

52. Attention : dans le Frido, la notation $f\colon A \to B$, signifie que f est définie sur tout l'ensemble A, mais pas qu'elle est surjective sur B.

que nous ordonnons par l'inclusion, c'est-à-dire par $(X_1, Y_1, \varphi_1) < (X_2, Y_2, \varphi_2)$ lorsque $X_1 \subset X_2$, $Y_1 \subset Y_2$ et $\varphi_2|_{X_1} = \varphi_1$.

Nous passons rapidement sur le fait que cet ensemble est inductif, et nous considérons tout de suite un élément maximal (X, Y, φ). Il y a deux possibilités : soit $\varphi(X) = B$, soit $\varphi(X) \neq B$.

(i) **Si $\varphi(X) = B$** Dans ce cas, $\varphi\colon X \to B$ est une surjection. L'ensemble A est donc surpotent à B. La proposition 1.110 conclut que B est subpotent à A.

(ii) **Si $\varphi(X) \neq B$** Nous subdivisons en deux nouveaux cas : soit $X = A$, soit $X \neq A$.

(i) **Si $X = A$** Alors nous avons une injection $\varphi\colon A \to B$, et c'est bon.

(ii) **Si $X \neq A$** Nous sommes dans le cas $X \neq A$ et $\varphi(X) \neq B$. Soient $a \in A \backslash X$ et $b \in B \backslash \varphi(X)$. Nous considérons l'application

$$\begin{aligned} \psi\colon X \cup \{a\} &\to Y \cup \{b\} \\ x &\mapsto \begin{cases} \varphi(x) & \text{si } x \in X \\ b & \text{si } x = a. \end{cases} \end{aligned} \tag{1.232}$$

C'est une application injective. Donc le triplet $(X \cup \{a\}, Y \cup \{b\}, \psi)$ majore (X, Y, φ). Nous avons une contradiction et ce cas n'est pas possible.

□

1.141.
Le théorème de comparabilité cardinale couplé au théorème de Cantor-Schröder-Bernstein nous indique que pour tout ensembles A et B, nous avons soit $A \preceq B$, soit $B \preceq A$. Et si $A \preceq B \preceq A$, alors $A \approx B$.

Nous ne sommes pas loin de dire que la relation $\preceq$ donne un ordre total sur l'ensemble des ensembles. C'est très beau sauf que l'ensemble des ensembles n'existe pas [53]. Il faudrait parler de *classe* des ensembles, mais ça nous mènerait trop loin. Toujours est-il que ces deux théorèmes montrent qu'on n'est pas loin d'avoir un ordre sur les ensembles, et que cela est une des bases possibles pour développer les nombres cardinaux.

1.5.6 Théorème de Cantor, ensemble des ensembles

Théorème 1.142 (Cantor[28])**.**
Un ensemble est toujours strictement subpotent à son ensemble des parties.

Démonstration. Soit un ensemble E et son ensemble des parties $\mathcal{P}(E)$. Nous commençons par prouver qu'il n'existe pas de surjection $E \to \mathcal{P}(E)$. Soit en effet une application $f\colon E \to \mathcal{P}(E)$. Nous posons

$$D = \{x \in E \text{ tel que } x \notin f(x)\}. \tag{1.233}$$

Nous prouvons que D ne peut pas être dans l'image de f. Supposons que $y \in E$ soit tel que $f(y) = D$.

(i) **Si $y \in D$** Alors par définition de D, nous avons $y \notin f(y) = D$. Contradiction.

(ii) **Si $y \notin D$** Alors $y \in f(y) = D$, contradiction.

Donc aucune surjection $f\colon E \to \mathcal{P}(E)$ n'existe. En particulier pas de bijections.

Par ailleurs, l'application $g\colon \mathcal{P}(E) \to E$ qui fait $g(\{a\}) = a$ (et n'importe quoi d'autre sur les autres éléments de $\mathcal{P}(E)$) est une surjection $\mathcal{P}(E) \to E$.

Donc $\mathcal{P}(E)$ est toujours strictement surpotent à E. □

1.143.
Le théorème de Cantor implique en particulier qu'il existe (au moins) une infinité dénombrable d'ensembles infinis de cardinalité différentes (plus évidemment une infinité dénombrable d'ensembles finis de cardinalité différentes).

Pour tout ensemble A, il est donc possible de dire « soit E, un ensemble strictement surpotent à A ».

53. Corolaire 1.144.

Corolaire 1.144.
Il n'existe pas d'ensemble contenant tous les ensembles.

Démonstration. Si E était un tel ensemble, nous aurions $\mathcal{P}(E) \subset E$ parce que les éléments de $\mathcal{P}(E)$ sont des ensembles. Or cela donnerait une surjection $E \to \mathcal{P}(E)$ alors que cela est impossible par le théorème de Cantor 1.142. □

1.5.7 Ajouter ou soustraire des cardinalités

Nous allons prouver une série de résultats que nous pourrions résumer en « ajouter ou retrancher des parties de cardinalité plus petite ne change pas la cardinalité ».

Lemme 1.145 ([1])**.**
Si A est infini et B est fini, alors $A \cup B \approx A$.

Démonstration. Nous supposons que A et B sont disjoints[54]. La proposition 1.120 nous permet de considérer une bijection $\psi\colon \{1,\dots,n\} \to B$.

Puisque A est infini, la proposition 1.135 nous permet de considérer $N \subset A$ et une bijection $\varphi\colon \mathbb{N} \to N$.

Maintenant, il s'agit seulement d'insérer B dans A en le mettant « au début » de N et en décalant les autres éléments. La bijection est

$$\begin{aligned} f\colon A \cup B &\to A \\ x &\mapsto \begin{cases} x & \text{si } x \in A\backslash N \\ \varphi\big(\varphi^{-1}(x)+n\big) & \text{si } x \in N \\ \varphi\big(\psi^{-1}(x)\big) & \text{si } x \in B. \end{cases} \end{aligned} \tag{1.234}$$

□

Lemme 1.146.
Si A est infini et si A est surpotent à B, alors $A \approx A \cup B$.

Démonstration. Il existe évidemment une injection $A \to A \cup B$. Donc le théorème de Cantor-Schröder-Bernstein 1.139 nous indique que trouver une injection $A \cup B \to A$ suffira pour la peine.

Nous allons utiliser le lemme de Zorn 1.22 avec l'ensemble

$$\mathcal{A} = \left\{ (X, \varphi_X) \text{ tel que } \begin{cases} X \subset B \\ \varphi_X\colon A \cup X \to A \text{ est injective.} \end{cases} \right\} \tag{1.235}$$

muni de l'ordre de l'inclusion : $(X,\varphi_X) < (Y,\varphi_Y)$ si $X \subset Y$ et $\varphi_Y(x) = \varphi_X(x)$ pour tout $x \in A \cup X$.

(i) <u>$\mathcal{A}$ **est inductif**</u> Soit une famille $\mathcal{F} = \{(X_i,\varphi_i)\}_{i\in I}$ complètement ordonnée indexée par l'ensemble I. En posant $X = \bigcup_{i\in I} X_i$ et $\varphi(x) = \varphi_i(x)$ dès que $x \in A \cup X_i$, l'élément (X,φ) majore $\mathcal{F}$.

(ii) <u>**Un maximum**</u> Le lemme de Zorn nous assure que $\mathcal{A}$ possède (au moins) un élément maximal. Soit un tel élément maximum (X,φ).

(iii) <u>$X \approx B$</u> Ah oui, vous auriez aimé avoir $X = B$. Mais non ; il n'y a pas de garantie. Nous allons montrer que $X \approx B$, et ça suffira.

Vu que $X \subset B$, si X n'est pas équipotent à B, il est strictement inclus dans B. Nous pouvons donc considérer

$$b \in B\backslash X \tag{1.236a}$$
$$a \in A\backslash\varphi(X). \tag{1.236b}$$

54. Adaptez la démonstration au cas où l'intersection n'est pas vide.

Nous considérons alors l'élément $(Y, \psi) \in \mathcal{A}$ défini par

$$Y = X \cup \{b\} \tag{1.237a}$$

$$\psi(x) = \begin{cases} a & \text{si } x = b \\ \varphi(x) & \text{sinon .} \end{cases} \tag{1.237b}$$

Cet élément majore (X, φ).
Donc $X \approx B$.

(iv) **Résumé de la situation** Nous avons $A \approx A \cup X$ ainsi qu'une injection $\varphi\colon A \cup X \to A$ et une bijection $\psi\colon B \to X$.

(v) **Conclusion si A est disjoint de B** Si A et B sont disjoints, nous avons une bijection

$$\begin{aligned} l\colon A \cup B &\to A \\ x &\mapsto \begin{cases} \varphi(x) & \text{si } x \in A \\ \varphi\big(\psi(x)\big) & \text{si } x \in B. \end{cases} \end{aligned} \tag{1.238}$$

(vi) **Conclusion si A n'est pas disjoint de B** Il suffit de poser $C = B \backslash A$ et nous avons

$$A \cup B = [A \cup (A \cap B)] \cup C. \tag{1.239}$$

Cette union est disjointe, $A \cup (A \cap B)$ est surpotent à A et C est subpotent à B. La conclusion est donc encore valable.

□

La proposition suivante sera utilisée en théorie de la mesure, dans l'exemple 14.68.

Proposition 1.147 ([29, 11])**.**
Si S est un ensemble infini alors il existe une bijection $\varphi\colon \{0,1\} \times S \to S$.

Démonstration. Nous posons $A = \{0\} \times S$ et $B = \{1\} \times S$. L'ensemble A est infini et surpotent à B (pas strictement, mais quand même).

Donc A est idempotent à $A \cup B$ par le lemme 1.145. Mais A est idempotent à S, donc

$$S \approx A \approx A \cup B = \{0,1\} \times S. \tag{1.240}$$

□

Corolaire 1.148.
Si A est un ensemble infini, alors A possède deux sous-ensembles disjoints A_1 et A_2 qui sont tous deux en bijection avec A.

Démonstration. La proposition 1.147 donne une bijection $\varphi\colon \{1,2\} \times A \to A$. Il suffit de poser $A_1 = \varphi(1, A)$ et $A_2 = \varphi(2, A)$. □

Maintenant que nous pouvons mettre dans A deux copies disjointes de A, il n'est pas très étonnant que nous puissions en mettre une infinité dénombrable. C'est en substance ce que signifie la proposition suivante.

Proposition 1.149.
Si A est infini, alors $A \times \mathbb{N} \approx A$.

Démonstration. La démonstration se base sur le fait qu'à l'intérieur de A, nous pouvons construire autant de copies de A deux à deux disjointes que nous le voulons. La k^{e} « copie » sera naturellement l'image de $k \times A$.

Voyons tout cela en détail.

(i) **Ce que nous allons faire** Nous allons construire, pour tout $i \geqslant 1$ des parties $A_i, B_i \subset A$ telles que

— $A_i \cap B_i = \varnothing$,

— $A_i, B_i \subset B_{i-1}$,

— $A_i \approx B_i \approx A$,

— $A_i \cap A_j = \varnothing$ si $i \neq j$

(ii) **La construction** Nous commençons à zéro en utilisant le corolaire 1.148 pour construire des parties disjointes A_0 et B_0 de A telles que $A_0 \approx B_0 \approx A$.

Ensuite, puisque $B_0 \approx A$, il existe A_1 et B_1 dans B_0 tels que $A_1 \cap B_1 = \varnothing$ et $A_1 \approx B_1 \approx B_0 \approx A$. Cela est notre construction pour $i = 1$.

Pour la récurrence, puisque $A_i \approx B_i \approx A$, nous considérons A_{i+1} et B_{i+1} dans B_i tels que $A_{i+1} \cap B_{i+1} = \varnothing$ et $A_{i+1} \approx B_{i+1} \approx B_i \approx A$. C'est encore le corolaire 1.148 qui fait le travail.

(iii) **Les propriétés** Nous avons $A_i \cap A_{i+1} = \varnothing$ parce que $A_i \cap A_{i+1} \subset A_i \cap B_i = \varnothing$.

Nous devons encore montrer que $A_i \cap A_j = \varnothing$ dès que $i \neq j$. Supposons que $j > i$. Nous avons les inclusions

$$A_j \subset B_{j-1} \subset B_{j-2} \subset \ldots \subset B_i. \tag{1.241}$$

Donc $A_j \cap A_i \subset B_i \cap A_i = \varnothing$.

(iv) **Une injection** Nous pouvons à présent écrire une injection qui termine presque la preuve. Pour cela nous considérons pour tout i, une bijection $\psi_i \colon A \to A_i$. Ensuite nous posons

$$\begin{aligned} \varphi \colon A \times \mathbb{N} &\to A \\ (a, k) &\mapsto \psi_k(a). \end{aligned} \tag{1.242}$$

Si $\varphi(a, k) = \varphi(b, l)$, alors $\psi_k(a) = \psi_l(b)$. L'élément $\psi_k(a)$ est donc dans $A_k \cap A_l$; ce n'est possible que si $k = l$. Donc $\psi_l(a) = \psi_l(b)$. Cette dernière égalité n'est possible que si $a = b$ parce que ψ_l est une bijection.

Donc φ est une injection, et nous avons prouvé que $A \times \mathbb{N} \preceq A$.

(v) **La bijection** Nous venons de prouver que $A \times \mathbb{N} \preceq A$. La surpotence $A \times \mathbb{N} \succeq A$ étant évidente, le théorème de Cantor-Schröder-Bernstein 1.139 conclut que $A \times \mathbb{N} \approx A$.

□

Lemme 1.150 ([1]).
Sois un ensemble A muni de deux sous-ensembles B et B' équipotents et disjoints. Alors $A \backslash B$ est équipotent à $A \backslash B'$.

Démonstration. Soit une bijection $\psi \colon B' \to B$. L'application

$$\begin{aligned} \varphi \colon A \backslash B &\to A \backslash B' \\ x &\mapsto \begin{cases} x & \text{si } x \notin B' \\ \psi(x) & \text{si } x \in B'. \end{cases} \end{aligned} \tag{1.243}$$

est la bijection cherchée. □

Lemme 1.151 ([30]).
Si A est un ensemble infini et si $B \prec A$, alors $A \approx A \backslash B$.

Démonstration. Nous pouvons écrire

$$A = (A \backslash B) \cup B. \tag{1.244}$$

Le théorème de comparabilité cardinale 1.140 nous indique que soit $A \backslash B \preceq B$, soit $A \backslash B \succeq B$. Nous allons étudier les deux cas.

(i) **Si $A\backslash B \geq B$** Dans ce cas, $(A\backslash B) \cup B \approx A\backslash B$ par le lemme 1.146. Alors, notre résultat est prouvé parce que $A = (A\backslash B) \cup B \approx A\backslash B$.

(ii) **Si $A\backslash B \leq B$** Dans ce cas, le lemme 1.146 nous indique que $A = (A\backslash B) \cup B \approx B$. Mais $A \approx B$ est exclu par l'hypothèse. Ce cas est donc impossible.

□

Lemme 1.152 ([1])**.**
Si A est infini et si B est une partie strictement subpotente de A, alors il existe $U \subset A$ disjoint de B et équipotent à B.

Démonstration. Le lemme 1.151 nous donne une bijection $\varphi\colon A \to A\backslash B$. Il suffit alors de poser $U = \varphi(B)$. Cette partie est disjointe de B parce que φ prend ses valeurs dans $A\backslash B$. □

Lemme 1.153 ([31])**.**
Soit un ensemble infini A ainsi qu'un sous-ensemble $B \subset A$. Nous supposons l'existence d'une fonction surjective $f\colon B \to B \times B$.

Alors $B \preceq B \times B \preceq B \preceq A$.

Démonstration. La première est l'hypothèse sur f. La seconde est l'existence (évidente) d'une surjection $B \times B \to B$. La troisième est le fait que B soit inclus dans A. □

Lemme 1.154 ([31])**.**
Soit un ensemble infini A ainsi qu'un sous-ensemble strictement subpotent $B \subset A$. Nous supposons l'existence d'une fonction surjective $f\colon B \to B \times B$.

Alors f peut être étendue en une injection $f\colon D \to D \times D$ où $D \subset A$ contient strictement B.

Démonstration. Par le lemme 1.152, nous considérons une partie $U \subset A$ disjointe de B et équipotente à B. Nous pouvons écrire le développement

$$(B \cup U) \times (B \cup U) = (B \times B) \cup (B \times U) \cup (U \times B) \cup (U \times U). \tag{1.245}$$

Nous savons que B est surpotent à U (il est même équipotent) ; donc le lemme 1.146 nous dit que $B \cup U \approx B$. De plus il existe une bijection $B \to U$, donc

$$U \approx B \approx B \times B \approx B \times U \approx U \times B \approx U \times U. \tag{1.246}$$

Chacun des ensembles $U \times B$, $B \times U$ et $U \times U$ est équipotent à U. Leur union est donc équipotente [55] à U et nous avons une bijection

$$\varphi\colon U \to U \times B \cup (B \times U) \cup (U \times U). \tag{1.247}$$

Et enfin nous définissons

$$\begin{aligned} g\colon B \cup U &\to (B \times U) \times (B \cup U) \\ x &\mapsto \begin{cases} f(x) & \text{si } x \in B \\ \varphi(x) & \text{si } x \in U. \end{cases} \end{aligned} \tag{1.248}$$

Cette définition est bonne parce que U et B sont disjoints, et g est injective. □

Le théorème suivant est une généralisation de la proposition 1.147. Elle implique, entre autres choses, qu'il existe une bijection entre $\mathbb{R}$ et $\mathbb{R} \times \mathbb{R}$. Pour le cas de $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$, il y a la proposition 1.129 qui donne une bijection explicite et donc sans axiome du choix et sans lemme de Zorn.

Théorème 1.155.
Si A est infini, alors $A \approx A \times A$.

55. Lemme 1.146.

Démonstration. Une fois de plus, ce sera le lemme de Zorn qui va s'y coller. Soit l'ensemble

$$\mathcal{A} = \left\{ (X, \varphi) \text{ tel que } \begin{cases} X \subset A \\ \varphi\colon X \to X \times X \text{ est surjective.} \end{cases} \right\} \tag{1.249}$$

Cet ensemble est non vide parce que A est infini ; il contient donc une partie dénombrable N, et nous connaissons la surjection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ du lemme 1.129.

Nous ordonnons $\mathcal{A}$ par l'inclusion : $(X, \varphi) \leqslant (Y, \phi)$ lorsque $X \subset Y$ et $\phi|_X = \varphi$. La tambouille usuelle montre que $\mathcal{A}$ est un ensemble inductif et le lemme de Zorn 1.22 donne l'existence d'un élément maximal que nous notons (B, φ).

Puisque B est subpotent à A (parce qu'il est inclus), soit il est strictement subpotent, soit il est équipotent. Nous commençons par montrer que B ne peut pas être strictement subpotent à A.

En effet, si nous avions une surjection $B \to B \times B$, alors que B est strictement subpotent à A. Le lemme 1.154 nous dit alors que φ peut être étendue, ce qui contredirait la maximalité de (B, φ).

Donc la partie B est équipotente à A : il existe une bijection $g\colon A \to B$. Mais nous avons une surjection $B \to B \times B$ et donc aussi une injection $B \times B \to B$. Vu que nous avons par ailleurs une injection $B \to B \times B$, le théorème de Cantor-Schröder-Bernstein 1.139 nous donne une bijection $\phi\colon B \times B \to B$. Avec ça, l'application

$$\begin{aligned} f\colon A \times A &\to A \\ (a, b) &\mapsto \phi\big(g(a), g(b)\big) \end{aligned} \tag{1.250}$$

est une bijection. Donc les ensembles A et $A \times A$ sont équipotents. □

Lemme 1.156.
Si A est infini, et si pour tout $i \in \mathbb{N}$ nous avons $A_i \approx A$, alors

$$\bigcup_{i \in \mathbb{N}} A_i \approx A. \tag{1.251}$$

Démonstration. Pour chaque $i \in \mathbb{N}$ nous avons une bijection $\varphi_i\colon A_i \to A$. Nous posons

$$\begin{aligned} \varphi\colon A \times \mathbb{N} &\to \bigcup_{i=0}^{\infty} A_i \\ (a, i) &\mapsto \varphi_i(a). \end{aligned} \tag{1.252}$$

Cette application est surjective mais peut-être pas injective parce que les A_i peuvent avoir des intersections non vides. Nous avons alors le calcul

$$A \approx A \times \mathbb{N} \tag{1.253a}$$

$$\geq \bigcup_{i=0}^{\infty} A_i \tag{1.253b}$$

$$\geq A \tag{1.253c}$$

Justifications :

— Pour (1.253a), c'est la proposition 1.149.

— Pour (1.253b), c'est la surjection (1.252).

— Pour (1.253c), c'est le fait que seulement A_0 possède déjà une surjection vers A.

Donc $\bigcup_i A_i$ est à la fois surpotent et subpotent à A. Il est donc équipotent par le théorème 1.139. □

1.6 Groupes

1.6.1 Définition, unicité du neutre

La définition d'un groupe est la définition 1.34.

Lemme-Définition 1.157 (Unicités)**.**
Dans un groupe, l'inverse et le neutre sont uniques. Plus précisément, si G est un groupe nous avons :

(1) il existe un unique élément $e \in G$ tel que $eg = ge = g$ pour tout $g \in G$,

(2) pour tout $g \in G$, il existe un unique élément $h \in G$ tel que $gh = hg = e$.

*Le e ainsi défini est nommé **neutre** de G. Le h tel que $gh = hg = e$ est nommé l'**inverse** de g et est noté g^{-1}.*

Démonstration. Chaque point séparément.

(1) Supposons que e_1 et e_2 vérifient la propriété. Nous avons pour tout $g \in G$: $e_1 g = g e_1 = g$. En particulier pour $g = e_2$ nous écrivons $e_1 e_2 = e_2 e_1 = e_2$. Mais en partant dans l'autre sens : $e_2 g = g e_2 = g$ avec $g = e_1$ nous avons $e_2 e_1 = e_1 e_2 = e_1$. En égalant ces deux valeurs de $e_2 e_1$ nous avons $e_1 = e_2$.

Pour la suite de la preuve nous écrivons e l'unique neutre de G.

(2) Supposons que k_1 et k_2 soient deux inverses de g. On considère alors le produit $k_1 g k_2$. Puisque $k_1 g = e$, on a $k_1 g k_2 = e k_2 = k_2$; mais, comme $g k_2 = e$, on a aussi $k_1 g k_2 = k_1 e = k_1$. Le produit est donc à la fois égal à k_1 et à k_2, et donc $k_1 = k_2$.

□

Lemme 1.158.
Soient deux groupes G et H. Si $\alpha \colon G \to H$ est un morphisme de groupes[56]*, alors*

(1) $\alpha(e_G) = e_H$.

(2) $\alpha(g^{-1}) = \alpha(g)^{-1}$.

Lemme-Définition 1.159.
Soit un groupe G. L'ensemble des automorphismes de G, noté $\mathrm{Aut}(G)$, est un groupe pour la composition.

Lemme 1.160.
Si G est un groupe et si $h \in G$, alors les applications

$$\begin{aligned} L_h \colon G &\to G \\ g &\mapsto hg \end{aligned} \tag{1.254}$$

et

$$\begin{aligned} R_h \colon G &\to G \\ g &\mapsto gh \end{aligned} \tag{1.255}$$

sont des bijections.

Démonstration. D'abord si $L_h(g_1) = L_h(g_2)$, alors $hg_1 = hg_2$ et en multipliant à gauche par h^{-1} nous avons $g_1 = g_2$; donc L_h est injective. Ensuite L_h est surjective parce que si $g \in G$, alors $g = L_h(h^{-1}g)$.

Pour l'application R_h, la preuve est une simple adaptation. □

56. Définition 1.35.

1.6.2 Groupe ordonné

Définition 1.161 ([32]).
Soient un groupe $(G,+)$, ainsi qu'une relation d'ordre $\leqslant$ sur G. Nous disons que la relation d'ordre est ***compatible*** *avec la structure de groupe si pour tout $x,y,z\in G$, si $x\leqslant y$ alors $x+z\leqslant y+z$ et $z+x\leqslant z+y$. Dans ce cas, le triple $(G,+,\leqslant)$ est un* ***groupe ordonné****.*

Si $(G,\leqslant)$ est totalement ordonné, nous disons que le groupe est totalement ordonné.

1.6.3 Classes de conjugaison

Définition 1.162 (classe de conjugaison).
Soit un groupe G et un élément $g\in G$. La ***classe de conjugaison*** *de g est la partie*

$$C_g=\{kgk^{-1}\text{ tel que }k\in G\}. \tag{1.256}$$

Lemme 1.163.
Un groupe est commutatif si et seulement si ses classes de conjugaison sont des singletons.

Démonstration. Supposons que G soit commutatif. Alors

$$C_g=\{kgk^{-1}\text{ tel que }k\in G\}=\{g\}. \tag{1.257}$$

Donc les classes de conjugaison sont des singletons.

Dans l'autre sens, si les classes sont des singletons, on a $kgk^{-1}=g$ pour tous $k,g\in G$. Cela signifie immédiatement que G est commutatif. □

Définition 1.164 (centralisateur[33]).
Soient un groupe G, un sous-groupe H et un élément $h\in H$. Le ***centralisateur*** *de h dans G est l'ensemble des éléments de G qui commutent avec h :*

$$Z_G(h)=\{z\in G\text{ tel que }hz=zh\}. \tag{1.258}$$

Le centralisateur de H dans G est l'ensemble des éléments de G qui commutent avec tous les éléments de H :

$$Z_G(H)=\bigcap_{h\in H}Z_G(h). \tag{1.259}$$

Le ***centre*** *d'un groupe G est l'ensemble des éléments de G qui commutent avec tous les autres :*

$$Z_G=Z_G(G)=\{z\in G\text{ tel que }gz=zg,\forall g\in G\}. \tag{1.260}$$

Définition 1.165 (normalisateur[33]).
Soient un groupe G et un sous-groupe H. Le ***normalisateur*** *de H dans G est*

$$\mathcal{N}_G(H)=\{g\in G\text{ tel que }gH=Hg\}. \tag{1.261}$$

Définition 1.166 (Sous-groupe normal).
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 1.167.
Un sous-groupe H de G est un sous-groupe ***caractéristique*** *si $\alpha(H)\subset H$ pour tout automorphisme*[57] *α de G.*

Lemme 1.168 ([34]).
Si H est un sous-groupe caractéristique de G, alors $\alpha(H)=H$ pour tout automorphisme α de G.

57. Automorphisme de groupe, définition 1.35.

Démonstration. Si α est un automorphisme de G, alors α^{-1} est encore un automorphisme de G. En particulier $\alpha^{-1}(H) \subset H$.

Soit $h \in H$. Nous devons prouver que $h \in \alpha(H)$. Pour cela :

$$h = \alpha\big(\alpha^{-1}(h)\big) \in \alpha\big(\alpha^{-1}(H)\big) \subset \alpha(H). \tag{1.262}$$

□

Définition 1.169 (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'élément neutre.*

1.7 Sous-groupe normal

Proposition 1.170.
Soit N un sous-groupe de G. Les propriétés suivantes sont équivalentes :

(1) N est normal[58] *dans G.*
(2) N est une union de classes de conjugaison[59] *de G,*
(3) $gNg^{-1} \subseteq N$ pour tout $g \in G$,
(4) $gNg^{-1} = N$ pour tout $g \in G$,
(5) $gN = Ng$ pour tout $g \in G$,
(6) Le normalisateur[60] *de N est G : $\mathcal{N}_G(N) = G$.*

Démonstration. En plusieurs parties.

(i) **(1) implique (3)** C'est la définition de sous-groupe normal.

(ii) **(3) implique (4)** Soit $g \in G$. Nous avons $gNg^{-1} \subset N$, mais aussi (en appliquant l'hypothèse à g^{-1}) $g^{-1}Ng \subset N$. En combinant nous avons

$$N \subset g(g^{-1}Ng)g^{-1} \subset gNg^{-1}. \tag{1.263}$$

Nous avons l'inclusion dans les deux sens. Donc l'égalité.

(iii) **(4) implique (5)** Soit $g \in G$. Un élément général de gN est de la forme gn avec $n \in N$. Nous devons trouver un $n' \in N$ tel que $gn = n'g$. En posant $n' = gng^{-1}$ nous avons

$$n' = gng^{-1} \in gNg^{-1} \subset N. \tag{1.264}$$

Il est immédiat de prouver que $gn = n'g$. Cela prouve que $gN \subset Ng$.
Le même raisonnement donne $Ng \subset gN$.

(iv) **(5) implique (3)** Un élément de gNg^{-1} est $a = gng^{-1}$ avec $n \in N$. Nous devons prouver que $a \in N$. Puisque $gn \in gN$, par hypothèse il existe n' tel que $gn = n'g$. En remplaçant dans la définition de a,

$$a = gng^{-1} = n'gg^{-1} = n' \in N. \tag{1.265}$$

(v) **(5) implique (2)** Pour chaque $a \in G$ nous notons C_a la classe de conjugaison de a dans G :

$$C_a = \{gag^{-1} \text{ tel que } g \in G\}. \tag{1.266}$$

Comme $a \in C_a$ (prendre $g = e$ dans (1.266).) nous avons forcément

$$N \subset \bigcup_{n \in N} C_n. \tag{1.267}$$

58. Définition 1.166.
59. Définition 1.162.
60. Définition 1.165.

Prouvons maintenant l'inclusion inverse. Nous avons déjà prouvé que (5) implique (3). Donc si $n \in N$, alors $gng^{-1} \in N$. Nous avons alors

$$C_n = \{gng^{-1} \text{ tel que } g \in G\} \subset N. \tag{1.268}$$

Donc il est vrai que $N = \bigcup_{n \in N} C_n$.

(vi) **(2) implique (1)** Nous supposons que $N \subset G$ est un sous-groupe de la forme

$$N = \bigcup_{a \in I} C_a \tag{1.269}$$

où I est une partie de G. Nous devons montrer que pour tout $g \in G$ et pour tout $n \in N$ nous avons $gng^{-1} \in N$. Puisque $n \in N$, il existe $a \in I$ tel que $n \in C_a$ et donc il existe $k \in G$ tel que $n = kak^{-1}$. Nous avons donc

$$gng^{-1} = g(kak^{-1})g^{-1} = (gk)a(gk)^{-1} \in C_a \subset N. \tag{1.270}$$

Le lecteur attentif aura remarqué l'utilisation de l'axiome du choix. La prudence l'incitera à ne pas le faire remarquer au jury.

(vii) **(6) si et seulement si (5)** C'est la définition du normalisateur.

□

Définition 1.171.
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 $g^n = (g^{-1})^{-n}$.

L'ordre d'un groupe et l'ordre d'un élément d'un groupe sont deux choses différentes.

Définition 1.172 (Ordre d'un groupe)**.**
Soit un groupe G.

*(1) Si G est un ensemble fini, l'**ordre** de G est son cardinal*[61]*, et nous le notons $|G|$.*

(2) Si l'ensemble G est infini, nous disons que $|G| = \infty$ et qu'il est d'ordre infini.

Oui : nous pourrions simplement toujours dire « cardinalité » et écrire $\mathrm{Card}(G)$*. Au lieu de ça, dans le cas particulier des groupes, il y a une tradition de dire « ordre » et d'écrire $|G|$.*

Définition 1.173 (Ordre d'un élément)**.**
*L'**ordre** d'un élément g de G est le naturel*

$$\min\{n \in \mathbb{N}\backslash\{0\} \text{ tel que } g^n = e\}, \tag{1.271}$$

si il existe ; dans le cas contraire, nous disons que l'ordre de g est infini.

1.174.
Nous verrons que le corolaire 2.14 au théorème de Lagrange dira que l'ordre d'un élément divise l'ordre du groupe.

Lemme 1.175 ([35, 11])**.**
Soient un groupe G et deux sous-groupes normaux[62] *H et K tels que $H \cap K = \{e\}$. Alors :*

(1) Tout élément de H commute avec tout élément de K.

(2) HK est un sous-groupe de G.

61. Définition 1.120.
62. Sous-groupe normal, définition 1.166.

(3) L'application

$$\begin{aligned}\varphi\colon H\times K &\to HK\\ (h,k) &\mapsto hk\end{aligned}\tag{1.272}$$

est un isomorphisme de groupes.

Démonstration. Point par point.

(i) **(1)** Soient $h\in H$ et $k\in K$. Nous voulons montrer que $hk=kh$. Pour cela nous considérons l'élément $a=hkh^{-1}k^{-1}$. Comme H est normal dans G, nous avons

$$kh^{-1}k^{-1}\in H\tag{1.273}$$

et donc $a\in H$. De même K étant normal dans G, nous avons $hkh^{-1}\in K$ et donc $a\in K$. Au final $a\in H\cap K=\{e\}$. Nous avons prouvé que

$$hkh^{-1}k^{-1}=e,\tag{1.274}$$

et donc que $hk=kh$.

(ii) **(2)** Puisque H et K sont des sous-groupes, $\{e\}$ est dans les deux, de telle sorte que $e\in HK$. De plus si $h_i\in H$ et $k_i\in K$, la commutativité du point (1) donne

$$(h_1k_1)(h_2k_2)=h_1h_2k_1k_2\in HK.\tag{1.275}$$

Donc le produit de deux éléments de HK est dans HK.

(iii) **(3)** En trois sous-parties.

(i) **Morphisme** Soient $h_i\in H$ et $k_i\in K$. En utilisant la commutativité du point (1) nous avons

$$\begin{aligned}\varphi\big((h_1,k_1)(h_2,k_2)\big)&=\varphi(h_1h_2,k_1k_2)&(1.276a)\\ &=(h_1h_2)(k_1k_2)&(1.276b)\\ &=(h_1k_1)(h_2k_2)&(1.276c)\\ &=\varphi(h_1,k_1)\varphi(h_2,k_2).&(1.276d)\end{aligned}$$

(ii) **Injectif** Si $\varphi(h_1,k_1)=\varphi(h_2,k_2)$ nous avons successivement

$$\begin{aligned}h_1k_1&=h_2k_2&(1.277a)\\ h_1k_1h_2^{-1}&=k_2&(1.277b)\\ h_1k_1h_2^{-1}k_1^{-1}&=k_2k_1^{-1}&(1.277c)\\ h_1h_2^{-1}&=k_2k_1^{-1}.&(1.277d)\end{aligned}$$

Le membre de gauche est un élément de H et le membre de droite un élément de K. Comme $H\cap K=\{e\}$ nous avons $h_1h_2^{-1}=e$ et $k_2k_1^{-1}=e$, c'est-à-dire $h_1=h_2$ et $k_1=k_2$.

(iii) **Surjectif** Un élément général de HK est hk avec $h\in H$ et $k\in K$, c'est à dire $\varphi(h,k)$.

□

Définition 1.176.
*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 un tel n, nous disons que l'exposant du groupe est infini.*

Proposition 1.177.
À propos d'exposant de groupe et de ppcm.

(1) *Si l'ensemble des ordres de tous les éléments d'un groupe est majoré, alors l'exposant du groupe est le plus petit commun multiple des ordres des éléments du groupe.*

(2) *Pour un groupe fini, l'exposant est le* ppcm *des ordres des éléments du groupe.*

Le théorème de Burnside 9.293 nous donnera un bon paquet d'exemples de groupes d'exposant fini dans $\mathrm{GL}(n, \mathbb{C})$.

Proposition 1.178.
Soit un groupe G. Nous considérons un sous-groupe normal H de G ainsi qu'un morphisme $\psi\colon G \to H$. Alors

(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

$$\begin{aligned}
[\psi(g_1)][\psi(g_2)] &= \big[\psi(g_1)\psi(g_2)\big] && (1.278a)\\
&= \big[\psi(g_1g_2)\big] && (1.278b)\\
&= \{\psi(g_1g_2)\psi(h) \text{ tel que } h \in H\} && (1.278c)\\
&= \{\psi(g_1g_2h) \text{ tel que } h \in H\} && (1.278d)\\
&= \psi\Big(\{g_1g_2h \text{ tel que } h \in H\}\Big) && (1.278e)\\
&= \psi\big(\overline{g_1g_2}\big) && (1.278f)\\
&= \psi\big(\overline{g_2g_1}\big) && (1.278g)\\
&= \text{refaire à l'envers} && (1.278h)\\
&= [\psi(g_2)][\psi(g_1)]. && (1.278i)
\end{aligned}$$

Par conséquent $\psi(G)/\psi(H)$ est abélien. □

1.7.1 Permutations, groupe symétrique

Nous donnons ici quelques éléments à propos du groupe symétrique. Beaucoup de choses supplémentaires sont reportées à la section 5.6. Voir aussi le thème 7.

Définition 1.179.
*Soit un ensemble E. Une **permutation** de l'ensemble E est une bijection $E \to E$. Le **groupe symétrique** de E est le groupe des bijections $E \to E$; 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 $\{1, \dots, n\} \to \{1, \dots, n\}$.*

Définition 1.180.
*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 1.181 ([35])**.**
*Soient une permutation $\sigma \in E$ ainsi que $a \in E$. La σ-**orbite** de a est l'ensemble*

$$\Omega_\sigma(a) = \{\sigma^i(a)\}_{i\in\mathbb{N}}. \tag{1.279}$$

Lemme 1.182 ([36])**.**
Le groupe symétrique S_n est un ensemble fini contenant $n!$ éléments.

Lemme 1.183 ([37]).
Deux résultats.

(1) Tout groupe est isomorphe à un 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

$$\begin{aligned} \tau_g \colon G &\to G \\ x &\mapsto gx. \end{aligned} \tag{1.280}$$

Commençons par prouver que cela est une bijection. D'une part, $\tau_g(x) = y$ pour $x = g^{-1}y$ (surjection) et, d'autre part, $\tau_g(x) = \tau_g(y)$ implique $gx = gy$ et donc $x = y$ (injection).

Nous avons donc $\tau_g \in S_G$. De plus l'application

$$\begin{aligned} \varphi \colon G &\to S_G \\ g &\mapsto \tau_g \end{aligned} \tag{1.281}$$

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 $\varphi \colon G \to \mathrm{Image}(\varphi)$ 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. □

1.7.2 Décomposition en cycles

Définition 1.184 (cycle[35]).
Soit E un ensemble de cardinal[63] *n. Soit un entier $1 \leqslant k \leqslant n$. Un élément $\sigma \in S_E$ est un k-**cycle** si il ne possède qu'une seule orbite*[64] *non réduite à un élément et qu'elle est de cardinal k.*

Lemme 1.185 ([1]).
Soient un k-cycle σ et $a \in \Omega$. Alors

$$\Omega_\sigma(a) = \{a, \sigma(a), \dots, \sigma^{k-1}(a)\} \tag{1.282}$$

et $\sigma^k(a) = a$.

En particulier, les éléments $\sigma^q(a)$ avec $q = 0, \dots, k-1$ sont tous distincts.

Démonstration. Soit l le plus grand entier tel que les $\sigma^i(a)$ avec $0 \leqslant i \leqslant l$ soient tous distincts, et notons $A = \{\sigma^i(a)\}_{i=0,\dots,l}$. Cet ensemble satisfait

— $\mathrm{Card}(A) = l + 1$

— $A \subset \Omega_\sigma(a)$, et donc $\mathrm{Card}(A) \leqslant \mathrm{Card}(\Omega_\sigma(a)) = k$ par le lemme 1.122(3).

Que vaut $\sigma^{l+1}(a)$? Par maximalité de l, $\sigma^{l+1}(a)$ est un des $\sigma^i(a)$ avec $i \leqslant l$. Par injectivité de σ, nous avons donc forcément $\sigma^{l+1}(a) = a$.

Donc pour tout $i > l$ il existe $j \leqslant l$ tel que $\sigma^i(a) = \sigma^j(a)$ (parce que $\sigma^i(a) = \sigma^{i-l-1}(a)$). Nous en déduisons que

$$\Omega_\sigma(a) = \{\sigma^i(a)\}_{0 \leqslant i \leqslant l} = A. \tag{1.283}$$

Le cardinal de $\Omega_\sigma(a)$ étant k par hypothèse nous avons $k = l + 1$, et donc $l = k - 1$. □

Lemme 1.186.
Si σ est une cycle de longueur k, et si $b \in \Omega_\sigma(a)$, alors

$$\Omega_\sigma(a) = \Omega_\sigma(b) = \{\sigma^i(b)\}_{i=0,\dots,k-1}. \tag{1.284}$$

63. Définition 1.120.
64. Définition 1.181.

Démonstration. Comme $b \in \Omega_\sigma(a)$, il existe $l \leqslant k-1$ tel que $b = \sigma^l(a)$. Pour tout i nous avons $\sigma^i(a) = \sigma^{k-l+1}(b)$, et donc $\Omega_\sigma(a) \subset \Omega_\sigma(b)$.

Mais pour tout i nous avons aussi $\sigma^i(b) = \sigma^{l+1}(a)$ et donc $\Omega_\sigma(b) \subset \Omega_\sigma(a)$.

Nous avons donc montré que $\Omega_\sigma(a) = \Omega_\sigma(b)$. La seconde égalité est le lemme 1.185 appliqué à b. □

Lemme 1.187 ([1]).
Soient un ensemble fini E, une permutation $\sigma \in S_E$ ainsi que $a \in E$. Si $b \in \Omega_\sigma(a)$, alors $\Omega_\sigma(b) = \Omega_\sigma(a)$.

Lemme 1.188 ([35]).
Tout k-cycle est d'ordre[65] *k.*

Démonstration. Soit le cycle $\{a, \sigma(a), \dots, \sigma^{k-1}(a)\}$. Tous les $\sigma^i(a)$ avec $i \leqslant k-1$ sont distincts et $\sigma^k(a) = a$. Donc σ^k est l'identité, et l'ordre de σ est plus petit ou égal à k.

Si $i \leqslant k-1$, alors $\sigma^i(a) \neq a$ parce que les éléments du cycle sont distincts. Donc $\sigma^i \neq \mathrm{Id}$ pour $i \leqslant k-1$. Nous en déduisons que l'ordre de σ est k. □

Lemme 1.189 ([35, 38]).
Tout élément du groupe symétrique S_n peut être décomposé en un nombre fini de cycles de supports disjoints.

Cette décomposition est unique à l'ordre près de l'écriture des cycles.

Plus précisément, si σ est une permutation, alors il existe un unique ensemble fini $\{\omega_i\}_{i\in I}$ de cycles de supports disjoints tels que[66] *$\sigma = \prod_{i\in I} \omega_i$.*

Démonstration. Soit $\sigma \in S_E$. Si les éléments $\{a, \sigma(a), \dots \sigma^k(a)\}$ sont distincts, alors soit $\sigma^{k+1}(a)$ est distincts des autres, soit $\sigma^{k+1}(a) = a$. Il n'est en effet pas possible d'avoir $\sigma^{k+1}(a) = \sigma^l(a)$ avec $l < k$ parce que ça contredirait l'injectivité de σ.

Soit donc $a \in E$. Nous considérons le cycle $\big(a, \sigma(a), \dots, \sigma^k(a)\big)$ où k est maximum tel que tous les éléments sont distincts.

Soit ce cycle contient tous les éléments de E, soit il existe un élément b hors de ce cycle. Dans le second cas, nous considérons le cycle commençant par b.

Et ça continue... □

Lemme 1.190 ([35]).
Deux cycles de support disjoint commutent.

Lemme 1.191 ([39]).
Tout cycle de longueur r est le produit de $r-1$ transpositions.

Démonstration. Il suffit de vérifier que

$$(a_1, \dots, a_r) = (a_1, a_r)(a_1, a_{r-1}) \dots (a_1, a_2). \tag{1.285}$$

□

Lemme 1.192 ([1]).
Soit une permutation $\sigma \in S_E$. Soit un cycle c et une permutation s de supports disjoints telles que $\sigma = s \circ c$. Alors

(1) Si $a \in \mathrm{supp}(\sigma)$, alors pour tout $q \in \mathbb{N}$ nous avons $c^q(a) = \sigma^q(a)$.

(2) Si $a \in \mathrm{supp}(c)$, alors

$$\Omega_\sigma(a) = \Omega_c(a) = \mathrm{supp}(c). \tag{1.286}$$

65. Définition 1.173.
66. Ici I est un ensemble fini et vu que les supports sont disjoints, le produit est commutatif.

(3) Si $a \in \operatorname{supp}(c)$, alors

$$c(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon.} \end{cases} \tag{1.287}$$

Démonstration. En plusieurs parties.

(i) **Si $a \in \operatorname{supp}(c)$, alors $c(a) = \sigma(a)$** Soit $a \in \operatorname{supp}(c)$. Nous savons que $c(a) \neq a$, et vu que c est injective, nous devons aussi avoir $c\big(c(a)\big) \neq c(a)$. Donc a et $c(a)$ sont dans $\operatorname{supp}(c)$. Étant donné que les supports de c et de s sont disjoints, nous déduisons que $c(a)$ n'est pas dans le support de s, et donc que

$$\sigma(a) = (s \circ c)(a) = s\big(c(a)\big) = c(a). \tag{1.288}$$

(ii) $\sigma^q(a) = c^q(a)$ Juste une récurrence sur le point précédent : si $b \in \operatorname{supp}(a)$, alors $\sigma(b) = c(b) \in \operatorname{supp}(a)$.

(iii) **Si $a \in \operatorname{supp}(a)$, alors $\Omega_\sigma(a) = \Omega_c(a)$** Utilisant le point précédent, ainsi que la définition 1.181 d'une orbite,

$$\Omega_\sigma(a) = \{\sigma^q(a)\} = \{c^q(a)\} = \Omega_c(a). \tag{1.289}$$

(iv) $\operatorname{supp}(c) \subset \Omega_c(a)$ Comme toujours, a est un élément de $\operatorname{supp}(c)$. Nous considérons $b \in \operatorname{supp}(c)$ et nous montrons que $b \in \Omega_c(a)$. Étant donné que $b \in \operatorname{supp}(c)$, nous avons $c(b) \neq b$, de telle sorte que $\Omega_c(b)$ contienne au moins deux éléments distincts.
Même chose pour a : l'ensemble $\Omega_c(a)$ contient au moins a et $c(a)$. Vu que c est un cycle, il n'existe qu'une seule orbite non triviale. Donc $\Omega_c(a) = \Omega_c(b)$. En particulier $b \in \Omega_c(b) = \Omega_c(a)$.

(v) $\Omega_c(a) \subset \operatorname{supp}(c)$ Soit $b \in \Omega_c(a)$. Le lemma 1.186 nous permet de dire que $\Omega_c(a) = \Omega_c(b)$. Comme $\Omega_c(b)$ contient au moins deux éléments (parce qu'il est égal à $\Omega_c(a)$ et que a est dans le support de c), nous savons que $c(b) \neq b$ et donc que $b \in \operatorname{supp}(c)$.

(vi) **La formule pour $c(x)$** Si $x \in \Omega_{\sigma)}(a)$, alors $c(x) = \sigma(x)$ par le point (1). Si x n'est pas dans $\Omega_c(a) = \operatorname{supp}(c)$, alors x n'est pas dans le support de c et donc $c(x) = x$.

□

Le lemme suivant permet d'extraire le cycle de σ associé à un élément de E.

Lemme 1.193 ([1]).
Soient un ensemble finie E ainsi que $\sigma \in S_E$, et $a \in E$ tel que $\sigma(a) \neq a$. Nous posons

$$\begin{aligned} c\colon E &\to E \\ x &\mapsto \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon} \end{cases} \end{aligned} \tag{1.290}$$

Alors

(1) Si $b \in \Omega_\sigma(a)$, nous avons $\Omega_c(b) = \Omega_\sigma(a)$.

(2) Si $b \notin \Omega_\sigma(a)$, nous avons $\Omega_c(b) = \{b\}$.

(3) c est un cycle.

Démonstration. Si $b \in \Omega_\sigma(a)$, alors $\sigma^q(b) \in \Omega_\sigma(a)$ pour tout $q \in \mathbb{N}$, et donc

$$c^q(b) = \sigma^q(b) \in \Omega_\sigma(a) \tag{1.291}$$

pour tout q. Donc nous avons

$$\Omega_c(b) = \{c^q(b) \text{ tel que } q \in \mathbb{N}\} = \{\sigma^q(b) \text{ tel que } q \in \mathbb{N}\} = \Omega_\sigma(b) = \Omega_\sigma(a). \tag{1.292}$$

La dernière égalité est le lemme 1.187.

Si $b \notin \Omega_\sigma(a)$, alors $c(b) = b$ et $\Omega_c(b) = \{b\}$.

Nous avons prouvé que c a une seule orbite de taille plus grands ou égale à 2. Donc c est un cycle. □

Théorème 1.194 ([38]).
Soit un ensemble fini E de cardinal au moins deux. Soit une permutation $\sigma \in S_E$.

— *Il existe des cycles $c_1, \dots, c_m$ à support disjoints tels que $\sigma = c_1 \circ \dots \circ c_m$.*

— *Cette décomposition est unique à l'ordre près.*

Démonstration. Plusieurs points.

(i) **Existence** Nous choisissons des éléments $\{a_i\}_{i=1,\dots,p}$ tels que les $\Omega_\sigma(a_i)$ forment une partition de E en sous-ensembles disjoints. En posant $l_k = \min\{r \text{ tel que } \sigma^r(a_k) = a_k\}$, nous avons

$$\Omega_\sigma(a_k) = \{\sigma^q(a_k)\}_{q=1,\dots,l_k-1} \tag{1.293}$$

et tous les $\sigma^q(a_k)$ sont distincts pour $q = 1, \dots, l_k - 1$.

Posons

$$\begin{aligned} c_k \colon E &\to E \\ x &\mapsto \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a_k) \\ x & \text{sinon.} \end{cases} \end{aligned} \tag{1.294}$$

Le lemme 1.192 dit que c_k est un cycle. Vu que $c(a_k) = \sigma(a_k)$, le cycle c est un l_k-cycle.

Nous montrons à présent que $\sigma = c_1 \circ \dots c_p$. Soit $x \in E$. Il existe un $k \in \{1, \dots, p\}$ tel que

$$x \in \Omega_\sigma(a_k) = \Omega_{c_k}(a_k) = \Omega_{c_k}(x), \tag{1.295}$$

la dernière égalité est parce que $x \in \Omega_{c_k}(a_k)$. Nous en déduisons que $c_k(x) = x$. D'autre part si $l \neq k$, alors x n'est pas dans $\Omega_\sigma(a_l)$, et donc $c_l(x) = x$. Au final,

$$(c_1 \circ \dots c_p)(x) = c_k(x) = \sigma(x). \tag{1.296}$$

(ii) **Unicité** Nous supposons avoir $\sigma = c_1 \circ \dots \circ c_p = \gamma_1 \circ \gamma_q$ où les c_i et les γ_j sont deux ensembles de cycles de supports disjoints. Nous avons

$$\operatorname{supp}(\sigma) = \bigcup_{i=1}^{p} \operatorname{supp}(c_i) = \bigcup_{j=1}^{q} \operatorname{supp}(\gamma_j). \tag{1.297}$$

Montrons que si $\operatorname{supp}(c_i) \cap \operatorname{supp}(\gamma_j) \neq \emptyset$, alors $\operatorname{supp}(c_i) = \operatorname{supp}(\gamma_j)$. En effet si $a \in \operatorname{supp}(c_i) \cap \operatorname{supp}(\gamma_j)$, alors

$$\operatorname{supp}(c_i) = \Omega_{c_i}(a) = \Omega_\sigma(a) = \Omega_{\gamma_j}(a) = \operatorname{supp}(\gamma_j). \tag{1.298}$$

Et comme les $\operatorname{supp}(\gamma_j)$ sont disjoints, l'ensemble $\operatorname{supp}(c_i)$ n'a d'intersection qu'avec un et un seul des $\operatorname{supp}(\gamma_j)$. Cela définit donc une application

$$\begin{aligned} u \colon \{1, \dots, p\} &\to \{1, \dots, q\} \\ i &\mapsto \text{l'unique } j \text{ tel que} \operatorname{supp}(c_i) = \operatorname{supp}(\gamma_j). \end{aligned} \tag{1.299}$$

Autrement dit, l'application u permet d'écrire

$$\operatorname{supp}(c_i) = \operatorname{supp}\left(\gamma_{u(i)}\right). \tag{1.300}$$

L'application u est injective. En effet si $u(i) = u(l)$, nous avons

$$\operatorname{supp}(c_i) = \operatorname{supp}\left(\gamma_{u(i)}\right) \tag{1.301a}$$

$$\operatorname{supp}(c_l) = \operatorname{supp}\left(\gamma_{u(l)}\right) \tag{1.301b}$$

$$u(i) = u(l). \tag{1.301c}$$

Donc $\text{supp}(c_i) = \text{supp}(c_l)$. Et comme les supports sont disjoints, $i = l$.

L'application u est surjective. En effet, soit $j \in \{1, \dots, q\}$. Soit $a \in \text{supp}(\gamma_j)$. Il existe un i tel que $a \in \text{supp}(c_i)$. Nous avons alors $a \in \text{supp}(\gamma_j) \cap \text{supp}(c_i)$, autrement dit $u(i) = j$.

Maintenant l'application $u\colon \{1, \dots, p\} \to \{1, \dots, q\}$ est bijective. Nous en déduisons que $p = q$. Concluons en montrant que $c_i = \gamma_{u(i)}$. Soit $a \in \text{supp}(c_i) = \text{supp}\left(\gamma_{u(i)}\right)$.

Nous avons

$$c_i(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon} \end{cases} \tag{1.302}$$

et

$$\gamma_j(x) = \begin{cases} \sigma(x) & \text{si } x \in \Omega_\sigma(a) \\ x & \text{sinon,} \end{cases} \tag{1.303}$$

et donc $c_i = \gamma_j$.

□

Lemme 1.195 ([40]).
Soit $\sigma = (i_1, \dots, i_k) \in S_n$, un cycle de longueur k et $\theta \in S_n$. Alors

$$\theta\sigma\theta^{-1} = \big(\theta(i_1), \dots, \theta(i_k)\big). \tag{1.304}$$

Tous les cycles de longueur k sont conjugués entre eux.

Proposition 1.196 (Classes de conjugaison et structure en cycles[41]).
Une classe de conjugaison[67] *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 c_i de supports disjoints. Si τ est une permutation, alors

$$\sigma' = \tau\sigma\tau^{-1} = (\tau c_1 \tau^{-1}) \dots (\tau c_m \tau^{-1}), \tag{1.305}$$

mais $\tau c_i \tau^{-1}$ est un cycle de même longueur que c_i, puisque le lemme 1.195 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 que σ.

Réciproquement, si $\sigma' = c'_1 \dots c'_m$ est une décomposition de σ' en cycles disjoints tels que la longueur des c_i est la même que la longueur des 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 1.195. Comme les supports des c_i et des c'_i sont disjoints, la permutation $\tau_1 \dots \tau_m$ conjugue σ et σ'. □

Exemple 1.197.
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 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.

67. Définition 1.162.

Ce sont donc

$$C_1 = \{\mathrm{Id}\} \tag{1.306a}$$
$$C_2 = \{(1,2),(1,3),(2,3)\} \tag{1.306b}$$
$$C_3 = \{(1,2,3),(2,1,3)\}. \tag{1.306c}$$

△

Définition 1.198 (transposition)**.**
*Une **transposition** est une permutation*[68] *qui échange deux éléments de E. Plus précisément, une bijection $\sigma\colon E \to E$ est une transposition si il existe $a, b \in E$ tels que*

$$\sigma(x) = \begin{cases} a & si\ x = b \\ b & si\ x = a \\ x & sinon. \end{cases} \tag{1.307}$$

Exemple 1.199.
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 1.196). Le groupe symétrique S_4 possède donc 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.

△

Proposition 1.200.
Tout élément de S_n peut être écrit sous la forme d'un produit fini de transpositions.

Démonstration. Un élément de S_n se décompose en un nombre fini de cycles par le lemme 1.189 et chacun des cycles peut être décomposé en un nombre fini de transpositions par le lemme 1.191. □

Cette décomposition n'est pas à confondre avec celle en cycles de support disjoints. Par exemple $(1, 2, 3) = (1, 3)(1, 2)$.

Le théorème suivant, qui donne la notion de parité d'une permutation, est la clef pour savoir quelles positions du jeu de taquin sont possibles ou impossibles[42, 43].

Proposition-Définition 1.201 (parité d'une permutation)**.**
À propos de décomposition ne permutations.

(1) Si une permutation peut être écrite sous forme d'un produit d'un nombre pair de transpositions, alors toute décomposition en permutations sera en quantité paire.

(2) Si une permutation peut être écrite sous forme d'un produit d'un nombre impair de transpositions, alors toute décomposition en permutations sera en quantité impaire.

*Une permutation qui se décompose en une quantité paire de transpositions est une **permutation paire** (et **impaire** sinon).*

68. Une permutation est une bijection, définition 1.179.

Définition 1.202.
*La **signature** est l'application*

$$\begin{aligned} \epsilon\colon S_E &\to \{-1,1\} \\ \sigma &\mapsto \begin{cases} 1 & \text{si } \sigma \text{ est paire} \\ -1 & \text{si } \sigma \text{ est impaire.} \end{cases} \end{aligned} \tag{1.308}$$

Lemme 1.203.
*Nous disons qu'un élément $\sigma \in S_n$ est une **inversion** pour 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 $\sigma(i) > \sigma(j)$). Nous avons*

$$\epsilon(\sigma) = (-1)^{N_\sigma} \tag{1.309}$$

où ϵ est la signature[69] *dans S_n.*

Lemme 1.204 ([35])**.**
Un k-cycle est une permutation impaire si k est pair et paire si k est impair.

Proposition 1.205 ([40])**.**
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

$$\epsilon(\sigma)\epsilon(\theta) = (-1)^{N_2+N_3}(-1)^{N_3+N_4} = (-1)^{N_2+N_4} = (-1)^{N_{\sigma\theta}} = \epsilon(\sigma\theta). \tag{1.310}$$

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 1.195 dit que

$$\tau_{ij} = \theta\tau_{j-1,j}\theta^{-1}. \tag{1.311}$$

La signature étant un homomorphisme,

$$\epsilon(\tau_{ij}) = \epsilon(\theta)\epsilon(\tau_{j-1,j})\epsilon(\theta)^{-1} = \epsilon(\tau_{j-1,j}) = -1. \tag{1.312}$$

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[70]). Si τ' est une autre transposition, il existe $\sigma \in S_n$ tel que $\tau' = \sigma\tau\sigma^{-1}$ (lemme 1.195). Dans ce cas, $\varphi(\tau') = \varphi(\tau) = -1$, et si $\sigma = (\tau_1 \dots \tau_k)$,

$$\varphi(\sigma) = (-1)^k = \epsilon(\sigma). \tag{1.313}$$

□

69. Définition 1.202.
70. Nous utilisons ici le fait que tous les éléments de S_n sont des produits de transpositions, proposition 1.200.

Corolaire 1.206.
Si $\sigma \in S_n$, alors

$$\epsilon(\sigma) = \epsilon(\sigma^{-1}). \tag{1.314}$$

Démonstration. Comme énoncé par la proposition 1.205, ϵ est un homomorphisme, donc

$$\epsilon(\sigma)\epsilon(\sigma^{-1}) = \epsilon(\sigma\sigma^{-1}) = \epsilon(\mathrm{Id}) = 1. \tag{1.315}$$

Puisque $\epsilon(\sigma)$ et $\epsilon(\sigma^{-1})$ ne peuvent valoir que ± 1, ils doivent être tous les deux égaux à 1 ou tous les deux à -1 pour que le produit soit 1. □

1.8 Anneaux

Définition 1.207.
*Un **isomorphisme d'anneaux** est un morphisme d'anneaux*[71]*, bijectif.*

La distributivité de la partie (3) de la définition 1.38 ne traite que de l'addition ; pas de la soustraction. Voici une lemme qui dit que ça fonctionne quand même.

Lemme 1.208 ([11])**.**
Soient un anneau A ainsi que $a, b, c \in A$. Alors

$$a(b - c) = ab - ac. \tag{1.316}$$

Démonstration. Nous avons le calcul suivant :

$$a(b - c) + ac = a\big((b - c) + c\big) \tag{1.317a}$$
$$= ab. \tag{1.317b}$$

Justifications :

— Pour 1.317a. Distributivité.

— Pour 1.317b. Parce que $(b - c) + c = b$.

Nous avons donc $a(b-c)+ac = ab$ et donc l'égalité demandée en ajoutant $-ac$ des deux côtés. □

Lemme 1.209.
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 sous la forme du lemme 1.208,

$$a \cdot 0 = a \cdot (1 - 1) = a - a = 0. \tag{1.318}$$

Notons que la dernière égalité s'écrit en détail $a-a = a+(-a)$ qui donne le neutre de l'addition. □

Proposition 1.210.
Dans un anneau[72] *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, $a = 0$. □

Lemme 1.211 ([1])**.**
Un peu d'arithmétique. Soit un anneau A et un élément $a \in A$.

(1) $1 \times 1 = 1$.

(2) $(-1) \times a = -a$.

71. Définition 1.39.
72. Définition 1.38.

(3) $-(-a) = a$.

(4) $(-1) \times (-1) = 1$.

Démonstration. En plusieurs parties.

(i) **Pour (1)** La définition de 1 est que $1 \times a = a$ pour tout a. En particulier pour $a = 1$ nous avons le résultat.

(ii) **Pour (2)** Nous avons

$$(-1) \times a + a = a \times \big((-1) + 1\big) = a \times 0 = 0. \tag{1.319}$$

Nous avons utilisé le fait que la multiplication était distributive et que le zéro était absorbant (lemme 1.209).

(iii) **Pour (3)** Nous avons $-a + a = 0$ par définition de la notation $-a$. Donc a est bien l'inverse de $-a$ pour l'addition.

(iv) **Pour (4)** En utilisant les points (2) et (3) nous avons

$$(-1) \times (-1) = -(-1) = 1. \tag{1.320}$$

□

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

$$(f + g)(x) = f(x) + g(x) \tag{1.321a}$$
$$(fg)(x) = f(x)g(x). \tag{1.321b}$$

C'est la **structure canonique** d'anneau sur $\mathrm{Fun}(X, A)$.

Définition 1.212.
*Le **centralisateur** de $x \in A$ dans A est l'ensemble*

$$\{y \in A \text{ tel que } xy = yx\}, \tag{1.322}$$

*le **centre** de A est*

$$\{y \in A \text{ tel que } xy = yx, \forall x \in A\}. \tag{1.323}$$

Définition 1.213 (Idéal dans un anneau)**.**
*Un sous-ensemble $I \subset A$ est un **idéal à gauche** si*

(1) *I est un sous-groupe pour l'addition,*

(2) *pour tout $a \in A$, $aI \subset I$.*

*De même nous disons que $I \subset A$ est une **idéal à droite** lorsque I est un sous-groupe pour l'addition et $Ia \subset I$ pour tout $a \in A$.*

*Lorsqu'un ensemble est idéal à gauche et à droite, nous disons que c'est un **idéal bilatère**. Lorsque nous parlons d'idéal sans précision, nous parlons d'idéal bilatère.*

Proposition-Définition 1.214.
Soit A, un anneau, I un idéal bilatère[73] *de A. Nous considérons la relation d'équivalence $x \sim y$ si et seulement si $x - y \in I$. Sur le quotient*[74]

$$A/\sim = A/I, \tag{1.324}$$

nous mettons les opérations

(1) $[x] + [y] = [x + y]$

73. Définition 1.213.
74. Définition 1.30.

(2) $[x][y] = [xy]$.

Nous avons alors les résultats suivants :

(1) Les opérations sont bien définies,

(2) l'ensemble A/I, muni de ces opérations, est un anneau. Le neutre pour l'addition est $[0]$, l'inverse de $[a]$ est $[-a]$ que nous noterons $-[a]$.

(3) la surjection canonique $\pi\colon A \to A/I$ est un morphisme.

Cet anneau est appelé ***anneau quotient****.*

Démonstration. En plusieurs parties.

(i) **Pour (1)** Nous savons que, par définition,

$$\bar{x} = \{x + i \text{ tel que } i \in I\}. \tag{1.325}$$

Calculons le produit de représentants génériques de $\bar{x}$ et de $\bar{y}$:

$$(x + i_1)(y + i_2) = xy + xi_2 + yi_1 + i_1 i_2. \tag{1.326}$$

Puisque I est un idéal, nous avons $xi_2 + yi_1 + i_1 i_2 \in I$ et donc bien

$$(x + i_1)(y + i_2) \in \overline{xy}. \tag{1.327}$$

(ii) **Pour (2)** Il s'agit de vérifier les conditions de la définition 1.38.
D'abord A/I est un groupe de neutre $[0]$. En effet, vu que $(A, +)$ est un groupe commutatif de neutre 0, nous avons

(1) Neutre : $[a] + [0] = [a + 0] = [a]$.

(2) Associativité : $[a] + ([b] + [c]) = [a] + [b + c] = [a + b + c] = [a + b] + [c]$.

(3) Inversibilité : l'inverse de $[a]$ est $[-a]$ parce que $[a] + [-a] = [a - a] = [0]$.

Nous pouvons noter $-[a]$ l'élément $[-a]$. Le groupe A/I est commutatif :

$$[a] + [b] = [a + b] = [b + a] = [b] + [a]. \tag{1.328}$$

Donc $(A/I, +)$ est un groupe commutatif de neutre $[0]$.
L'associativité de A donne l'associativité dans A/I :

$$\big([a][b]\big)[c] = [ab][c] = [abc] = [a][bc] = [a]\big([b][c]\big). \tag{1.329}$$

Et enfin pour la distributivité,

$$[a]\big([b] + [c]\big) = [a][b + c] = [a(b + c)] = [ab + ac] = [ab] + [ac] = [a][b] + [a][c]. \tag{1.330}$$

Nous avons prouvé que A/I est un anneau de neutre $[0]$ et d'unité $[1]$.

(iii) **Pour (3)** Nous devons vérifier les trois conditions de la définition 1.39. Cela est immédiat parce que $\pi(x) = [x]$.

□

Définition 1.215.
Soient A un anneau commutatif et $S \subset A$. Nous disons que $\delta \in A$ est un ***PGCD*** *de S si*

(1) δ divise tous les éléments de S.

(2) si d divise également tous les éléments de S, alors d divise δ.

Nous disons que $\mu \in A$ est un ***PPCM*** *de S si*

(1) $S \mid \mu$,

(2) si $S \mid m$, alors $\mu \mid m$.

Remarque 1.216.
Au sens de la définition 1.215, le pgcd n'est pas unique. Dans $\mathbb{Z}$ par exemple les nombres 4 et -4 sont tous deux pgcd de $\{4, 16\}$.

Dans $\mathbb{Z}$ cependant, nous modifions implicitement la définition et nous n'acceptons que les positifs, de telle sorte à ce que l'unique pgcd soit effectivement le plus grand pour l'ordre usuel sur $\mathbb{Z}$.

Pour l'unicité dans $\mathbb{Z}$, voir 3.21.

1.8.1 Anneau intègre

Définition 1.217 (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 si 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 $c \in A$.*

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 1.11.

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 1.218 (Éléments nilpotents, unipotents)**.**
*On dit que $a \in A$ est **nilpotent** si 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 $n \in \mathbb{N}$.*

Définition 1.219 (Éléments inversibles)**.**
*Un élément $a \in A$ est dit **inversible** si 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 $A^* = A \backslash \{0\}$.

Conformément à la définition 1.217 de diviseur, nous posons la définition suivante pour les diviseurs de zéro.

Définition 1.220 ([44])**.**
*Un élément $a \neq 0$ est un **diviseur de zéro à gauche** si il existe $x \neq 0$ tel que $ax = 0$. L'élément a est un diviseur de zéro **à droite** si il existe $y \neq 0$ tel que $ya = 0$.*

*Nous disons que a est un **diviseur de zéro** si il est un diviseur de zéro à gauche ou à droite.*

Proposition-Définition 1.221 (Anneau intègre[1])**.**
Soit A un anneau non réduit à $\{0\}$. Les assertions suivantes sont équivalentes :

(1) A ne possède pas de diviseurs de zéro.

(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 anneau non réduit à $\{0\}$ qui vérifie ces propriétés est dit **intègre**.*

Démonstration. En trois implications.

(i) **(1) implique (2)** Si $ab = 0$ avec $b \neq 0$ alors a est un diviseur de zéro. Vu que nous supposons que A n'a pas de diviseurs de zéros, a est nul. De même, si $a \neq 0$ b devrait être nul.

(ii) **(2) implique (3)** Si $ab = ac$, alors $a(b - c) = 0$ et l'hypothèse dit que soit $a = 0$, soit $b - c = 0$. Donc si $a \neq 0$, alors $b - c = 0$.

(iii) **(3) implique (1)** Si $A = \{0\}$, le point (3) n'est pas applicable.
Si $a \neq 0$ et $ax = 0$, alors nous avons aussi $ax = a \times 0$. Par propriété de simplification, $x = 0$. Donc a n'est pas un diviseur de zéro à gauche. Nous prouvons de la même façon qu'il n'y a pas de diviseurs de zéro à droite.

□

1.8.2 Fonction puissance

Voici une première définition de la fonction puissance. Il y en aura d'autres, de plus en plus générales. Voir le thème 50.

Définition 1.222.
Si A est un anneau, si $a \in A$ et si $n \in \mathbb{N}$, nous définissons a^n par récurrence :

(1) $a^0 = 1$ *(l'unité pour la multiplication dans A),*

(2) $a^{k+1} = a \cdot a^k$.

Le lemme suivant dit que le point (2) de la définition 1.222 aurait pu être écrit $a^k \cdot a$ au lieu de $a \cdot a^k$.

Lemme 1.223 ([1]).
Si A est un anneau, si $a \in A$ et si $n \in \mathbb{N}$, alors

$$a^n = a \cdot a^{n-1} = a^{n-1} \cdot a. \tag{1.331}$$

Démonstration. Cela se prouve par récurrence. Pour $n = 1$ c'est l'égalité $a = a^0 a$ qui est correcte parce que par définition $a^0 = 1$.

Supposons que le résultat soit bon pour n et voyons ce que ça donne pour $n + 1$:

$$\begin{aligned}
a^{n+1} &= aa^n && \text{Définition de } a^{n+1} && (1.332a)\\
&= a(a^{n-1}a) && \text{hypothèse de récurrence pour } a^n && (1.332b)\\
&= (aa^{n-1})a && \text{associativité} && (1.332c)\\
&= a^n a && \text{Définition de } a^n. && (1.332d)
\end{aligned}$$

□

1.9 Le groupe et anneau des entiers

Certes $(\mathbb{Z}, +)$ est un groupe mais en ajoutant la multiplication, $(\mathbb{Z}, +, \times)$ devient un anneau [75].

1.9.1 Division euclidienne

Théorème-Définition 1.224 (Division euclidienne[45]).
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

$$a = bq + r. \tag{1.333}$$

*L'opération $(a, b) \mapsto (q, r)$ ainsi définie est la **division euclidienne**. Le nombre q est le **quotient** et r est le **reste** de la division de a par b.*

Démonstration. Remarquons que $r = a - bq$, et donc, une fois l'existence et l'unicité de q établie, celle de r suivra.

(i) **Unicité** Nous supposons avoir $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tels que

$$\begin{cases} 0 \leqslant r < b & (1.334a)\\ a = qb + r. & (1.334b)\end{cases}$$

Ce système implique que

$$0 \leqslant a - qb < b. \tag{1.335}$$

75. Définition 1.38.

En ajoutant qb dans les trois membres de cette inégalité,

$$qb \leqslant a < (q+1)b. \tag{1.336}$$

Cela implique que

$$q = \max\{k \in \mathbb{Z} \text{ tel que } kb \leqslant a\}. \tag{1.337}$$

Donc q est unique et la relation $a = bq + r$ implique que r est également unique.
Soit

$$E = \{q \in \mathbb{Z} | bq \leqslant a\}.$$

La partie E est non vide (parce qu'elle contient $-|a|$) et admet un majorant : l'élément $|a|$. Elle admet donc un maximum q par le lemme 1.106. Ce maximum vérifie

$$bq \leqslant a < b(q+1). \tag{1.338}$$

Cela donne $0 \leqslant a - bq < b$ et le résultat, en posant $r = a - qb$.

□

1.9.2 PGCD, PPCM et Bézout

Puisque $\mathbb{Z}$ est un anneau intègre, nous avons la définition 1.215 de pgcd et de ppcm.

Proposition 1.225 (PPCM et PGCD)**.**
Soient $p, q \in \mathbb{Z}^*$.

(1) Le pgcd de p et q est le plus grand diviseur commun de p et q.

(2) Le ppcm de p et q est leur plus petit multiple commun.

Démonstration. Démontrons le premier point. Notons δ le pgcd de p et q. Si d est un diviseur commun de p et q, alors d divise δ. Dans $\mathbb{Z}$, $d \mid \delta$ implique $d \leqslant \delta$ (proposition 1.107). □

Lemme 1.226.
Soient $p, q \in \mathbb{Z}^*$. Les entiers $\operatorname{ppcm}(p,q)$ *et* $\operatorname{pgcd}(p,q)$ *fournissent les isomorphismes de groupes suivants :*

$$p\mathbb{Z} \cap q\mathbb{Z} = \operatorname{ppcm}(p,q)\mathbb{Z} \tag{1.339a}$$

$$p\mathbb{Z} + q\mathbb{Z} = \operatorname{pgcd}(p,q)\mathbb{Z}. \tag{1.339b}$$

Définition 1.227.
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 plus grand que 1, au triplet $(2, 4, 7)$.

Théorème 1.228 (Théorème de Bézout [76][46], thème 3)**.**
Deux entiers non nuls $a, b \in \mathbb{Z}^*$ sont premiers entre eux si et seulement si il existe $u, v \in \mathbb{Z}$ tels que

$$au + bv = 1 \tag{1.340}$$

76. Il y a une super application ici : https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/mauvais_prix.pdf.

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

$$E = \{au + bv \text{ tel que } u, v \in \mathbb{Z}\} \cap \mathbb{N}^*. \tag{1.341}$$

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

$$m = au_1 + bv_1. \tag{1.342}$$

Par le théorème de division euclidienne [77] (avec a et m), il existe des entiers uniques q et r tels que

$$a = mq + r \tag{1.343}$$

avec $0 \leqslant r < m$. En remplaçant m par sa valeur (1.342), $a = (au_1 + bv_1)q + r$ et

$$r = a(1 - u_1 q) - bv_1 q, \tag{1.344}$$

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. Puisque m divise à la fois a et b nous avons $m = 1$. □

Une généralisation de Bézout 1.228 à plus de 2 variables.

Proposition 1.229.
Si $\{a_i\}_{i=1,\dots,N}$ sont des entiers tels que $\text{pgcd}(a_1, \dots, a_N) = 1$, alors il existe des entiers $\{u_i\}_{i=1,\dots,N}$ tels que

$$\sum_i a_i u_i = 1. \tag{1.345}$$

Corolaire 1.230.
Soient p et q deux entiers premiers entre eux. Alors

$$p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}; \tag{1.346}$$

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 $kp + lq = 1$. Alors, $(xk)p + (xl)q = x$. □

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 chaine de Markov peuvent être atteints à tout moment (assez grand), voir la définition 38.47 et ce qui suit.

Proposition 1.231.
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 positifs.

77. Théorème 1.224.

(i) **Une décomposition pour x** On applique le théorème 1.224 de division euclidienne à x et $a+b$: il existe des entiers p_x, r_x, uniques, tels que

$$\begin{cases} x = (p_x - 1)(a+b) + r_x & (1.347a) \\ 0 \leqslant r_x < a+b. & (1.347b) \end{cases}$$

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

$$ua + vb = p_x(a+b) - x \tag{1.348}$$

par le corolaire 1.230. Ainsi, x peut s'écrire

$$x = (p_x - u)a + (p_x - v)b. \tag{1.349}$$

(ii) **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 corolaire 1.230, nous considérons les nombres u_i et v_i définis par

$$u_i a + v_i b = i \tag{1.350}$$

pour $i = 1, \dots, a+b$. Nous posons $u^* = \max\{u_i\}$, $v^* = \max\{v_i\}$, et $p^* = \max\{u^*, v^*\}$. Nous posons alors $N = p^*(a+b)$, et considérons $x > N$.

(iii) **Nouvelle décomposition pour x** Nous voulons écrire

$$x = (p_x - u_k)a + (p_x - v_k)b \tag{1.351}$$

pour un certain k. Cela demande $u_k a + v_k b = ua + vb = p_x(a+b) - x$ par l'équation (1.348). Vu que $p_x(a+b) - x > 0$, les nombres u_k et v_k existent : il suffit de prendre $k = p_x(a+b) - x$.

(iv) **Conclusion** Avec tous ces choix, nous avons d'abord $x > p^*(a+b)$ et donc

$$x = (p_x - 1)(a+b) + r_x > p^*(a+b), \tag{1.352}$$

ce qui donne

$$(p_x - 1)(a+b) > p^*(a+b) - r_x > (p-1)(a+b). \tag{1.353}$$

ou encore $p_x > p^*$. Nous avons finalement

$$p_x \geqslant p^* \geqslant u^* \geqslant u_k \tag{1.354}$$

et

$$p_x \geqslant p^* \geqslant v^* \geqslant v_k. \tag{1.355}$$

De ce fait, la décomposition (1.351) est celle que nous voulions.

□

1.232.
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 1.233.
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

$$1004 = 72 \cdot 7 + 100 \cdot 5. \tag{1.356}$$

Ensuite $4 = 25 - 21 = -3 \cdot 7 + 5 \cdot 5$. Au final,

$$1000 = 75 \cdot 7 + 95 \cdot 5. \tag{1.357}$$

△

1.9.3 Sous-groupes de $(\mathbb{Z}, +)$

Proposition 1.234 (liste des sous groupes de $\mathbb{Z}$)**.**
À propos de sous-groupes de $\mathbb{Z}$.

(1) Une partie H du groupe $(\mathbb{Z}, +)$ est un sous-groupe si et seulement si il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.

(2) Si H est une sous-groupe de $(\mathbb{Z}, +)$, il existe un unique 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 1.224, 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 $x = nq \in n\mathbb{Z}$.

En ce qui concerne l'unicité, supposons que $n\mathbb{Z} = m\mathbb{Z}$. Le nombre n divise m (parce que $m \in m\mathbb{Z} \subset n\mathbb{Z}$) et le nombre m divise n parce que $n \in m\mathbb{Z}$. Par conséquent $n = m$. □

1.9.4 Résultats supplémentaires sur l'anneau des entiers

Corolaire 1.235.
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 1.234, 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[78] : si $a \in n\mathbb{Z}$ et si $k \in \mathbb{Z}$ alors $ak \in n\mathbb{Z}$. □

Proposition 1.236.
Soient $n \geqslant 2$ un entier et $\phi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ la surjection canonique. Nous noterons $\overline{a} = \phi(a)$. Alors l'ensemble des inversibles de $\mathbb{Z}/n\mathbb{Z}$ est donné par

$$U(\mathbb{Z}/n\mathbb{Z}) = \phi(P_n) = \{\overline{x} \text{ tel que } 0 \leqslant x \leqslant n \text{ tel que } \operatorname{pgcd}(x, n) = 1\}. \tag{1.358}$$

où P_n est l'ensemble $P_n = \{x \in \{0, \dots, n-1\} \text{ tel que } \operatorname{pgcd}(x, n) = 1\}$.

De plus,

$$\operatorname{Card}\big(U(\mathbb{Z}/n\mathbb{Z})\big) = \phi(n). \tag{1.359}$$

Démonstration. Soit $0 \leqslant x \leqslant n$ tel que $\operatorname{pgcd}(x, n) = 1$. Il existe donc[79] $u, v \in \mathbb{Z}$ tels que $ux + vn = 1$. En passant aux classes,

$$\overline{ux} = \overline{1}, \tag{1.360}$$

donc $\overline{u}$ est l'inverse de $\overline{x}$. Cela prouve que $\phi(P_n) \subset U(\mathbb{Z}/n\mathbb{Z})$.

Nous prouvons maintenant l'inclusion inverse. Soient $\overline{x}$ et $\overline{y}$ inverses l'un de l'autre : $\overline{xy} = \overline{1}$. Il existe donc $q \in \mathbb{Z}$ tel que $xy - qn = 1$, ce qui prouve[80] que $\operatorname{pgcd}(x, n) = 1$. □

1.10 Corps

1.10.1 Définitions, morphismes

Définition 1.237 ([47])**.**
*Un **corps** est un anneau*[81] *$(A, +, \times)$ dans lequel tout élément non nul est inversible pour l'opération $\times$ (pour l'opération $+$, tous les éléments sont inversibles parce que $(A, +)$ est un groupe).*

78. Définition 1.213.
79. Théorème de Bézout 1.228
80. À nouveau avec le Théorème de Bézout.
81. Définition 1.38.

1.238.
Dans le Frido, nous ne parlons que de corps commutatifs ; nous ne le répéterons pas toujours.

1.239.
Pour savoir ce qu'est un « *ring* » ou « *field* » en anglais, voir -2.3.

La proposition suivante donne une caractérisation d'un corps, en disant un tout petit peu plus que la définition 1.237.

Proposition 1.240.
L'anneau A est un corps si et seulement si $U(A) = A^*$.

Démonstration. En deux parties.

(i) **Sens direct** Nous supposons que A est un corps. D'une part tous les éléments non nuls sont inversibles, c'est-à-dire $A^* \subset U(A)$.
Pour l'inclusion inverse, nous montrons qu'une élément inversible ne peut pas être nul. Cela n'est autre que le lemme 1.209 couplé à la proposition 1.210 : $a \cdot 0 = 0 \neq 1$ pour tout a.

(ii) **Sens inverse** Si $U(A) = A^*$, nous avons immédiatement que tous les éléments non nuls sont inversibles et donc que A est un corps.

□

Lemme 1.241.
Si $\mathbb{K}$ est un corps et si $a \in \mathbb{K}$ vérifie $a^2 = 1$, alors $a = \pm 1$.

Lemme 1.242.
Un corps non nul est un anneau intègre[82].

Démonstration. Soit un produit nul $ab = 0$. Si $a \neq 0$, alors il est inversible et nous multiplions $ab = 0$ par a^{-1}. Nous trouvons $b = 0$ parce que $0a^{-1} = 0$. □

Conséquence : dans un corps nous avons toujours la règle du produit nul, et l'élément nul n'est jamais inversible.

Définition 1.243 (Morphisme de corps)**.**
*Un corps étant un anneau sans plus de structure, un **morphisme de corps** n'est qu'un morphisme des anneaux*[83].

Le lemme suivant montre que définir un morphisme de corps comme étant simplement un morphisme des anneaux est une bonne idée.

Lemme 1.244.
Si $\varphi\colon \mathbb{K} \to \mathbb{K}'$ est un morphisme de corps, alors

(1) pour tout $a \in \mathbb{K}$ nous avons $\varphi(a^{-1}) = \varphi(a)^{-1}$;

(2) le morphisme φ est injectif.

Démonstration. Vu que $\varphi(1) = 1$, nous avons aussi

$$1 = \varphi(aa^{-1}) = \varphi(a)\varphi(a^{-1}). \tag{1.361}$$

Donc, par unicité de l'inverse[84], $\varphi(a^{-1}) = \varphi(a)^{-1}$.

Pour l'injectivité nous supposons $\varphi(a) = \varphi(b)$. Étant donné que $\mathbb{K}'$ est un corps, nous pouvons multiplier par $\varphi(b)^{-1}$:

$$\varphi(a)\varphi(b)^{-1} = 1. \tag{1.362}$$

En utilisant le premier point nous avons $1 = \varphi(a)\varphi(b^{-1})$, puis le morphisme d'anneaux : $1 = \varphi(ab^{-1})$, et encore le morphisme d'anneaux nous permet de déduire $ab^{-1} = 1$ et donc $a = b$. □

82. Définition 1.221.
83. Définition 1.39.
84. Lemme 1.157 (2).

1.11 Anneau intègre

La définition d'un anneau intègre est la définition 1.221.

Lemme 1.245.
Un corps[85] *est un anneau intègre.*

Démonstration. En effet, soient un corps $\mathbb{K}$ et deux éléments $x, y \in \mathbb{K}$ tels que $xy = 0$. Si y est inversible, alors nous pouvons multiplier par y^{-1} pour trouver $x = 0$. Cela prouve que $\mathbb{K}$ est un anneau intègre. □

Exemple 1.246.
L'ensemble $\mathbb{Z}$ avec les opérations usuelles est un anneau intègre[86]. △

Exemple 1.247.
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. △

Nous verrons au théorème 3.126 que l'anneau A est intègre si et seulement si $A[X]$ est intègre.

Corolaire 1.248.
L'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre si et seulement si n est premier.

Démonstration. Supposons que n soit premier. La proposition 1.236 donne les inversibles de $\mathbb{Z}/n\mathbb{Z}$ par

$$U(\mathbb{Z}/n\mathbb{Z}) = \{\overline{x} \text{ tel que } 0 \leqslant x \leqslant n \text{ tel que } \operatorname{pgcd}(x, n) = 1\}. \tag{1.363}$$

Mais comme n est premier, $\operatorname{pgcd}(x, n) = 1$ pour tout x, et donc tous les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont inversibles. Donc $\mathbb{Z}/n\mathbb{Z}$ est intègre.

Si n n'est pas premier, alors $n = pq$ avec $1 < p \leqslant q < n$. Alors

$$[p]_n[q]_n = [pq]_n = [0]_n. \tag{1.364}$$

Donc lorsque n n'est pas premier, l'anneau $\mathbb{Z}/n\mathbb{Z}$ possède des diviseurs de zéro et n'est alors pas intègre. □

1.11.1 Élément premier

Définition 1.249 ([48])**.**
Soit un anneau commutatif A. Un élément $p \in A$ est ***premier*** *si il est*

(1) non nul,

(2) non inversible,

(3) si p divise un produit ab, alors il divise soit a soit b (ou le deux).

Le lemme suivant est souvent pris pour la définition d'un nombre premier lorsqu'on parle de $\mathbb{N}$ ou $\mathbb{Z}$.

Lemme 1.250 ([49, 1])**.**
Dans $\mathbb{N}$, un nombre est premier si et seulement si il admet exactement deux diviseurs entiers distincts.

Démonstration. En deux parties.

(i) $\Rightarrow$ Soit un élément premier $p \in \mathbb{N}$. Il y a trois possibilités : $p = 0$, $p = 1$ et $p > 1$.
Le nombre $p = 0$ n'est pas premier parce qu'il est nul. Le nombre $p = 1$ n'est pas premier parce qu'il est inversible. Donc nous savons que si p est premier, alors $p > 1$.

85. Définition 1.237.
86. Anneau intègre, définition 1.221.

Un élément $p > 1$ dans $\mathbb{N}$ a toujours au moins deux diviseurs distincts : 1 et p. Soit un diviseur k de p. Il existe $l \in \mathbb{N}$ tel que $p = kl$. Vu que p est premier et divise le produit kl, il divise k ou l. Disons que p divise k. De cette façon p divise k et k divise p.

Il existe donc $n \in \mathbb{N}$ tel que $k = np$. En y substituant $p = kl$, on trouve $k = np = nkl$. En simplifiant par k, il vient

$$1 = nl, \tag{1.365}$$

ce qui prouve que $n = l = 1$ et donc que $k = p$ et donc que p n'a pas d'autres diviseurs que 1 et p.

(ii) $\Leftarrow$ Nous supposons que $p \in \mathbb{N}$ ait exactement deux diviseurs entiers distincts. Nous vérifions que p vérifie les trois conditions de la définition 1.249.

(1) $p \neq 0$ parce que 0 a nettement plus que deux diviseurs distincts.

(2) $p \neq 1$ parce que 1 a exactement un diviseur. Donc p n'est pas inversible dans $\mathbb{N}$.

(3) Soit p admettant exactement deux diviseurs distincts. Soit p divisant le produit ab' pour certains a et b' dans $\mathbb{N}$. Nous supposons que p ne divise pas a, et nous allons prouver que p divise b' en supposant d'abord que p ne divise pas b'.

(i) **Un ensemble** Pour cela nous posons

$$E = \{x \in \mathbb{N} \text{ tel que } p \mid ax, p \nmid x\}. \tag{1.366}$$

Nous posons $b = \min(E)$. Nous avons pour hypothèse que E est non vide ; en particulier $0 < b$.

(ii) $b < p$ On vérifie que si $p + k \in E$ alors $k \in E$. Donc b ne peut pas être plus grand que p. Vu que p lui-même n'est pas dans E, nous avons $b < p$.

(iii) **Division euclidienne** Nous effectuons la division euclidienne du théorème 1.224 :

$$p = mb + r. \tag{1.367}$$

En multipliant par a, $ar = ap - mab$. Vu que ab est un multiple de p $ap - mab$ est un multiple de p. En particulier ar est divisible en p.

(iv) **La contradiction** Nous avons donc $r \in E$, alors que $r < b$. Impossible.

□

Proposition 1.251 ([50]).
Dans un anneau intègre[87] *tout élément premier est irréductible*[88].

Démonstration. Soit p, un élément premier dans un anneau intègre A.

(i) **p n'est pas inversible** Cela fait partie de la définition d'un élément premier.

(ii) **p n'est pas un produit d'inversibles** Soient $a, b \in A$ tels que $p = ab$. Par le point (3) de la définition 1.249, p divise soit a soit b. Supposons que p divise a. Alors il existe $x \in A$ tel que $a = px$. En remettant dans $p = ab$ nous avons :

$$p = pxb. \tag{1.368}$$

Mais l'anneau est intègre et permet donc des simplifications par tout élément non nul. La relation 1.368 donne donc

$$1 = xb, \tag{1.369}$$

ce qui signifie que b est inversible.

Un travail similaire montre que a est inversible si p divise b.

87. Si pas intègre, voir l'exemple 1.253.
88. Toutes les définitions dans le thème 21.

□

Exemple 1.252.
Si nous avons l'égalité $7 = ab$ dans $\mathbb{Z}$, alors soit a soit b vaut 1. Mettons $a = 1$. Dans ce cas, $b = 7$ et n'est donc pas inversible. △

Sur un anneau non intègre, la notion d'élément premier n'est pas aussi intéressante que sur un anneau intègre. Par exemple la proposition 1.251 devient fausse.

Exemple 1.253.
Soit l'anneau $\mathbb{Z}^2$. L'élément $(1,0)$ est premier mais pas irréductible.

(i) **$(1,0)$ est premier** L'élément $(1,0)$ est non nul ; ça c'est pas cher. Pour qu'il soit inversible, il faudrait $(1,0)(x,y) = (1,1)$. Entre autres, $0 \times y = 1$, ce qui est impossible. Donc il n'est pas inversible.

Supposons que $(1,0)$ divise le produit $(a,b)(c,d) = (ac,bd)$. Alors il existe (x,y) tel que $(1,0)(x,y) = (ac,bd)$. Cela signifie que $x = ac$ et $0 \times y = bd$. En particulier, soit $b = 0$ soit $d = 0$. Si $b = 0$, nous avons $(a,b) = (a,0)$ et effectivement, $(1,0)$ le divise.

(ii) **$(1,0)$ n'est pas irréductible** Nous avons $(1,0) = (1,0)(1,0)$. Donc l'élément $(1,0)$ est le produit de deux éléments non inversibles.

△

1.12 Symbole de sommation

1.12.1 Somme à valeurs dans un groupe commutatif

Si S est un ensemble fini, nous savons de la proposition 1.120 qu'il existe un unique $N \in \mathbb{N}$ pour lequel il existe une bijection $\varphi\colon \{0,\dots,N\} \to S$. Cette bijection n'est à priori pas unique.

Lemme-Définition 1.254 ([1]).
Soient un groupe commutatif $(G,+)$ ainsi qu'un ensemble fini I contenant n éléments. Soit une application $f\colon I \to G$. Si $\sigma_1,\sigma_2\colon \{1,\dots,n\} \to I$ sont deux bijections, alors[89]

$$\sum_{i=1}^{n} f\big(\sigma_1(i)\big) = \sum_{i=1}^{n} f\big(\sigma_2(i)\big). \tag{1.370}$$

La valeur commune est notée

$$\sum_{i\in I} f(i) \tag{1.371}$$

Démonstration. Nous commençons par considérer une transposition σ (qui permute k et l avec $k < l$). Nous avons

$$\sum_{i=1}^{n} f(i) = \sum_{i=1}^{k-1} f(i) + f(k) + \sum_{i=k+1}^{l-1} f(i) + f(l) + \sum_{i=l+1}^{n} f(i) \tag{1.372a}$$

$$= \sum_{i=1}^{k-1} f(i) + f(l) + \sum_{i=k+1}^{l-1} f(i) + f(k) + \sum_{i=l+1}^{n} f(i) \tag{1.372b}$$

$$= \sum_{i=1}^{n} f\big(\sigma(i)\big). \tag{1.372c}$$

Pour cela nous avons utilisé le fait que G est commutatif pour permuter $f(l) \in G$ et $f(k) \in G$ avec $\sum_{i=k+1}^{l-1} f(i) \in G$.

89. Pour rappel, le symbole $\sum_{i=1}^{n}$ est défini par 1.81.

Une permutation quelconque est un produit de telles transpositions (proposition 1.200). Donc pour toute permutation σ nous avons

$$\sum_{i=1}^{n} f\big(\sigma(i)\big) = \sum_{i=1}^{n} f(i). \tag{1.373}$$

□

La définition 1.254 donne lieu à un certain nombre de remarques.

(1) Elle donne la somme sur un ensemble fini. Un problème avec les ensembles infinis (outre la convergence) est l'ordre de sommation. Si vous voulez sommer sur $\mathbb{Z}$, dans quel ordre le faire ?

(2) Pour aller plus loin, et sommer sur des ensembles infinis, rendez-vous dans le thème 51.

Proposition 1.255.
Soient un groupe commutatif $(G,+)$, un ensemble fini I, une application $f\colon I\to G$ et une bijection $\sigma\colon I\to I$. Alors

$$\sum_{i\in I} f(i) = \sum_{i\in I} f\big(\sigma(i)\big). \tag{1.374}$$

Si nous avons une application $L\colon S\to S$, nous notons

$$\sum_{s\in S} f\big(L(s)\big) = \sum_{s\in S}(f\circ L)(s). \tag{1.375}$$

Cette façon d'écrire donne une interprétation pour la notation $\sum_{g\in G} f(hg)$ qui arrive dans la proposition 1.257. Il s'agit de considérer l'application L_h du lemme 1.160, de considérer [90]

$$\sum_{g\in G} f(hg) = \sum_{g\in G}(f\circ L_h)(g) \tag{1.376}$$

et de faire tourner la définition 1.254. La même chose tient pour définir $\sum_{g\in G}(gh)$ à l'aide de R_h.

Lemme 1.256.
Soit un ensemble A fini pouvant être écrit comme une union disjointe $A=\bigcup_{k=1}^{n} A_k$; nous supposons que les A_i sont non vides. Soient un groupe commutatif $(G,+)$ et une application $f\colon A\to G$. Alors

$$\sum_{a\in A} f(a) = \sum_{k=1}^{n}\sum_{a\in A_k} f(a). \tag{1.377}$$

Démonstration. Le lemme 1.113 nous indique que les parties A_k sont des ensembles finis. Nous notons

(1) $N_0=0$, et $N_k=\mathrm{Card}(A_k)$,

(2) $S_k=\sum_{i=1}^{k} N_k$.

(3) $\varphi_k\colon\{1,\dots,N_k\}\to A_k$, une bijection (l'existence est dans la proposition 1.120).

Nous avons $\mathrm{Card}(A)=S_n$ par le lemme 1.122(4). Nous définissons une belle bijection comme il faut :

$$\begin{aligned}\alpha\colon\{1,\dots,S_n\}&\to A\\ i&\mapsto \varphi_{k+1}(i-S_k)\end{aligned} \tag{1.378}$$

pour $i\in\,]S_k,S_{k+1}]$.

(i) <u>**α est bien définie**</u> Puisque $i>S_k$ et $i\leqslant S_{k+1}$ nous avons $i-S_k\in\{1,\dots,N_{k+1}\}$, et donc φ_{k+1} s'applique bien à $i-S_k$.

90. Le fait que L_h soit une bijection n'a pas d'importance ici.

(ii) **α est injective** Supposons que $\alpha(i) = \alpha(j)$. Si $i \in]S_k, S_{k+1}]$ et $j \in]S_l, S_{l+1}]$, alors $\alpha(i) = \varphi_{k+1}(i - S_k) \in A_{k+1}$ et $\alpha(j) = \varphi_{l+1}(j - S_l) \in A_{l+1}$. Vu que les A_i sont disjoints, nous avons $k = l$, et donc

$$\varphi_{k+1}(u - S_k) = \varphi_{k+1}(j - S_k). \tag{1.379}$$

Étant donné que φ_{k+1} est injective, nous avons $i - S_k = j - S_k$, ce qui montre que $i = j$.

(iii) **α est surjective** Soit $a \in A$. Il existe k tel que $a \in A_k$. Nous avons donc un $s \in \{1, \dots, N_k\}$ tel que $a = \varphi_k(s)$. En posant $i = s + S_k$, nous avons bien $a = \alpha(s + S_k)$ parce que $s + S_k \in]S_{k-1}, S_k]$.

Vu que α est une bijection, nous avons l'égalité

$$\sum_{a \in A} f(a) = \sum_{i=1}^{S_n} (f \circ \alpha)(i). \tag{1.380}$$

Nous avons encore besoin d'introduire une bijection. Nous posons

$$\begin{aligned} \beta_k \colon]S_{k-1}, S_k] &\to A_k \\ i &\mapsto \varphi_k(i - S_{k-1}). \end{aligned} \tag{1.381}$$

C'est une bijection parce que φ_k en est une, et que $i \mapsto i - S_{k-1}$ est une bijection de $]S_{k-1}, S_k]$.

Nous pouvons maintenant terminer :

$$\begin{aligned} \sum_{a \in A} f(a) &= \sum_{i=1}^{S_n} (f \circ \alpha)(i) && \text{(1.382a)} \\ &= \sum_{k=1}^{n} \left(\sum_{i=S_{k-1}-1}^{S_k} (f \circ \alpha)(i) \right) && \text{(1.382b)} \\ &= \sum_{k=1}^{n} \left(\sum_{i \in]S_{k-1}, S_k]} f\big(\varphi_k(i - S_{k-1})\big) \right) && \text{(1.382c)} \\ &= \sum_{k=1}^{n} \left(\sum_{i \in]S_{k-1}, S_k]} f\big(\beta_k(i)\big) \right) && \text{(1.382d)} \\ &= \sum_{i=1}^{n} \left(\sum_{a \in A_k} f(a) \right). && \text{(1.382e)} \end{aligned}$$

Justifications :

— Pour (1.382b). Associativité de la somme.

□

Proposition 1.257 ([1]).
Soient un groupe fini G et une fonction $f \colon G \to A$ où A est un anneau commutatif. Alors

$$\sum_{g \in G} f(g) = \sum_{g \in G} f(gh) = \sum_{g \in G} f(hg) \tag{1.383}$$

pour tout $h \in G$.

Démonstration. Nous avons une bijection $\varphi \colon \{0, \dots, N\} \to G$ garantie par la proposition 1.120. Sa définition est

$$\sum_{g \in G} f(g) = \sum_{i=0}^{N} f\big(\varphi(i)\big). \tag{1.384}$$

Par ailleurs, le lemme 1.160 donne une bijection $L_h\colon G \to G$ et permet de considérer la composée

$$\begin{aligned} \varphi'\colon \{0,\dots,N\} &\to G \\ \varphi' &= L_h \circ \varphi. \end{aligned} \tag{1.385}$$

La proposition 1.254 nous permet d'utiliser la bijection φ' au lieu de φ pour exprimer la somme $\sum_{g\in G}$. Ensuite un jeu de notation utilisant (1.376) donne

$$\begin{aligned} \sum_{g\in G} f(g) &= \sum_{i=0}^{N} f\big(\varphi(i)\big) = \sum_{i=0}^{N} f\big(\varphi'(i)\big) = \sum_{i=0}^{N} (f\circ L_h\circ\varphi)(i) \\ &= \sum_{i=0}^{N} (f\circ L_h)\big(\varphi(i)\big) = \sum_{g\in G} (f\circ L_h)(g) = \sum_{g\in G} f(hg). \end{aligned} \tag{1.386}$$

En ce qui concerne $\sum_{g\in G} f(gh)$, c'est la même chose, en utilisant R_h au lieu de L_h. □

Lemme 1.258.
Soit un groupe totalement ordonné[91] $(A,+,\leqslant)$. *Soient deux suites* (a_i) *et* (b_i) *dans* G *telles que* $a_i \leqslant b_i$ *pour tout* i. *Alors pour tout* n *nous avons*

$$\sum_{i=0}^{n} a_i \leqslant \sum_{i=0}^{n} b_i. \tag{1.387}$$

Tout cela nous permet de définir une somme sympathique et bien connue.

Lemme 1.259.
Soit $n\in\mathbb{N}$. *Nous avons*

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}. \tag{1.388}$$

Démonstration. La preuve est pratiquement immédiate par récurrence. Nous allons donner une preuve plus « constructive », qui formalise l'idée classique d'écrire la somme à l'endroit et à l'envers.

Nous notons S la somme $\sum_{k=0}^{n} k$. Le lemme 1.254 dit que si les $\sigma_i\colon \{0,\dots,n\}\to\{0,\dots,n\}$ sont des bijections, alors $\sum_{k=0}^{n} f\big(\sigma_1(k)\big) = \sum_{k=0}^{n} f\big(\sigma_2(k)\big)$. Nous sommes intéressé au cas $f(i)=i$.

En prenant $\sigma_1(k)=k$ et $\sigma_2(k)=n-k$, nous avons

$$S = \sum_{k=0}^{n} k = \sum_{k=0}^{n} (n-k). \tag{1.389}$$

Donc

$$2S = \sum_{k=0}^{n} \big(k+(n-k)\big) = \sum_{k=0}^{n} n = n\sum_{k=0}^{n} 1 = n(n+1). \tag{1.390}$$

En divisant par deux, nous obtenons le résultat annoncé. □

1.13 Symbole de produit

1.260.
Si $(G,\ \cdot\ )$ est un groupe et si $H\subset G$, nous notons le produit des éléments de H par

$$\prod_{g\in H} g = \sum_{g\in H} g \tag{1.391}$$

où à droite, c'est la somme déjà définie. La différence entre $\prod$ et $\sum$ est que nous utilisons $\prod$ pour les groupes notés « multiplicativement » comme $(G,\ \cdot\ )$ alors que nous utilisons $\sum$ lorsque le groupe est noté « additivement » comme $(G,+)$.

91. Définition 1.161.

Dans le cas d'un anneau $(A, +, \cdot)$, la distinction est importante pour savoir quelle opération est sous-entendue.

La définition 1.81(1) signifie qu'une somme vide vaut zéro : $\sum_{x\in\emptyset} x = 0$. Vu que zéro est la façon usuelle de noter le neutre pour une opération notée « + », lorsque l'opération est notée $\cdot$ nous avons

$$\prod_{x\in\emptyset} x = 1 \tag{1.392}$$

parce que 1 est la façon usuelle de noter le neutre d'une opération notée « $\cdot$ ».

Notez que (1.392) n'est pas une nouvelle définition ou une nouvelle convention. C'est seulement l'égalité $\sum_{x\in\emptyset x} x = 0$, avec des notations adaptées à un groupe dont l'opération est notée multiplicativement.

Proposition 1.261.
Si E est un ensemble fini et si G est un groupe commutatif, alors pour toute fonction $f\colon E \to G$ et pour toute permutation[92] *σ de E,*

$$\prod_{i\in E} f(i) = \prod_{i\in E} f\big(\sigma(i)\big) \tag{1.393}$$

Démonstration. C'est exactement la proposition 1.254, sauf qu'ici la loi de groupe est notée multiplicativement au lieu d'additivement. □

1.13.1 Sous-groupe engendré

Définition 1.262 (Sous-groupe engendré)**.**
Soit A une partie du groupe G. Le sous-groupe ***engendré*** *par A est l'intersection de tous les sous-groupes de G contenant A. Nous notons ce groupe $\mathrm{gr}_G(A)$.*

Lorsque A est fini (disons $A = \{a_1, \dots, a_n\}$), on note aussi le sous-groupe engendré $\langle a_1, \dots, a_n\rangle$.

1.263.
Un sous-groupe engendré n'est jamais vide parce qu'il contient toujours au moins le neutre (parce que c'est un sous-groupe). Si G est un groupe, le sous-groupe $\mathrm{gr}_G(\emptyset)$ lui-même contient e[93].

1.264.
Dans de nombreux cas, le groupe « ambiant » G est entendu par le contexte et nous noterons $\mathrm{gr}(A)$ au lieu de $\mathrm{gr}_G(A)$.

Si par exemple A est la matrice $\begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$, le groupe $\mathrm{gr}(A)$ est à comprendre dans $\mathrm{GL}(2, \mathbb{R})$. Il faudrait être fou pour avoir en tête un autre groupe que $\mathrm{GL}(2, \mathbb{R})$ sans le préciser.

D'ailleurs, connaissez-vous un groupe contenant la matrice A et n'étant pas un sous-groupe de $\mathrm{GL}(2, \mathbb{C})$?

Lemme 1.265.
Si G est un groupe et A une partie de G, alors $\mathrm{gr}(A)$ est un sous-groupe de G.

Le sous-groupe engendré par A est le plus petit (pour l'inclusion) groupe de G contenant A. Plus formellement, nous avons le résultat suivant :

Lemme 1.266.
Tout sous-groupe de G contenant A contient $\mathrm{gr}(A)$.

Démonstration. Si H est un sous-groupe de G contenant A, alors $\mathrm{gr}(A)$ est l'intersection de H avec tous les autres sous-groupes de G contenant A. Il contient donc $\mathrm{gr}(A)$. □

92. Une permutation est une bijection, définition 1.179.
93. Demandez-vous si il est possible que $\mathrm{gr}(\emptyset)$ contienne d'autres éléments que e.

Lemme 1.267 ([51]).
Si A est une partie du groupe G, alors le sous-groupe $\mathrm{gr}(A)$ engendré[94] *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).*

C'est-à-dire que tout élément de $\mathrm{gr}(A)$ peut être écrit sous la forme[95]

$$\prod_{i=1}^{n} g_i^{a_i} \tag{1.394}$$

où $a_i \in \mathbb{Z}$ et $g\colon \mathbb{N} \to A$ n'est pas spécialement injective : il peut arriver que $g_i = g_j$.

Démonstration. Puisqu'un produit vide est égal à l'identité[96], le lemme est vrai (un peu trivialement) dans le cas où $A = \emptyset$. À partir de maintenant, nous supposons que A est non vide.

Nous nommons $\mathrm{gr}(A)$ le groupe engendré par A et H, l'ensemble

$$H = \{g_1 \dots g_n \text{ tel que } g_i \in A \cup A^{-1}\}. \tag{1.395}$$

Nous commençons par prouver que H est un groupe.

— Puisque A est non vide, nous considérons $a \in A$. Dans ce cas, $e = aa^{-1} \in H$. Donc $e \in H$.

— L'inverse de $g_1 \dots g_n$ est $g_n^{-1} \dots g_1^{-1}$ qui est également dans H.

— Le produit de $g_1 \dots g_n$ par $h_1 \dots h_n$, tous éléments de H, est également dans H[97].

Comme H est un groupe contenant A, nous avons $\mathrm{gr}(A) \subset H$ parce que $\mathrm{gr}(A)$ est une intersection dont un des éléments est H.

Par ailleurs tout groupe contenant A doit contenir les inverses et les produits finis, donc $H \subset \mathrm{gr}(A)$.

Au final, $H = \mathrm{gr}(A)$, ce qu'il fallait. □

Lemme 1.268.
Soit un groupe G et un sous-groupe $H = \mathrm{gr}(h_1, \dots, h_n)$. Si $\alpha \in G$, alors

$$\alpha H \alpha^{-1} = \mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1}). \tag{1.396}$$

Démonstration. Il s'agit d'une conséquence du lemme 1.267. Un élément de $\mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1})$ est un produit d'éléments de G de la forme $\alpha h_i \alpha^{-1}$ ou $(\alpha h_j \alpha^{-1})^{-1} = \alpha h_j^{-1} \alpha^{-1}$. Or nous avons

$$\alpha h_i \alpha^{-1} \alpha h_j \alpha^{-1} = \alpha h_i h_j \alpha^{-1} \in \alpha H \alpha^{-1}. \tag{1.397}$$

Donc

$$\mathrm{gr}(\alpha h_1 \alpha^{-1}, \dots, \alpha h_n \alpha^{-1}) \subset \alpha H \alpha^{-1}. \tag{1.398}$$

L'inclusion dans l'autre sens est du même tonneau. □

Définition 1.269 (Partie génératrice, groupe monogène).
Soient un groupe G, et une partie $A \subset G$. Si $\mathrm{gr}(A) = G$, alors nous disons que A est une ***partie génératrice*** *du groupe G.*

Un groupe est ***monogène*** *si il a une partie génératrice réduite à un seul élément.*

Définition 1.270 (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****.*

94. Définition 1.262.
95. Les a_i négatifs correspondent aux inverses. Notons que si $g \in A$, il n'y a pas de garanties que g^{-1} soit également dans A.
96. Voir 1.260.
97. Et c'est ici qu'on se rend compte que la décomposition n'est probablement que rarement unique.

1.271.
La différence entre un groupe monogène et un groupe cyclique est qu'un groupe cyclique est fini. Dans un groupe cyclique, à force d'itérer le générateur, nous finissons par tourner en rond – d'où le nom.

Exemple 1.272.
Soit le groupe $\big(\mathbb{Z}/10\mathbb{Z},+\big)$. L'élément $[2]_{10}$ n'est pas générateur parce que ses puissances[98] sont

$$\mathrm{gr}([2]_{10}) = \{[2]_{10},[4]_{10},[6]_{10},[8]_{10},[0]_{10}\}. \tag{1.399}$$

Par contre l'élément $[3]_{10}$ est générateur : ses puissances sont dans l'ordre

$$[3]_{10},[6]_{10},[9]_{10},[2]_{10},[5]_{10},[8]_{10},[1]_{10},[4]_{10},[7]_{10},[0]_{10}. \tag{1.400}$$

△

Un exemple presque identique, mais un peu masqué sera l'exemple 18.149.

Lemme 1.273 ([1])**.**
Si $n = dr$, alors

(1) $r\mathbb{Z}/n\mathbb{Z}$ est un groupe.

(2) Le groupe $r\mathbb{Z}/n\mathbb{Z}$ est cyclique[99]*.*

(3) $\mathrm{Card}(r\mathbb{Z}/n\mathbb{Z}) = d = n/r$.

1.14 Module sur un anneau

Définition 1.274 (module sur un anneau[52])**.**
*Soit un anneau A. Un **module à gauche** sur A est la donnée d'un triplet $(M,+,\ \cdot\ )$ où*

(1) $+$ est une loi de composition interne à M, c'est-à-dire $+\colon M\times M\to M$,

(2) $\cdot$ est une loi de composition externe, c'est-à-dire $\cdot\ : A\times M\to M$

telles que

(1) $(M,+)$ est un groupe[100]*.*

(2) $a\cdot(x+y) = a\cdot x + a\cdot y$,

(3) $(a+b)\cdot x = a\cdot x + b\cdot x$,

(4) $(ab)\cdot x = a\cdot(b\cdot x)$

(5) $1\cdot x = x$.

pour tout $a,b\in A$ et $x,y\in M$.

*Si M et N sont des A-modules, un **morphisme** de M vers N est une application $f\colon M\to N$ qui*

(1) est un morphisme de groupes entre $(M,+)$ et $(N,+)$

(2) vérifie $f(a\cdot x) = a\cdot f(x)$ pour tout $a\in A$, $x\in M$.

L'ensemble des morphismes entre M et N est noté $\mathrm{Hom}_A(M,N)$. Si B st une sous-anneau de A, nous parlons de $\mathrm{Hom}_B(M,N)$ pour parler des morphismes de groupes qui ne vérifient $f(a\cdot x) = a\cdot f(x)$ que pour $a\in B$.

Proposition 1.275.
Si M est un module sur un anneau, alors $(M,+)$ est un groupe commutatif.

98. Attention aux notations ; en général on écrit la loi de groupe de façon multiplicative et on parle des puissances d'un élément, mais ici on écrit la loi de groupe additivement, donc les « puissances » sont en réalité les multiples.
99. Définition 1.270.
100. Nous verrons dans la proposition 1.275 qu'il est forcément commutatif.

Démonstration. Il suffit de calculer $(1+1)\cdot(x+y)$ de deux façons différentes :

$$(1+1)\cdot(x+y) = 1\cdot(x+y) + 1\cdot(x+y) = x+y+x+y \tag{1.401}$$

d'une part et

$$(1+1)\cdot(x+y) = (1+1)\cdot x + (1+1)\cdot y = x+x+y+y, \tag{1.402}$$

d'autre part. En égalant les deux expressions, il vient

$$x+y+x+y = x+x+y+y, \tag{1.403}$$

qui se simplifie (nous sommes dans un groupe) en $y+x = x+y$. □

Définition 1.276.
*Un **espace vectoriel** est un module*[101] *sur un corps commutatif*[102].

Définition 1.277 ([53])**.**
*Soient un A-module M et un ensemble I. Une famille $\{m_i\}_{i\in I}$ est **libre** si les m_i sont **linéairement indépendants**, c'est-à-dire si pour tout choix d'une partie finie J dans I et d'éléments $(a_j)_{j\in J}$ dans A, si nous avons*

$$\sum_{j\in J} a_j m_j = 0, \tag{1.404}$$

alors $a_j = 0$ pour tout j.

Définition 1.278 ([54])**.**
*Soit S, une partie du A-module M. Le **sous-module engendré** par S est l'ensemble des éléments de M qui sont des combinaisons linéaires finies d'éléments de S, c'est-à-dire de sommes de la forme*

$$\sum_{t\in T} a_t t \tag{1.405}$$

où T est fini dans S et $a_t \in A$.

1.14.1 Module produit

Lemme-Définition 1.279 ([53])**.**
*Soient un anneau A et un ensemble I. Le A-**module produit** A^I est l'ensemble des applications $I \to A$.*

En termes de notations, nous écrivons ceci :

$$A^I = \{(a_i)_{i\in I}, a_i \in A\}. \tag{1.406}$$

L'ensemble A^I devient un module par les définitions, pour $x, y \in A^I$ et $a \in A$:

$$ax = (ax_i)_{i\in I} \tag{1.407a}$$

$$x+y = (x_i+y_i)_{i\in I}. \tag{1.407b}$$

En d'autres termes, $A^I = \mathrm{Fun}(I, A)$.

Lemme 1.280.
Pour chaque $i \in I$ nous considérons l'élément $e_i \in A^I$ donné par

$$\begin{aligned} e_i \colon I &\to A \\ j &\mapsto \begin{cases} 1 & \text{si } j = i \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{1.408}$$

La famille $\{e_i\}_{i\in I}$ est libre[103] *dans A^I.*

101. Définition 1.274.
102. La condition de commutativité n'est pas indispensable, mais comme nous ne parlerons que de corps commutatifs...
103. Définition 1.277.

Démonstration. Soient J fini dans I ainsi que des éléments $a_j \in A$ ($j \in J$). Nous supposons que[104] $\sum_{j\in J} a_j e_j = 0$. Calculons un peu :

$$\sum_{j\in J} a_j e_j = \sum_{j\in J} (a_j \delta_{ji})_{i\in I} = \left(\sum_{j\in J} a_j \delta_{ji} \right)_{i\in I}. \tag{1.409}$$

Pour que le tout soit nul dans A^I, il faut que

$$\sum_{j\in J} a_j \delta_{ji} \tag{1.410}$$

soit nul pour tout $i \in I$. Si nous fixons $i \in I$, la somme sur j possède un seul terme non annulé par δ_{ji}, et c'est le terme $j = i$. Nous avons donc $a_i = 0$. □

Définition 1.281.
Nous notons $A^{(I)}$ le sous-module de A^I engendré[105] *par les e_i.*

Lemme 1.282 ([1])**.**
L'ensemble $A^{(I)}$ est l'ensemble des applications $I \to A$ de support fini.

Démonstration. En deux sens.

(i) **Si $x \in A^{(I)}$** Pour rappel, la définition 1.279 nous dit que x est une application $I \to A$. Vu que x est dans le sous-module engendré par les e_i, il existe une partie finie $J \subset I$ telle que

$$x = \sum_{j\in J} x_j e_j. \tag{1.411}$$

Pour $i \in I$ nous avons

$$x(i) = \sum_{j\in J} x_j \delta_{ij} = \begin{cases} x_i & \text{si } i \in J \\ 0 & \text{sinon.} \end{cases} \tag{1.412}$$

Donc le support de x est dans J qui est fini. Vu que toute partie d'un ensemble fini est fini (lemme 1.113), le support de x est fini.

(ii) **Si x est de support fini** Supposons que le support de $x\colon I \to A$ soit la partie finie $J \subset I$. En notant $x_j = x(j)$ pour tout $j \in J$, nous avons

$$x = \sum_{j\in J} x_j e_j. \tag{1.413}$$

□

Théorème 1.283 (Propriété universelle de $A^{(I)}$[53])**.**
Soient un anneau A ainsi qu'un A-module P. Pour $\phi \in \mathrm{Hom}_A(A^{(I)}, P)$, nous considérons

$$\begin{aligned} \phi|_I \colon I &\to P \\ i &\mapsto \phi(e_i). \end{aligned} \tag{1.414}$$

(1) L'application

$$\begin{aligned} f\colon \mathrm{Hom}_A(A^{(I)}, P) &\to \mathrm{Fun}(I, P) \\ \phi &\mapsto \phi|_I \end{aligned} \tag{1.415}$$

est une bijection.

(2) L'application inverse est $g\colon \mathrm{Fun}(I, P) \to \mathrm{Hom}_A(A^{(I)}, P)$ donnée par

$$g(\psi)\big(\sum_{j\in J} a_j e_j\big) = \sum_{j\in J} a_j \psi(j) \tag{1.416}$$

pour tout J fini dans I et choix de $a_j \in A$.

104. Pour rappel, les sommes finies sont définies par 1.254.
105. Définition 1.278.

Démonstration. Nous allons montrer que $g\big(f(\phi)\big) = \phi$ et que $f\big(g(\psi)\big) = \psi$ pour tout $\phi \in \mathrm{Hom}_A(A^{(I)}, P)$ et pour tout $\psi \in \mathrm{Fun}(I, P)$.

Dans un premier sens nous avons :

$$g\big(f(\phi)\big)\big(\sum_j a_j e_j\big) = \sum_j a_j f(\phi)(j) \tag{1.417a}$$
$$= \sum_j a_j \phi(e_j) \tag{1.417b}$$
$$= \phi(\sum_j a_j e_j). \tag{1.417c}$$

Justifications :

— Pour (1.417b), nous avons utilisé le fait que $f(\phi)(i) = \phi|_I(i) = \phi(e_i)$.

— Pour (1.417c), nous utilisons le fait que ϕ est un morphisme de modules.

Et pour l'autre sens,

$$f\big(g(\psi)\big)(i) = g(\psi)(e_i) = \psi(i). \tag{1.418}$$

Vérifions que cela est suffisant pour que f soit une bijection.

(i) **Surjectif** Soit $\psi \in \mathrm{Fun}(I, P)$. Nous avons $f\big(g(\psi)\big) = \psi$, ce qui prouve que ψ est dans l'image de f.

(ii) **Injectif** Supposons que $f(\phi_1) = f(\phi_2)$. Alors en appliquant g des deux côtés, il vient $\phi_1 = \phi_2$.

□

1.14.2 Sous-module

Soient 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

$$\begin{aligned} \mu_x \colon A^{(I)} &\to M \\ (a_i)_{i\in I} &\mapsto \sum_{i\in I} a_i x_i. \end{aligned} \tag{1.419}$$

Ici $A^{(I)}$ désigne l'ensemble de toutes les applications $I \to A$ de support fini (définition 1.281).

Définition 1.284.
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 1.285.
*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 1.286.
Un anneau A est lui-même un A-module et ses sous-modules sont les idéaux. △

Définition 1.287.
*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 1.288.
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

$$p_i(x_1 + \dots + x_n) = x_i \tag{1.420}$$

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

$$M = \bigoplus_{i=1}^n p_i(M). \tag{1.421}$$

Définition 1.289.
*Un module est **simple** ou **irréductible** si il n'a pas d'autres sous-modules que $\{0\}$ et lui-même. Un module est **indécomposable** si 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 1.290.
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 1.291 (Algèbre[55])**.**
Si $\mathbb{K}$ est un corps commutatif[106]*, une $\mathbb{K}$-**algèbre** A est un espace vectoriel*[107] *muni d'une opération bilinéaire $\times\colon A \times A \to A$, 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 $f\colon A \to B$ 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 1.292 ([1])**.**
Soient 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 1.293.
*L'**algèbre engendrée** par X est l'intersection de toutes les sous-algèbres de A contenant X (qui est une algèbre par le lemme 1.292).*

106. Définition 1.237
107. Définition 1.276.

1.15 Caractéristique d'un anneau

Lemme-Définition 1.294.
Soit l'application

$$\begin{aligned} \mu\colon \mathbb{Z} &\to A \\ n &\mapsto n \cdot 1_A \end{aligned} \tag{1.422}$$

où $n \cdot 1_A$ *signifie* $\sum_{k=1}^{n} 1_A$.

(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 de $\mathbb{Q}$ est zéro parce qu'aucun multiple de l'unité n'est nul.
À propos de diagonalisation en caractéristique 2, voir l'exemple 9.208.

Lemme 1.295.
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 1.296.
Soit un anneau A *de caractéristique* p.

(1) Si $p > 0$, *alors nous avons l'isomorphisme d'anneaux*

$$\mathbb{Z}1_A \simeq \mathbb{Z}/p\mathbb{Z}. \tag{1.423}$$

(2) Si $p = 0$, *alors nous avons l'isomorphisme d'anneaux*

$$\mathbb{Z}1_A \simeq \mathbb{Z} \tag{1.424}$$

Démonstration. Pour (1), l'isomorphisme est donné par l'application $n1_A \mapsto \phi(n)$ si ϕ est la projection canonique $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$.

Pour (2), la preuve est encore à faire. Écrivez-moi. □

Proposition 1.297.
La caractéristique d'un anneau fini divise son cardinal.

Démonstration. Si A est un anneau, le groupe $\mathbb{Z}$ agit sur A par

$$n \cdot a = a + n1_A. \tag{1.425}$$

Chaque orbite de cette action est de la forme

$$\mathcal{O}_a = \{a + n1_A \text{ tel que } n = 0, \dots, p-1\} \tag{1.426}$$

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 1.298 ([56])**.**
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 1.299.
Soit A un anneau commutatif unitaire de caractéristique p. L'application

$$\begin{aligned} \mathrm{Frob}_A \colon A &\to A \\ x &\mapsto x^p \end{aligned} \tag{1.427}$$

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 1.300.
Soit à factoriser $X^p - 1$ dans $\mathbb{F}_p$. Grâce au morphisme de Frobenius, nous avons immédiatement

$$X^p - 1 = (X - 1)^p. \tag{1.428}$$

△

Lemme 1.301.
La caractéristique[108] *d'un anneau intègre est zéro ou un élément premier*[109].

Démonstration. Si A est intègre, alors $\mathbb{Z}1_A$ est a fortiori intègre. Notons p la caractéristique de A. Si $p = 0$, la preuve est finie ; supposons donc que $p \neq 0$. Alors, l'anneau $\mathbb{Z}/p\mathbb{Z}$ est isomorphe à $\mathbb{Z}1_A$, et est donc intègre. Or, la proposition 1.248 dit que $\mathbb{Z}/p\mathbb{Z}$ est intègre si et seulement si p est premier, ce qui conclut la preuve. □

Exemple 1.302.
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 :

$$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = 0. \tag{1.429}$$

△

1.15.1 Caractéristique deux

Beaucoup de résultats demandent une caractéristique différente de deux. Qu'a donc de particulier la caractéristique deux ?

Si $\mathbb{K}$ est un corps de caractéristique 2, alors l'égalité $x = -x$ n'implique pas $x = 0$, puisque $2x = 0$ 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.

1.16 Polynômes

1.16.1 Polynômes d'une variable

Et voilà la définition que tout le monde attendait ; la définition des anneaux de polynômes. Pour ne pas taper trop fort du premier coup, nous commençons par les polynômes d'une seule variable[110].

L'ensemble des polynômes sur A sera simplement $A^{(\mathbb{N})}$ (notation 1.281). Puisque $\mathbb{N}$ est un ensemble bien particulier possédant plein de structure, nous allons pouvoir installer sur $A^{(\mathbb{N})}$ une structure non seulement de A-module (ça c'est déjà fait), mais en plus d'anneau, ainsi qu'une évaluation.

108. Définition 1.294.
109. Définition 1.249.
110. Pour les polynômes à plusieurs variables, voir la définition 3.53.

Définition 1.303.
L'ensemble des ***polynômes*** *en une indéterminée sur l'anneau A est l'anneau*

$$\mathcal{P}(A) = A^{(\mathbb{N})} \tag{1.430}$$

défini en 1.281.

1.304.
En ce qui concerne la notation $A[X]$, voir 1.16.2. Pour $\mathbb{K}(X)$ lorsque $\mathbb{K}$ est un corps, voir 6.78.

Proposition-Définition 1.305 ([57]).
Soit P non nul dans $\mathcal{P}(A)$. Nous notons a_n la valeur[111] *de P en $n \in \mathbb{N}$: $P = (a_n)_{n\in\mathbb{N}}$.*

(1) L'ensemble $\{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est fini dans $\mathbb{N}$.

(2) Cet ensemble possède un minimum et un maximum.

Le ***degré*** *de P est*

$$\deg(P) = \max\{n \in \mathbb{N} \text{ tel que } a_n \neq 0\}, \tag{1.431}$$

et la ***valuation*** *de P est*

$$\mathrm{val}(P) = \min\{n \text{ tel que } a_n \neq 0\}. \tag{1.432}$$

Dans le cas du polynôme nul, l'ensemble $\{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est vide, et les définitions ne s'appliquent pas. Nous convenons que

$$\mathrm{val}(0) = +\infty \tag{1.433a}$$

$$\deg(0) = -\infty. \tag{1.433b}$$

Démonstration. Le fait que P soit non nul implique que $A = \{n \in \mathbb{N}$ tel que $a_n \neq 0\}$ est non vide. De plus cet ensemble est fini parce que $P \in A^{(\mathbb{N})}$. Toute partie finie non vide de $\mathbb{N}$ étant majorée et minorée (lemme 1.61), le lemme 1.62 définit correctement le minimum et le maximum de A. □

Vu que $A^{(\mathbb{N})}$ est engendré par les e_i, tout polynôme sur A s'écrit $P = \sum_{i=1}^{n} a_i e_i$.

Définition 1.306.
Nous ajoutons deux structures à $A^{(\mathbb{N})}$.

L'évaluation *Si $\alpha \in A$ et si $P \in A^{(\mathbb{N})}$, nous définissons $P(\alpha)$ par*

$$P(\alpha) = (\sum_{i=0}^{n} a_i e_i)(\alpha) = \sum_{i=0}^{n} a_i \alpha^i, \tag{1.434}$$

étant entendu que $\alpha^0 = 1$ dans A.

Cette définition s'étend immédiatement au cas où B est un anneau qui étend A. Dans ce cas nous pouvons définir $P(b)$ pour tout $P \in A^{(\mathbb{N})}$ et $b \in B$ avec la même formule (1.434).

Le produit *C'est ici que la structure particulière de $\mathbb{N}$ est utilisée. Nous définissons le produit $A^{\mathbb{N}} \times A^{(\mathbb{N})} \to A^{(\mathbb{N})}$ de la façon suivante. Si $(P_k)_{k\in\mathbb{N}}$ est la suite (presque partout nulle) d'éléments de A qui définit P et si $(Q_k)_{k\in\mathbb{N}}$ est celle de Q, nous notons*

$$(PQ)_n = \sum_{k=0}^{n} P_k Q_{n-k}, \tag{1.435}$$

et donc $PQ = \sum_i (PQ)_i e_i$. Plus explicitement,

$$(\sum_{i=0}^{n} a_i e_i)(\sum_{j=0}^{m} b_j e_j) = \sum_{k=0}^{\infty} \Big(\sum_{\substack{(i,j)\in\mathbb{N}^2 \\ i+j=k}} a_i b_j \Big) e_k. \tag{1.436}$$

Notons qu'à droite, la somme sur k est une somme finie.

111. Ici il y a une énorme subtilité de terminologie. Formellement, P est une application $\mathbb{N} \to A$. Cela n'a rien à voir avec le fait que P puisse être évalué sur A avec des formule du type $P(x) = \sum_n a_n x^n$. D'ailleurs nous n'avons pas encore vu cette évaluation.

Proposition 1.307.
Soit un anneau A. À propos de structure sur $A^{(\mathbb{N})}$.

(1) Avec le produit, l'ensemble $A^{(\mathbb{N})}$ devient un anneau.

(2) L'application

$$\begin{aligned} g\colon A^{(\mathbb{N})} &\to A \\ P &\mapsto P(\alpha) \end{aligned} \tag{1.437}$$

est un morphisme d'anneaux[112]*. En particulier, $(PQ)(\alpha) = P(\alpha)Q(\alpha)$.*

Démonstration. En plusieurs points

(i) **Anneau** L'identité pour le produit dans $A^{(\mathbb{N})}$ est le polynôme donné par $a_0 = 1$ et $a_i = 0$ pour $i \neq 0$. Cela se vérifie en utilisant directement la définition (1.436). La distributivité aussi[113].

(ii) **Le morphisme** Nous notons P_k les éléments de la suite définissant P et Q_k ceux de Q. Alors nous avons

$$(P+Q)(\alpha) = \sum_k (P_k + Q_k)\alpha^k = \sum_k P_k\alpha^k + \sum_k Q_k\alpha^k = P(\alpha) + Q(\alpha). \tag{1.438}$$

Vous aurez noté que la première égalité était la définition (1.407b). De même,

$$P(\alpha)Q(\alpha) = \big(\sum_n P_n\alpha^n\big)\big(\sum_k Q_k\alpha^k\big) = \sum_k Q_k\big(\sum_n P_n\alpha^n\big)\alpha^k = \sum_k\sum_n Q_kP_n\alpha^{n+k} \tag{1.439a}$$

$$= \sum_m \big(\sum_{l=0}^m P_lQ_{m-l}\big)\alpha^m = \sum_m (PQ)_m\alpha^m = (PQ)(\alpha). \tag{1.439b}$$

□

Lemme 1.308.
Si A est commutatif, alors $A^{(\mathbb{N})}$ est commutatif.

Démonstration. Soient $P, Q \in A^{(\mathbb{N})}$; pour rappel, le produit est donné par la définition 1.435. L'application

$$\begin{aligned} \varphi\colon \{0,\dots,n\} &\to \{0,\dots,n\} \\ k &\mapsto n-k \end{aligned} \tag{1.440}$$

est une bijection. Voici maintenant le calcul :

$$(PQ)_n = \sum_{k=0}^n P_kQ_{n-k} \tag{1.441a}$$

$$= \sum_{k=0}^n P_{\varphi(k)}Q_{n-\varphi(k)} \tag{1.441b}$$

$$= \sum_{k=0}^n P_{n-k}Q_k \tag{1.441c}$$

$$= \sum_{k=0}^n Q_lP_{n-k} \tag{1.441d}$$

$$= (QP)_n. \tag{1.441e}$$

Justifications

— Pour (1.441b). Lemme 1.254 et le fait que φ soit une bijection.

— Pour (1.441d). Commutativité de A.

□

112. Définition 1.39.

113. Je n'ai pas fait les calculs, écrivez-moi pour me dire si ça va facilement.

1.16.2 La notation $A[X]$

Si A est un anneau, nous avons déjà défini les polynômes en une indéterminée sur A comme étant le module $A^{(\mathbb{N})}$ qui est devenu un anneau par la proposition 1.307.

Le polynôme donné par la suite $(a_n)_{n\in\mathbb{N}}$ est souvent notée

$$\sum_k a_k X^k. \tag{1.442}$$

Par exemple avec $a = (4, 2, 8)$ nous avons $a = 8X^2 + 2X + 4$. Nous utiliserons souvent cette notation, qui est très pratique parce qu'elle s'adapte bien aux règles de multiplication et d'addition, en particulier la distributivité.

Il y a (au moins) deux façons de comprendre ce que signifie réellement « X » dans cette notation.

1.16.2.1 Première façon (qui botte en touche)

La première est de dire qu'il n'a pas de significations, et que X^2 est un simple abus de notations pour écrire $(0, 0, 1, 0, \cdots)$. Avec cette façon de voir, nous notons l'anneau des polynômes sur A par « $A[X]$ » où le X n'a pas d'autres raisons d'être que d'avertir le lecteur que nous réservons la lettre « X » pour utiliser la notation pratique des polynômes.

1.16.2.2 Seconde façon (la bonne)

1.309.
La seconde façon de voir le « X » est de nous rappeler que $A^{(\mathbb{N})}$ a une base en tant que module : les e_k dont nous avons parlé plus haut. Nous posons $X = e_1$, et nous prenons la convention $X^0 = 1$. Alors nous avons $e_k = X^k$ et nous notons $A[X]$ l'anneau $A^{(\mathbb{N})}$ exprimé avec X.

Dans les deux cas, il n'est pas vraiment légitime d'écrire des égalités comme « $P(X) = X^2 + 2X - 3$ », et encore moins de dire « Le polynôme P, *évalué* en X vaut $X^2 + 2X - 3$ » : il est plus correct d'écrire « $P = X^2 + 2X - 3$ ».

Le lemme suivant montre que ces notations tombent vraiment à point. La véritable difficulté de l'énoncé est de comprendre qu'il n'est pas trivial.

Nous avons vu dans la définition 1.306 que si B est un anneau qui étend A, et si $P \in A[X]$, alors nous avons une définition de $P(b)$ pour tout $b \in B$. Nous appliquons cela à $B = A[X]$, qui est un anneau qui étend A. Autrement dit, si P et Q sont des polynômes, ça a un sens d'écrire $P(Q)$ et le résultat sera un élément de $A[X]$.

Dans le cas particulier $Q = X$, nous avons une chouette formule.

Lemme 1.310.
Nous avons

$$P(X) = P \tag{1.443}$$

pour tout $P \in A[X]$.

Démonstration. Si $P = (a_k)_{k\in\mathbb{N}}$ alors par définition $P(\alpha) = \sum_k a_k \alpha^k$ dès que α est dans un anneau B qui étend A. Nous considérons le cas particulier $B = A[X]$ et $\alpha = X$, c'est-à-dire $Q = (0, 1, 0, \ldots)$, l'élément $P(X)$ de $A[X]$ vaut

$$\sum_k a_k X^k, \tag{1.444}$$

qui est exactement P lui-même. □

Mais il faut bien comprendre que si P est le polynôme $(-3, 2, 1, 0, \ldots)$, noté $X^2 + 2X - 3$, écrire $P(X) = X^2 + 2X - 3$ est une pirouette de notations que rien ne justifie par rapport à simplement écrire $P = X^2 + 2X - 3$.

1.16.3 Action du groupe symétrique

Définition 1.311 (Thème 8).
*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.*

Par souci de notations, nous notons $\mathcal{P}_n(A)$ l'anneau des polynômes de n variables sur A. La propriété universelle de $\mathcal{P}_n(A) = A^{(\mathbb{N}^n)}$ du théorème 1.283 nous donne une application

$$g\colon \operatorname{Fun}\big(\mathbb{N}^n, \mathcal{P}_n(A)\big) \to \operatorname{Hom}_A\big(\mathcal{P}_n(A), \mathcal{P}_n(A)\big) \tag{1.445}$$

Avec cela nous pouvons énoncer et démontrer le lemme qui donne l'action de S_n [114] sur $\mathcal{P}_n(A)$.

Lemme 1.312 ([58]).
Pour $\sigma \in S_n$ nous définissons

$$\begin{aligned} \phi_\sigma\colon \mathbb{N}^n &\to \mathcal{P}_n(A) \\ m &\mapsto e_{\sigma(m)}. \end{aligned} \tag{1.446}$$

Alors l'application

$$\begin{aligned} \rho\colon S_n &\to \operatorname{Hom}_A\big(\mathcal{P}_n(A), \mathcal{P}_n(A)\big) \\ \sigma &\mapsto g(\phi\sigma) \end{aligned} \tag{1.447}$$

est une action [115].

Démonstration. Nous commençons par donner une expression à notre ρ. Un élément de $\mathcal{P}_n(A)$ est de la forme $\sum_{m\in\mathbb{N}^n} a_m e_m$, et nous avons [116]

$$\rho(\sigma)\big(\sum_{m\in\mathbb{N}^n} a_m e_m\big) = \sum_m a_m \phi_\sigma(m) = \sum_m a_m e_{\sigma(m)}. \tag{1.448}$$

Nous avons tout de suite $\rho(\mathrm{Id}) = \mathrm{Id}$.

En ce qui concerne la composition, nous avons d'une part

$$\rho(\sigma_1\sigma_2)\big(\sum_m a_m e_m\big) = g(\phi_{\sigma_1\sigma_2})\big(\sum_m a_m e_m\big) = \sum_m a_m e_{\sigma_1\sigma_2(m)}, \tag{1.449}$$

et d'autre part,

$$\begin{aligned} \rho(\sigma_1)\rho(\sigma_2)\big(\sum_m a_m e_m\big) &= \rho(\sigma_1)\big(\sum_m a_m e_{\sigma_2(m)}\big) && \text{(1.450a)} \\ &= \rho(\sigma_1)\big(\sum_m a_{\sigma_2^{-1}(m)} e_m\big) && \text{(1.450b)} \\ &= \sum_m a_{\sigma_2^{-1}(m)} e_{\sigma_1(m)} && \text{(1.450c)} \\ &= \sum_m a_m e_{\sigma_1\sigma_2(m)} && \text{(1.450d)} \end{aligned}$$

La proposition 1.255 est utilisée pour (1.450b) et pour (1.450d). □

114. Définition du groupe symétrique S_n en 1.179.
115. Définition 1.311.
116. La somme est définie par 1.254, et ça va être important. Ah oui, en réalité partout, les sommes sont finies parce que les a_m ($m \in \mathbb{N}^n$) sont presque tous nuls. Il faudrait écrire sur la somme sur $\{m \in \mathbb{N}^2$ tel que $a_m \neq 0\}$, mais vous vous imaginez la complication dans la notation.

1.16.4 Corps des fractions

Définition 1.313 ([59]).
Soit un anneau commutatif et intègre[117] *A. Nous posons $E = A \times A\backslash\{0\}$, et nous définissons les deux opérations suivantes sur E :*

(1) $(a,b)+(c,d) = (ad+cb, bd)$;

(2) $(a,b)(c,d) = (ac, bd)$.

Et aussi la relation d'équivalence $(a,b) \sim (c,d)$ si et seulement si $ad = bc$.

*Le **corps des fractions** de A est le quotient*

$$\mathrm{Frac}(A) = \big(A \times A\backslash\{0\}\big)/\sim . \tag{1.451}$$

Nous notons a/b la classe de (a,b).

Lorsque A est un anneau de polynômes[118]*, alors les éléments de* $\mathrm{Frac}(A)$ *sont des **fractions rationnelles**.*

Le fait que A soit intègre est important pour être certain que $bd \neq 0$ sous l'hypothèse que $b, d \neq 0$.

La proposition suivante montre encore que le corps des fractions est le plus petit corps que l'on puisse imaginer à partir d'un anneau.

Proposition 1.314 ([11, 1]).
Soit un anneau commutatif A. Tout corps commutatif contenant un sous-anneau isomorphe[119] *à A contient un sous-corps isomorphe à* $\mathrm{Frac}(A)$.

Démonstration. Soit un corps $\mathbb{K}$ contenant un sous-anneau A' isomorphe à A. Nous notons $\sigma\colon A' \to A$ un isomorphisme d'anneaux entre A' et A.

(i) **Une partie bien choisie** Nous considérons la partie suivante de $\mathbb{K}$:

$$S = \{ab^{-1} \text{ tel que } a, b \in A'\}. \tag{1.452}$$

(ii) **S est un corps** Deux éléments arbitraires de S sont ab^{-1} et xy^{-1}. Nous devons prouver plusieurs choses.

(i) **Neutres** En prenant $a = b = 1$ nous avons $ab^{-1} = 1 \in S$. En prenant $a = 0$ et $b = 1$ nous avons $ab^{-1} = 0 \in S$.

(ii) **Somme** Il faut remarquer que $ab^{-1} + xy^{-1} = (ay+xb)(by)^{-1}$. En effet,

$$(ay+xb)(xb)^{-1} = (ay+xb)y^{-1}b^{-1} \tag{1.453a}$$
$$= ayy^{-1}b^{-1} + xby^{-1}b^{-1} \tag{1.453b}$$
$$= ab^{-1} + xy^{-1} \tag{1.453c}$$

Justifications :

— Pour (1.453b). Distributivité.

— Pour (1.453c). Commutativité dans A.

(iii) **Produit** Il s'agit du même genre de calculs en utilisant les mêmes propriétés. Nous avons

$$(ab^{-1})(xy^{-1}) = (ax)(by)^{-1}. \tag{1.454}$$

(iii) **Ce qui va être notre isomorphisme** Ensuite nous montrons que l'application

$$\begin{aligned} \varphi\colon S &\to \mathrm{Frac}(A) \\ ab^{-1} &\mapsto \sigma(a)/\sigma(b) \end{aligned} \tag{1.455}$$

est bien définie et est un isomorphisme de corps.

117. Définition 1.221.
118. Définition 1.303.
119. Morphisme d'anneaux, définition 1.39.

(iv) **Bien définie** Si $ab^{-1} = xy^{-1}$ alors $ay = xb$. Puisque σ est un isomorphisme nous avons aussi $\sigma(a)\sigma(y) = \sigma(x)\sigma(b)$ et donc $\sigma(a)/\sigma(b) = \sigma(x)/\sigma(y)$ par définition des classes de $\mathrm{Frac}(A)$.

(v) **Morphisme** Deux éléments arbitraires de S sont ab^{-1} et xy^{-1}. Calculons un peu :

$$\begin{aligned}
\varphi\big((ab^{-1})(xy^{-1})\big) &= \varphi(axy^{-1}b^{-1}) && (1.456a)\\
&= \varphi\big((ax)(by)^{-1}\big) && (1.456b)\\
&= \sigma(ax)/\sigma(by) && (1.456c)\\
&= \big(\sigma(a)/\sigma(b)\big)\big(\sigma(x)/\sigma(y)\big) && (1.456d)\\
&= \varphi(ab^{-1})\varphi(xy^{-1}). && (1.456e)
\end{aligned}$$

Justifications :

— Pour (1.456a). Commutativité dans A.

— Pour (1.456b). Associativité dans A.

— Pour (1.456d). Définition 1.313(2) de la multiplication de fractions.

(vi) **Surjectif** Tout élément de $\mathrm{Frac}(A)$ est de la forme a'/b' avec $a', b' \in A$, et donc de la forme $\sigma(a)/\sigma(b)$ avec $a, b \in A'$. Un tel élément est l'image par φ de $ab^{-1} \in S$.

(vii) **Injectif** Si $\varphi(ab^{-1}) = \varphi(xy^{-1})$ alors $\sigma(a)/\sigma(b) = \sigma(x)/\sigma(y)$, et par définition des classes nous avons $\sigma(a)\sigma(y) = \sigma(b)\sigma(x)$. De là nous avons $\sigma(ay) = \sigma(bx)$ et donc $ay = bx$ (parce que σ est un isomorphisme). Nous en déduisons que $ab^{-1} = xy^{-1}$.

□

1.315.
Soit un anneau A et son anneau des polynômes $\mathcal{P}(A)$. Si $\alpha \in A$, nous avons la définition 1.306 qui donne l'évaluation $P(\alpha)$.

Si par contre P et Q sont des polynômes sur A, nous n'avons pas encore défini ce que serait l'évaluation de la fraction rationnelle P/Q en α. Nous comblons à présent ce manque.

Définition 1.316 (Évaluation d'une fraction rationnelle)**.**
Soit un corps $\mathbb{K}$ contenant l'anneau A. Si $R = P/Q \in \mathrm{Frac}(A)$ et si $\alpha \in \mathbb{K}$ nous définissons[120]

$$R(\alpha) = (P/Q)(\alpha) = P(\alpha)Q^{-1}(\alpha). \tag{1.457}$$

Dans cette formule, les polynômes, l'inverse et le produit sont calculés dans $\mathbb{K}$ et non dans A.

Théorème-Définition 1.317.
Soit A un anneau commutatif intègre.

(1) Il existe un couple $(\mathbb{K}, \epsilon)$ où $\mathbb{K}$ est un corps commutatif et $\epsilon\colon A \to \mathbb{K}$ est un morphisme injectif d'anneaux tels que pour tout $\lambda \in \mathbb{K}$, il existe $(a, b) \in A \times A^*$ tels que

$$\lambda = \epsilon(a)\big(\epsilon(b)\big)^{-1} \tag{1.458}$$

*(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 A est le **corps des fractions** de A, et sera noté $\mathrm{Frac}(A)$.*

(3) Nous posons

$$\begin{aligned}
\sigma\colon A \times A^* &\to \mathbb{K}\\
(a, b) &\mapsto \epsilon(a)\big(\epsilon(b)\big)^{-1}.
\end{aligned} \tag{1.459}$$

Nous avons

$$\sigma(xa, xb) = \sigma(a, b) \tag{1.460}$$

pour tout $a, b, x \in A$.

120. Les fractions rationnelles, définition 1.313.

1.16.5 Corps totalement ordonné

Définition 1.318.
Ordre et choses reliées dans un corps.

(1) *Un corps $\mathbb{K}$ est* ***totalement ordonné*** *si il existe une relation d'ordre total*[121] *tel que*
   (1a) $x \leqslant y$ *implique* $x + z \leqslant y + z$ *pour tout* $x, y, z \in \mathbb{K}$
   (1b) $x \geqslant 0$ *et* $y \geqslant 0$ *implique* $xy \geqslant 0$.

(2) *Si $\mathbb{K}$ est un corps totalement ordonné, nous y définissons la valeur absolue par*

$$|x| = \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{si } x \leqslant 0. \end{cases} \tag{1.461}$$

(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*** *si 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 totalement ordonné est* ***complet*** *si toute suite de Cauchy y est convergente.*

(6) *Si $a, \epsilon \in \mathbb{K}$ avec $\epsilon > 0$ alors nous définissons la* ***boule ouverte*** *de centre a et de rayon ϵ par*

$$B(a, \epsilon) = \{x \in \mathbb{K} \text{ tel que } |a - x| < \epsilon\}, \tag{1.462}$$

et la ***boule fermée*** *par*

$$\overline{B(a, \epsilon)} = \{x \in \mathbb{K} \text{ tel que } |a - x| \leqslant \epsilon\}. \tag{1.463}$$

Lemme 1.319.
Une suite (x_k) converge vers q si et seulement si pour tout $\epsilon > 0$, il existe $N > 0$ tel que $x_k \in B(q, \epsilon)$ pour tout $k \geqslant N$.

Démonstration. Il s'agit de mettre côte à côte les points (4) et (6) de la définition 1.318. □

1.320.
Ces boules prendront une nouvelle force avec le super-théorème 7.98.

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 1.321 (Règle des signes[60])**.**
Soit un corps totalement ordonné $\mathbb{K}$ ainsi que $x, y \in \mathbb{K}$. Nous avons :

(1) *Si $x \leqslant 0$ et $y \leqslant 0$ alors $xy \geqslant 0$.*
(2) *Si $x \leqslant 0$ et $y \geqslant 0$ alors $xy \leqslant 0$.*
(3) *Si $x \geqslant 0$ et $y \leqslant 0$ alors $xy \leqslant 0$.*
(4) $0 \leqslant 1$.
(5) *Si $x \geqslant 0$ alors $x^{-1} \geqslant 0$.*

Lemme 1.322 (Propriétés de la valeur absolue)**.**
Soit $\mathbb{K}$ un corps totalement ordonné. Si $x, y \in \mathbb{K}$ alors[122]

(1) *Si $x \geqslant 0$ alors $-x \leqslant 0$.*

121. Définition 1.11.
122. La « valeur absolue » est définie en (1.318)(2).

(2) Si $x \leqslant 0$ alors $-x \geqslant 0$.

(3) $|x| \geqslant 0$

(4) $|x| = 0$ si et seulement si $x = 0$

(5) $|-x| = |x|$.

(6) $|x+y| \leqslant |x| + |y|$.

(7) $|xy| = |x||y|$

Démonstration. Point par point

(i) **(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 $-x \leqslant 0$.

(ii) **(2)** Nous partons de $x \leqslant 0$ et nous ajoutons $-x$ des deux côtés.

(iii) **(3)** Si $x \geqslant 0$ alors c'est vrai. Sinon, $x \leqslant 0$ et $|x| = -x \geqslant 0$ par le point (1).

(iv) **(4)** 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 $\pm x \neq 0$.

(v) **(5)** Il faut décomposer en deux cas selon que $x \geqslant 0$ et $x \leqslant 0$. Supposons $x \geqslant 0$. Alors d'une part $|x| = x$. D'autre part $-x \leqslant 0$ par le point (1), de telle sorte que

$$|-x| = -(-x) = x. \tag{1.464}$$

Nous avons donc $|x| = |-x| = x$.

Le même raisonnement tient avec $x \leqslant 0$.

(vi) **(6)** Nous supposons que $x \leqslant y$ et nous distinguons divers cas suivant la positivité de x et y.

(1) Si $x, y \geqslant 0$. Dans ce cas, $x + y \geqslant y \geqslant 0$, donc $|x+y| = x + y = |x| + |y|$.

(2) Si $x, y \leqslant 0$. Dans ce cas, $x + y \leqslant 0$ et nous avons $|x+y| = -x - y = |x| + |y|$.

(3) 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

$$x \leqslant 0 \tag{1.465a}$$

$$x + y \leqslant y \leqslant y + |x| = |x| + |y| \tag{1.465b}$$

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.

(vii) **Pour (7)** Il suffit de prendre les 4 cas suivant les signes de x et y, et d'utiliser les règles de signes du lemme 1.321 dans la définition 1.461.

□

Remarque 1.323.
La partie (6) 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

$$|x - y| = |(x - z) + (z - y)| \leqslant |x - z| + |z - y|. \tag{1.466}$$

Lemme 1.324 (À propos de boules)**.**
Soient un corps totalement ordonné $\mathbb{K}$ et des éléments $x, y \in \mathbb{K}$. Soit aussi $\epsilon > 0$ dans $\mathbb{K}$. Nous avons :

(1) $y \in B(x, \epsilon)$ si et seulement si $x - \epsilon < y < x + \epsilon$.

(2) Si $y \in \overline{B(x, \epsilon)}$ alors $y \in B(x, \epsilon')$ pour tout $\epsilon' > \epsilon$.

Démonstration. Pour rappel,

$$|x-y| = \begin{cases} x-y & \text{si } x-y \geqslant 0 \\ y-x & \text{si } x-y \leqslant 0. \end{cases} \tag{1.467}$$

Nous pouvons maintenant démontrer nos assertions.

(i) **(1)** En deux parties.

(i) $\Rightarrow$ Nous supposons que $|x-y| < \epsilon$.

Si $x-y \geqslant 0$ alors l'hypothèse signifie $x-y < \epsilon$, ce qui donne $y > x-\epsilon$. Mais l'inégalité $x-y \geqslant 0$ donne également $x \geqslant y$ et donc $x+\epsilon \geqslant y+\epsilon > y$. Notez le jeu de l'inégalité non stricte qui se change en inégalité stricte.

Si $x-y \leqslant 0$ nous pouvons faire le même raisonnement.

(ii) $\Leftarrow$ Des inégalités $x-\epsilon < y$ et $y < x+\epsilon$ nous tirons $x-y < \epsilon$ et $y-x < \epsilon$. Donc quel que soit le signe de $x-y$ nous avons toujours $|x-y| < \epsilon$.

(ii) **(2)** C'est immédiat parce que

$$|x-y| \leqslant \epsilon < \epsilon'. \tag{1.468}$$

□

Lemme 1.325.
Tout corps totalement ordonné est de caractéristique nulle.

1.17 Les rationnels

Une construction très explicite est faite dans [15]. Ici nous allons prendre plus court :

Définition 1.326.
*Le corps des fractions de $\mathbb{Z}$ (définition 1.313) est noté $\mathbb{Q}$ et ses éléments sont les **rationnels**.*

1.327.
Les résultats énoncés ici sont utilisés plus bas et servent de guide à un éventuel contributeur qui voudrait écrire une partie dédiée à $\mathbb{Q}$ et ses propriétés de base[123]. Nous espérons que des preuves se trouvent dans [15]. En tout cas, cher lecteur, je t'invite à ne rien prendre pour évident.

Proposition 1.328.
L'application

$$\begin{aligned} f\colon \mathbb{Z} &\to \mathbb{Q} \\ z &\mapsto z/1 \end{aligned} \tag{1.469}$$

est une injection.

Démonstration. Supposons que $f(a) = f(b)$, c'est-à-dire que $a/1 = b/1$. En vertu de la relation d'équivalence donnée en 1.313, nous avons $a1 = b1$, c'est-à-dire $a = b$. □

1.329.
À partir de maintenant, nous allons identifier la partie $f(\mathbb{Z})$ à $\mathbb{Z}$. Nous nous autorisons donc à dire que $4 \in \mathbb{Q}$ ou que $-7 \in \mathbb{Q}$, et même que $0 \in \mathbb{Q}$.

Proposition 1.330.
L'ensemble des rationnels est infini dénombrable[124].

123. Par exemple, définir une relation d'ordre sur $\mathbb{Q}$ et expliciter l'inclusion de $\mathbb{Z}$ dans $\mathbb{Q}$.
124. Ouais, je sais, dans les définitions prises ici, un ensemble dénombrable est toujours infini. Mais l'excès de précision ne tue pas, loin s'en faut.

Démonstration. L'ensemble $\mathbb{Z}$ est infini[125] et la proposition 1.328 donne une injection $f\colon \mathbb{Z} \to \mathbb{Q}$. Donc $f(\mathbb{Z})$ est infini.

L'ensemble $\mathbb{Q}$ contient une partie infinie. Il est donc infini par le lemme 1.113.

L'application

$$\begin{aligned} g\colon \mathbb{Z}^2 &\to \mathbb{Q} \\ a, b &\mapsto a/b \end{aligned} \tag{1.470}$$

est surjective alors que $\mathbb{Z}^2$ est dénombrable. Le lemme 1.124 dit alors que $\mathbb{Q}$ est fini ou dénombrable. Vu que nous avons déjà prouvé que $\mathbb{Q}$ était infini, nous déduisons que $\mathbb{Q}$ est infini dénombrable. □

Lemme 1.331.
Soient $a, b, x, y \in \mathbb{Z}$. Nous avons $ay = xb$ dans $\mathbb{Z}$ si et seulement si nous avons $ab^{-1} = xy^{-1}$ dans $\mathbb{Q}$.

1.17.1 Relation d'ordre

Proposition-Définition 1.332.
Pour $a, b, c, d \in \mathbb{Z}$ nous disons que

$$\frac{a}{b} \geqslant \frac{c}{d} \tag{1.471}$$

si et seulement si $ad \geqslant bc$ dans $\mathbb{Z}$.

Avec cette définition, $(\mathbb{Z}, \geqslant)$ est un ensemble totalement ordonné.

Lemme 1.333.
Tout rationnel est majoré par un naturel.

Proposition 1.334.
Si $q < 1$ dans $\mathbb{Q}$, alors $qx < x$ pour tout $x \in \mathbb{Q}^+$.

Proposition 1.335.
Le corps $\mathbb{Q}$ est archimédien[126].

1.17.2 Caractéristique

Lemme 1.336.
Le corps $\mathbb{Q}$ est de caractéristique[127] *nulle.*

1.18 Suite de Cauchy dans un corps totalement ordonné

Lemme 1.337 ([56, 1])**.**
Tout corps commutatif de caractéristique nulle contient un sous-corps isomorphe à $\mathbb{Q}$.

Démonstration. Soit un corps $\mathbb{K}$ de caractéristique nulle. Nous savons du lemme 1.294 que

$$\begin{aligned} \mu\colon \mathbb{Z} &\to \mathbb{K} \\ n &\mapsto n1_{\mathbb{K}} \end{aligned} \tag{1.472}$$

est un morphisme d'anneaux vérifiant $\ker(\mu) = \{0\}$. Nous posons $Z = \mu(\mathbb{Z})$. L'application $\mu\colon \mathbb{Z} \to Z$ est un isomorphisme d'anneaux. Prouvons cela :

(i) **Morphisme** L'application μ est un morphisme par le lemme 1.294.

(ii) **Surjectif** Par définition les éléments de Z sont dans l'image de $\mathbb{Z}$.

(iii) **Injectif** Si $x, y \in \mathbb{Z}$ vérifient $\mu(x) = \mu(y)$, alors $\mu(x - y) = 0$ parce que μ est un morphisme. Mais $\mathbb{K}$ est de caractéristique nulle, c'est-à-dire $\ker(\mu) = \{0\}$. Donc $x - y = 0$.

125. Lemme 1.108.
126. Définition 1.69.
127. Définition 1.294.

Le corps $\mathbb{K}$ contient donc un sous-anneau isomorphe à $\mathbb{Z}$. Puisque $\mathbb{Z}$ et $\mathbb{K}$ sont commutatifs, la proposition 1.314 s'applique et $\mathbb{K}$ contient un sous-corps isomorphe à $\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q}$. □

La proposition suivante donne des précisions à propos du lemme 1.337.

Proposition 1.338 ([1]).
Soit un corps totalement ordonné $\mathbb{K}$. Nous considérons l'application

$$\begin{aligned} \mu\colon \mathbb{Z} &\to \mathbb{K} \\ n &\mapsto n \cdot 1_{\mathbb{K}} \end{aligned} \tag{1.473}$$

et ensuite

$$\begin{aligned} \sigma\colon \mathbb{Q} &\to \mathbb{K} \\ a/b &\mapsto \mu(a)\mu(b)^{-1}. \end{aligned} \tag{1.474}$$

Alors

(1) L'application σ est bien définie.
(2) L'application σ est un morphisme de corps.
(3) Si $q \leqslant q'$ dans $\mathbb{Q}$, alors $\sigma(q) \leqslant \sigma(q')$.

Démonstration. En plusieurs morceaux.

(i) **σ est bien définie** Montrons que σ est bien définie. Pour cela nous considérons $a, b, x, y \in \mathbb{Z}$ tels que $a/b = x/y$ dans $\mathbb{Q}$. Par définition des classes (définition 1.313 du corps des fractions), nous avons $ay = bx$ dans $\mathbb{Q}$. Vu que μ est un morphisme nous avons alors

$$\mu(a)\mu(y) = \mu(b)\mu(x) \tag{1.475}$$

et donc $\mu(a)\mu(b)^{-1} = \mu(x)\mu(y)^{-1}$, c'est-à-dire $\sigma(a/b) = \sigma(x/y)$. L'application σ est donc bien définie.

(ii) **Morphisme pour la somme** L'application μ est un morphisme d'anneaux, comme déjà dit depuis le lemme 1.294. Notons aussi que, parce que $\mathbb{K}$ est commutatif,

$$\mu(qy)^{-1} = \mu(q)^{-1}\mu(y)^{-1}. \tag{1.476}$$

En utilisant la définition 1.313(1) de la somme nous avons

$$\begin{aligned} \sigma(p/q + x/y) &= \sigma\big((py + qx)/qy\big) && (1.477a)\\ &= \big[\mu(py) + \mu(qx)\big]\mu(qy)^{-1} && (1.477b)\\ &= \mu(py)\mu(qy)^{-1} + \mu(qx)\mu(qy)^{-1} && (1.477c)\\ &= \mu(p)\mu(q)^{-1} + \mu(x)\mu(y)^{-1} && (1.477d)\\ &= \sigma(p/q) + \sigma(x/y). && (1.477e) \end{aligned}$$

(iii) **Morphisme pour le produit** Même genre de calculs que pour la somme.

(iv) **Croissante** Nous savons aussi par le lemme 1.321(4) que $1 \geqslant 0$. Puisque μ est un morphisme d'anneaux,

$$\mu(n+1) = \mu(n) + \mu(1) = \mu(n) + 1 \tag{1.478}$$

La définition 1.318(1a) dit alors que $\mu(n) \geqslant 0$ pour tout $n \in \mathbb{N}$. Nous avons pour la même raison que si $m \geqslant n$ dans $\mathbb{N}$, alors $\mu(m) \geqslant \mu(n)$ dans $\mathbb{K}$.
Soient maintenant $p, q \in \mathbb{N}$, et prouvons que $\sigma(p/q) \geqslant 0$. D'abord

$$\sigma(p/q) = \mu(p)\mu(q)^{-1} \tag{1.479}$$

où $\mu(p) \geqslant 0$ et $\mu(q) \geqslant 0$. Ensuite le lemme 1.321(5) nous indique que $\mu(q)^{-1} \geqslant 0$. Enfin la condition 1.318(1b) nous permet de conclure que $\sigma(p/q) \geqslant 0$.
Finalement, si $q_1 \geqslant q_2$ dans $\mathbb{Q}$, alors $q_1 - q_2 \geqslant 0$, et nous avons

$$\sigma(q_1) = \sigma(q_2 + q_1 - q_2) = \sigma(q_2) + \sigma(q_1 - q_2) \geqslant \sigma(q_2) \tag{1.480}$$

par la condition 1.318(1a).

□

1.339.
Si $\mathbb{K}$ est un corps totalement ordonné, la proposition 1.338 nous donne un morphisme de corps $\sigma\colon \mathbb{Q} \to \mathbb{K}$ qui respecte l'ordre. Pour $q \in \mathbb{Q}$ et $k \in \mathbb{K}$ nous notons

$$qk = \sigma(q)k. \tag{1.481}$$

Nous pourrons donc écrire $\frac{k}{2}$ pour $\sigma(1/2)k$.

Le lemme suivant explique que la notation (1.481) n'est pas complètement idiote.

Lemme 1.340.
Soit un corps commutatif totalement ordonné $\mathbb{K}$. Soit $k \in \mathbb{K}$. Nous avons

$$k + k = 2k. \tag{1.482}$$

Démonstration. Vu que $\sigma\colon \mathbb{Q} \to \mathbb{K}$ est un morphisme, il vérifie $\sigma(1) = 1$, donc

$$\begin{aligned} k + k &= \sigma(1)k + \sigma(1)k && (1.483a)\\ &= \big(\sigma(1) + \sigma(1)\big)k && (1.483b)\\ &= \sigma(2)k && (1.483c)\\ &= 2k. && (1.483d) \end{aligned}$$

□

Proposition 1.341.
Toute suite convergente dans un corps totalement ordonné est de Cauchy.

Démonstration. Soit un corps totalement ordonné $\mathbb{K}$ et une suite $x_n \xrightarrow{\mathbb{K}} x$. Soit $\epsilon > 0$. Il est important de se rendre compte que $\epsilon \in \mathbb{K}$ et que l'inégalité est au sens de l'ordre dans $\mathbb{K}$; en particulier ce n'est pas $\epsilon \in \mathbb{R}$ ni $\epsilon \in \mathbb{Q}$. D'ailleurs nous n'avons pas encore défini $\mathbb{R}$.

Vu que (x_n) converge vers x, il existe $N \in \mathbb{N}$ tel que pour tout $k > N$,

$$|x_k - x| < \epsilon. \tag{1.484}$$

Soient $p, q > N$. Alors en utilisant la majoration du lemme 1.322(6),

$$|x_p - x_q| = \big|(x_p - x) + (x - x_q)\big| \leqslant |x_p - x| + |x - x_q| \leqslant 2\epsilon. \tag{1.485}$$

En analyse en général, on s'arrête là et on dit que (x_n) est de Cauchy parce qu'il n'y a pas vraiment de différence entre réaliser une majoration avec ϵ ou avec 2ϵ. Détaillons toutefois comment ça se passe dans le cas où ϵ est un élément d'un corps totalement ordonné.

Le 2ϵ arrivant à la fin de (1.485) est en réalité $\epsilon + \epsilon = \sigma(2)\epsilon$ en vertu de ce qui est raconté en 1.339 et en vertu du lemme 1.340.

Considérons $\epsilon' = \sigma(1/2)\epsilon$, que nous pouvons noter $\epsilon' = \epsilon/2$. Vu que $\epsilon' > 0$, il existe un N' tel que pour tout $p, q > N'$ nous ayons

$$|x_p - x_q| \leqslant 2\epsilon' = \sigma(2)\sigma(1/2)\epsilon = \sigma(1)\epsilon = \epsilon. \tag{1.486}$$

Ce dernier ϵ étant bien celui fixé au début de la preuve, nous en déduisons que (x_n) est de Cauchy. □

1.18.1 Suites de Cauchy dans les rationnels

Proposition 1.342 ([15]).
Principales propriétés des suites de Cauchy dans $\mathbb{Q}$.

(1) Toute suite convergente est de Cauchy[128].

(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 il existe $a, b \in \mathbb{Q}$ tels que $x_n \to a$ 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) C'est la proposition 1.341.

(2) Soit (x_n) une suite de Cauchy dans $\mathbb{Q}$. Avec $\epsilon = 1$ dans la définition, si $q > N$, nous avons

$$|x_q - x_N| \leqslant 1. \tag{1.487}$$

Et donc pour tout q plus grand que N, $x_N - 1 \leqslant x_q \leqslant x_N + 1$, ou encore, pour tout n :

$$|x_n| \leqslant \max\{|x_1|, |x_2|, \dots, |x_N|, |x_N + 1|\}. \tag{1.488}$$

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,

$$|x_p + y_p - x_q - y_q| \leqslant |x_p - x_q| + |y_p - y_q|. \tag{1.489}$$

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 (1.489) 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 :

$$|x_p y_p - x_q y_q| \leqslant |x_p y_p - x_q y_p| + |x_q y_p - x_q y_q| \leqslant |y_p||x_p - x_q| + |x_q||y_p - y_q|. \tag{1.490}$$

Puisque (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

$$|x_p y_p - x_q y_q| \leqslant M\epsilon + M\epsilon = 2M\epsilon, \tag{1.491}$$

ce qui prouve que $(x_n y_n)$ est de Cauchy.

(5) En ce qui concerne la somme, nous pouvons tout de suite calculer

$$|x_n + y_n - (a + b)| \leqslant |x_n - a| + |y_n - b|. \tag{1.492}$$

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.

128. Et non la réciproque, qui sera justement la grande innovation des nombres réels.

En ce qui concerne le produit,

$$|x_n y_n - ab| \leqslant |x_n y_n - a y_n| + |a y_n - ab| \leqslant |y_n||x_n - a| + |a||y_n - b|. \tag{1.493}$$

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

$$y_n|x_n - a| + a|y_n - b| \to 0. \tag{1.494}$$

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

$$|x_p - x_q| \leqslant \frac{\alpha}{2}. \tag{1.495}$$

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 tout $k \geqslant n$. 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

$$\left|\frac{1}{x_p} - \frac{1}{x_q}\right| \leqslant \frac{|x_q - x_p|}{|x_p x_q|} \leqslant \frac{4}{\alpha^2}|x_q - x_p| \leqslant \frac{4}{\alpha^2}\epsilon. \tag{1.496}$$

Donc $\left(\frac{1}{x_n}\right)$ est de Cauchy.

□

1.19 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 1.343.
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. □

Le théorème 3.36 nous dira que tous les $\sqrt{n}$ sont irrationnels dès que n n'est pas un carré parfait. Voici déjà le résultat pour $n = 2$. Le fait que $\sqrt{2}$ existe dans $\mathbb{R}$ sera la proposition 1.404.

Proposition 1.344 (Irrationalité de $\sqrt{2}$)**.**
Il n'existe pas de fractions d'entiers dont le carré soit égal à 2.

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 1.343 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 [129]. Il n'y a donc pas de fractions d'entiers dont le carré vaut 2. □

Lemme 1.345 (Série géométrique)**.**
Si $q \neq 1$ dans $\mathbb{Q}$ et $p \in \mathbb{N}$ nous avons

$$\sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{1.497}$$

Démonstration. En posant $S_p = 1 + q + q^2 + \cdots + q^p$, nous avons $S_p - qS_p = 1 - q^{p+1}$ et donc

$$S_p = \sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{1.498}$$

□

Proposition 1.346.
La suite donnée par

$$x_n = 1 + \frac{1}{1!} + \cdots + \frac{1}{n!} \tag{1.499}$$

est de Cauchy et ne converge pas dans $\mathbb{Q}$.

Démonstration. Si $p > q > 0$ nous avons

$$\begin{aligned}
x_p - x_q &= \sum_{k=q+1}^{p} \frac{1}{k!} && (1.500a)\\
&\leqslant \sum_{k=q+1}^{p} \frac{1}{(q+1)!}\frac{1}{(q+1)^{k-q-1}} && (1.500b)\\
&\leqslant \frac{1}{(q+1)!}\lim_{p\to\infty}\sum_{k=0}^{p}\frac{1}{(q+1)^k} && (1.500c)\\
&\leqslant \frac{1}{(q+1)!}\frac{1}{1-\frac{1}{q+1}} && (1.500d)\\
&\leqslant \frac{1}{(q+1)!}\frac{q+1}{q} && (1.500e)\\
&\leqslant \frac{1}{q!q}. && (1.500f)
\end{aligned}$$

Justifications :

— Pour (1.500b), 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 (1.500c), il y a une inégalité parce que la suite $p \mapsto \sum_{k=0}^{p} 1/(q+1)^k$ est une suite strictement croissante.

129. Voir [15], et attention : ce n'est pas tout à fait évident.

— Pour (1.500d), le lemme 1.345 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

$$x_p - x_q \leqslant \frac{1}{b!b} < \frac{1}{b} < \frac{a}{b} = \epsilon. \tag{1.501}$$

Montrons par l'absurde que cette suite ne converge pas dans $\mathbb{Q}$. Pour cela, nous supposons que $\lim_{n\to\infty} x_n = \frac{a}{b} \in \mathbb{Q}$. Pour tout $p > q$ nous avons établi

$$0 < x_p - x_q < \frac{1}{qq!}. \tag{1.502}$$

Prenons la limite $p \to \infty$; par stricte croissance de la suite, les inégalités restent strictes :

$$0 < \frac{a}{b} - x_q < \frac{1}{qq!}. \tag{1.503}$$

Si $n > b$ alors nous pouvons écrire

$$\frac{a}{b} - x_n = \frac{\alpha}{n!} \tag{1.504}$$

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 (1.503) nous pouvons écrire

$$0 < \frac{\alpha}{q!} < \frac{1}{qq!}, \tag{1.505}$$

c'est-à-dire $0 < \alpha < \frac{1}{q}$, ce qui est impossible pour $\alpha \in \mathbb{Z}$. □

Lemme 1.347.
Soit $A > 0$ dans $\mathbb{Q}$. Il existe un rationnel $q > 0$ tel que $q^2 < A$.

Démonstration. Vu que $\mathbb{Q}$ est archimédien (proposition 1.335), il existe $n \in \mathbb{N}$ tel que $1 < nA$. Pour ce n, nous avons

$$\left(\frac{1}{n}\right)^2 < \frac{1}{n} < A. \tag{1.506}$$

□

La proposition suivante donne une suite de rationnels qui convergerait dans $\mathbb{R}$ vers $\sqrt{A}$ (non encore défini à ce stade). Il est expliqué dans [61] que la suite peut être vue comme une forme de méthode de Newton 34.54 ; voir l'exemple 34.57. Si vous aimez les dessins et les approches géométriques, il y a une explication sur Wikipédia[62].

Proposition 1.348 ([61])**.**
Soient $A > 0$ dans $\mathbb{Q}$ et $x_0 \in \mathbb{Q}$. La suite (x_k) définie par

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{A}{x_k}\right) \tag{1.507}$$

a les propriétés suivantes :

(1) La suite $y_k = x_k^2$ converge dans $\mathbb{Q}$ vers A.

(2) La suite (x_k) est de Cauchy dans $\mathbb{Q}$.

(3) La suite (x_k) ne converge pas dans $\mathbb{Q}$ dans le cas de $A = 2$.

Démonstration. En plusieurs points.

(i) **La suite s_k** En posant $y_k = x_k^2$ nous calculons que

$$y_{k+1} - A = \frac{(y_k - A)^2}{4y_k}. \tag{1.508}$$

Autrement dit, la suite $s_k = y_k - A$ vérifie

$$s_{k+1} = \frac{s_k^2}{4(A + s_k)}. \tag{1.509}$$

Quelle que soit la valeur de $s_0 = x_0^2 - A$, nous avons

$$s_1 = \frac{s_0^2}{4(A + s_0)} = \frac{(x_0^2 - A)^2}{4(A + x_0^2 - A)} = \frac{(x_0^2 - A)^2}{4x_0^2} > 0. \tag{1.510}$$

Donc à partir de s_1, tous les éléments sont positifs. Vu que $A > 0$ nous avons aussi

$$s_{k+1} < \frac{s_k^2}{4s_k} = \frac{s_k}{4} \tag{1.511}$$

et donc $s_k < s_0/4^k$. Donc $s_k \to 0$.

(ii) **La suite (y_k)** Nous venons de prouver que si $y_k = A + s_k$, alors $s_k \to 0$. Autrement dit, la suite y_k converge vers A dans $\mathbb{Q}$.
La suite (y_k) est donc de Cauchy par la proposition 1.342(1).

(iii) **La suite (x_k) est de Cauchy** Soit $\epsilon > 0$ dans $\mathbb{Q}$. Puisque (y_k) est de Cauchy, il existe $n_0 \in \mathbb{N}$ tel que

$$|x_r^2 - x_s^2| < \epsilon \tag{1.512}$$

pour tout $r, s \geqslant n_0$.
Soit par ailleurs $q \neq 0$ dans $\mathbb{Q}$ tel que $q^2 < A$, assuré par le lemme 1.347. Quitte à augmenter la valeur de n_0, nous supposons que $x_r, x_s > q$, et en particulier que $x_r + x_s \neq 0$. Cela permet d'écrire d'abord

$$x_r^2 - x_s^2 = (x_r + x_s)(x_r - x_s) \tag{1.513}$$

et ensuite de prendre la valeur absolue et de diviser par $|x_r + x_s|$:

$$|x_r - x_s| = \frac{|x_r^2 - x_s^2|}{|x_r + x_s|} < \frac{\epsilon}{2q}. \tag{1.514}$$

Donc (x_k) est une suite de Cauchy.

(iv) **Pas de convergence pour $A = 2$** Supposons que $x_k \to a \in \mathbb{Q}$. Dans ce cas nous aurions $x_k^2 \to a^2 = A = 2$ (proposition 1.342(5)). Mais nous savons par la proposition 1.344 que $a^2 = 2$ est impossible dans $\mathbb{Q}$.

□

Notons que cette proposition ne présume en rien de l'existence ou de la non-existence dans $\mathbb{Q}$ d'un élément qui pourrait décemment être nommé $\sqrt{A}$. Il se fait que le théorème 3.36 dira que $\sqrt{n}$ est soit entier, soit irrationnel.

1.349.
Un petit programme en python pour explorer la suite de la proposition 1.348.

```
#!/usr/bin/python3

def rec(A,x):
    return ((x**2+A)/x)/2

```

```
A = 3          # Compute square root of 3
x = 1000       # Initial guess: 1000

for i in range(1,100):
    print(i, x, x**2, x**2-A)
    x = rec(A, x)
```

tex/frido/codeSnip_4.py

1.20 Les réels

Une construction des réels via les coupures de Dedekind est donnée dans [63].

1.350.
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 7.225) alors que la topologie métrique (celle entre autres de $\mathbb{Q}$) ne sera définie que par le théorème 7.98. Nous avons donc dû définir en la définition 1.318 *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 notion d'espace métrique que nous allons voir que sur $\mathbb{R}$, ces trois notions coïncident avec celles topologiques [130]. Et par conséquent que $\mathbb{R}$ sera un espace métrique complet [131].

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[64].

1.20.1 L'ensemble

Soit $\mathcal{E}$ l'ensemble des suites de Cauchy [132] dans $\mathbb{Q}$. Soit aussi l'ensemble $\mathcal{E}_0$ constituée des suites qui convergent vers zéro [133].

En posant

$$x + y = (x_n + y_n) \tag{1.515}$$

et

$$xy = (x_n y_n), \tag{1.516}$$

l'ensemble $\mathcal{E}$ devient un anneau [134] commutatif dont le neutre de l'addition est la suite constante $x_n = 0$ et le neutre pour la multiplication est la suite constante $x_n = 1$.

Proposition 1.351.
La partie $\mathcal{E}_0$ est un idéal [135] *de l'anneau $\mathcal{E}$.*

Démonstration. Nous savons par la proposition 1.342(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 1.342 (il s'agit du fait que la somme de deux suites convergeant vers zéro est une suite convergente vers zéro).

130. Proposition 7.233.
131. Théorème 1.389 pour la complétude en tant que corps et théorème 1.341 pour la complétude en tant que espace métrique.
132. Définition 1.318(3)
133. Nous rappelons qu'à ce niveau nous n'avons pas encore prouvé que toutes les suites de Cauchy convergent.
134. Définition 1.38.
135. Définition 1.213.

En ce qui concerne la propriété fondamentale des idéaux, si $x \in \mathcal{E}_0$ et $y \in \mathcal{E}$ nous devons prouver que $xy \in \mathcal{E}_0$. Puisque $(\mathcal{E}_0, \cdot)$ 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 1.342(3)). □

Théorème-Définition 1.352 (L'anneau des réels[15]).
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.

(i) **Les opérations sont bien définies** La partie $\mathcal{E}_0$ est un idéal par la proposition 1.351. Le quotient est donc bien défini et est un anneau par la proposition 1.214(2).

(ii) **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

$$\bar{x} = \{\text{suites qui convergent vers } q\}. \tag{1.517}$$

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

$$y_n = x_n + (y_n - x_n), \tag{1.518}$$

mais $y_n - x_n \to 0$. Donc la suite $y - x \in \mathcal{E}_0$ ce qui signifie que $y \in \bar{x}$.

(iii) **Neutre et unité** Il est vite vérifié que $\bar{0}$, la classe de la suite constante égale à zéro est neutre pour l'addition. De même, $\bar{1}$, est un neutre pour la multiplication.

(iv) **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[136] alors il existe $y \in \mathcal{E}$ tel que $xy \in \bar{1}$.

Nous savons par la proposition 1.342(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

$$y_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ \frac{1}{x_n} & \text{si } n > n_0. \end{cases} \tag{1.519}$$

Nous avons alors

$$(xy)_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ 1 & \text{si } n > n_0 \end{cases} \tag{1.520}$$

et donc $xy \in \bar{1}$.

□

1.353 (Quelques notations entre $\mathbb{Q}$ et $\mathbb{R}$).
Si $k \mapsto x_k$ est une suite, nous notons (x_k) avec des parenthèses la suite elle-même. Le k dans (x_k) est un indice muet, et dans les cas où il peut y avoir une ambiguïté, nous pouvons noter $(x_k)_{k \in \mathbb{N}}$. Cette dernière notation est plus lourde, mais plus exacte.

Le mieux est d'écrire simplement x la suite, mais alors il faut être prudent et ne pas noter x la limite. Nous éviterons donc d'écrire $x_k \to x$.

Si (x_k) est une suite de Cauchy dans $\mathbb{Q}$, nous notons $\bar{x}$ l'élément de $\mathbb{R}$ qui lui correspond. En fait $\bar{x} = (x_k)$: $\bar{x}$ est la suite-elle même, mais pour nous souvenir de l'origine nous allons adopter cette notation.

136. $x \in \mathcal{E}$ peut soit ne pas converger du tout, soit converger vers autre chose que zéro.

D'autre part nous définissons

$$\begin{aligned}\varphi\colon \mathbb{Q} &\to \mathbb{R}\\ q &\mapsto \overline{[k \mapsto q]},\end{aligned} \tag{1.521}$$

c'est-à-dire que $\varphi(q)$ est la classe de la suite constante $x_k = q$.

Proposition 1.354.
Soit l'application

$$\begin{aligned}\varphi\colon \mathbb{Q} &\to \mathbb{R}\\ q &\mapsto \bar{q}.\end{aligned} \tag{1.522}$$

où par $\bar{q}$ nous entendons la classe de la suite constante égale à q (qui est de Cauchy).

(1) C'est un homomorphisme de corps 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

$$\varphi(q) = \{\text{suites de Cauchy qui convergent vers zéro}\} \tag{1.523}$$

Mais nous savons aussi que [137]

$$\varphi(q) = \bar{q} = \{\text{suites de Cauchy qui convergent vers } q\} \tag{1.524}$$

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.

Lemme 1.355.
Soient $q, l \in \mathbb{Q}$ tels que $\bar{q} = \bar{l}$. Alors $q = l$ dans $\mathbb{Q}$.

Démonstration. La suite constante $x_n = q$ est un représentant de $\bar{q}$, et la suite constante $y_n = l$ est représentant de $\bar{l}$. Dire que $\bar{l} = \bar{q}$ signifie qu'il existe une suite $z \in \mathcal{E}_0$ tel que

$$x = y + z. \tag{1.525}$$

Pour tout n nous avons donc $x_n = y_n + z_n$, ou encore

$$z_n = q - l \tag{1.526}$$

pour tout n. Puisque z est une suite constante, elle ne peut appartenir à $\mathcal{E}_0$ que si elle est la suite constante nulle, c'est-à-dire si $q = l$. □

1.20.2 Relation d'ordre

Définition 1.356.
Nous définissons les parties $\mathcal{E}^+$ et $\mathcal{E}^-$ de $\mathcal{E}$ par

(1) $x \in \mathcal{E}^+$ si et seulement si pour tout $\epsilon > 0$ (ϵ est dans $\mathbb{Q}$), 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 $\mathcal{E}^{++} = \mathcal{E}^+ \backslash \mathcal{E}_0$.

137. Voir dans la démonstration du théorème 1.352.

Dans le lemme suivant, le point (2) peut sembler perturbant. Il s'agit de dire que si x est la classe de la suite constante 0, alors il est le neutre pour l'addition dans $\mathbb{R}$.

Lemme 1.357 (À propos du zéro)**.**
Nous avons

(1) $\mathcal{E}^+ \cap \mathcal{E}^- = \{\bar{0}\}$.

(2) Si $x = \bar{0}$ *alors* $x = 0$.

Lemme 1.358.
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

$$x_n > \frac{\alpha}{2} > 0 \tag{1.527}$$

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. □

Lemme 1.359 ([1])**.**
Si $x \in \mathcal{E}^{++}$*, alors il existe* N *tel que* $x_n > 0$ *pour tout* $n > N$.

Démonstration. La suite x ne tend pas vers zéro. Donc il existe $\delta > 0$ tel que pour tout $N > 0$ il existe $n > N$ vérifiant $x_n > \delta$.

Mais la suite x est également de Cauchy. Écrivons cette condition pour $\delta/2$. Il existe $N_2 > 0$ tel que $p, q > N_2$ implique $|x_p - x_q| < \delta/2$.

Nous fixons $n > N_2$ tel que $x_n > \delta$. Alors pour tout $p > N_2$ nous avons aussi

$$|x_p - x_n| < \frac{\delta}{2}. \tag{1.528}$$

Cela implique que $x_p > \delta/2 > 0$ pour tout $p > N_2$. □

La proposition suivante est une version plus précise du lemme 1.359.

Proposition 1.360 ([11])**.**
Soit $x \in \mathcal{E}^{++}$*. Il existe* $r \in \mathbb{Q}^+ \backslash \{0\}$ *et* $N \in \mathbb{N}$ *tel que* $x_n \geqslant r$ *pour tout* $n \geqslant N$.

Démonstration. Fixons $\epsilon \in \mathbb{Q}^+ \backslash \{0\}$, et procédons par l'absurde. Du coup nous savons trois choses sur la suite (x_n).

(1) Hypothèse absurde : pour tout $q \in \mathbb{Q}^+ \backslash \{0\}$ et pour tout $N \in \mathbb{N}$, il existe $p \geqslant N$ vérifiant $x_p < q$.

(2) Vu que $x \in \mathcal{E}^{++} \subset \mathcal{E}^+$, il existe $P_1 \in \mathbb{N}$ tel que $x_n > -\epsilon$ pour tout $n \geqslant P_1$.

(3) La suite x est de Cauchy. Donc il existe $P_2 \in \mathbb{N}$ tel que si $m, n \geqslant P_2$, nous avons $|x_m - x_n| < \epsilon/2$.

Nous considérons $P \geqslant \max\{P_1, P_2\}$ et nous prenons $q = \epsilon/2$, $N = P$ dans la propriété (1). Il existe donc $p > P$ tel que $x_p < \epsilon/2$.

Prenons aussi $n > p$ et écrivons les deux autres propriétés :

(1) $x_n > -\epsilon$ parce que $n > p > P > P_1$.

(2) $|x_n - x_p| < \epsilon/2$ parce que $n, p > P > P_2$.

Du coup nous avons

$$x_n < x_p + \frac{\epsilon}{2} < \epsilon, \tag{1.529}$$

et donc $-\epsilon < x_n < \epsilon$.

Au final, nous avons prouvé que pour tout $\epsilon \in \mathbb{Q}^+\backslash\{0\}$, il existe P tel que $n > P$ implique $|x_n| < \epsilon$. Cela signifie que

$$x_n \xrightarrow{\mathbb{Q}} 0, \tag{1.530}$$

c'est-à-dire $x \in \mathcal{E}_0$, ce qui est contraire à l'hypothèse. □

Lemme 1.361 ([15])**.**
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 $x, y \in \mathcal{E}^{++}$ alors le lemme 1.359 nous indique que 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

$$\begin{cases} x_n > -\dfrac{\epsilon}{2} & \text{(1.531a)} \\ |x_n - y_n| < \dfrac{\epsilon}{2}. & \text{(1.531b)} \end{cases}$$

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 1.362 (Positivité dans $\mathbb{R}$)**.**
Vocabulaire et notations.

(1) Nous notons $\mathbb{R} = \mathcal{E}/\mathcal{E}_0$.

(2) Nous notons $\mathbb{R}^+ = \bar{\mathcal{E}}^+$.

(3) Nous notons $\mathbb{R}^- = \bar{\mathcal{E}}^-$.

*(4) Un élément de $\mathbb{R}$ est **positif** si il est la classe d'une suite de Cauchy appartenant à $\mathcal{E}^+$.*

*(5) Un élément de $\mathbb{R}$ est **négatif** si 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 $\mathcal{E}_0$.

1.363.
Avec les conventions de la définition 1.362, 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 ». Ce sont des conventions qui sont également celles de Wikipédia[65].

Définition 1.364 (Ordre sur $\mathbb{R}$)**.**
Si $a, b \in \mathbb{R}$ nous notons $a \leqslant b$ si et seulement si $b - a \in \overline{\mathcal{E}^+}$.

Proposition 1.365.
Le couple $(\mathbb{R}, \leqslant)$ est un corps totalement ordonné[138]

Démonstration. Il s'agit de vérifier, dans l'ordre, les définitions 1.10, 1.11 et 1.318(1). Pour la suite nous considérons $x, y, z \in \mathbb{R}$ et des suites de Cauchy a, b, c telles que $x = \bar{a}$, $y = \bar{b}$ et $z = \bar{c}$.

(i) **Réflexivité** Pour savoir si $x \geqslant x$, nous devons nous demander si $x - x \in \overline{\mathcal{E}^+}$. Nous avons $x - x = \bar{a} - \bar{a} = \bar{0} = 0$.

(ii) **antisymétrie** Nous supposons que $x \geqslant y$ et $y \geqslant x$. Du côté des suites de Cauchy, cela signifie que $a - b \in \mathcal{E}^+$ et $b - a \in \mathcal{E}^+$. Le lemme 1.361(1) nous indique alors que $a - b = -(b - a) \in \mathcal{E}^-$. Donc

$$a - b \in \mathcal{E}^+ \cap \mathcal{E}^- = \{\bar{0}\}. \tag{1.532}$$

Donc le réel $x - y$ est la classe de la suite constante 0. Le lemme 1.357(2) dit alors que $x - y = 0$ ou encore que $x = y$.

(iii) **transitivité** Nous supposons que $x \leqslant y$ et $y \leqslant z$. Alors nous avons

$$\begin{aligned} c - a &= c - b + b - a && (1.533a)\\ &= (c - b) + (b - a) && (1.533b)\\ &\in \mathcal{E}^+ + \mathcal{E}^+ && (1.533c)\\ &\subset \mathcal{E}^+ && (1.533d) \end{aligned}$$

où nous avons utilisé le lemme 1.361(2). Puisque $c - a \in \mathcal{E}^+$, nous avons $z - x = \overline{c - a} \geqslant 0$.

(iv) **Ordre total** Nous devons prouver que pour $x, y \in \mathbb{R}$ nous avons toujours $x \leqslant y$ ou $y \leqslant x$. Supposons que nous n'ayons pas $x \leqslant y$, c'est-à-dire $\overline{b - a} \notin \mathcal{E}^+$. Vu le lemme 1.358(2) nous avons $\overline{b - a} \in \mathcal{E}^-$, ce qui donne, par le lemme 1.361(1) que $\overline{a - b} \in \mathcal{E}^+$, c'est-à-dire $y \leqslant x$.

(v) **Corps ordonné** Enfin nous devons vérifier les deux conditions de la définition 1.318(1). Pour la première condition, nous supposons $x \leqslant y$, c'est-à-dire $b - a \in \mathcal{E}^+$. Nous avons donc

$$(b + c) - (a + c) = b + x - a - c = b - a \in \mathcal{E}^+, \tag{1.534}$$

donc $\overline{a + c} \leqslant \overline{b + c}$, c'est-à-dire $x + z \leqslant y + z$.
Pour la seconde condition, c'est le lemme 1.361(3).

□

Définition 1.366.
Puisque $\mathbb{R}$ est un corps totalement ordonné (proposition 1.365), si $x \in \mathbb{R}$, nous définissons $|x|$ conformément à 1.318(2).

Lemme 1.367.
L'application

$$\begin{aligned} \varphi\colon \mathbb{Q} &\to \mathbb{R}\\ q &\mapsto \bar{q} \end{aligned} \tag{1.535}$$

dont nous avons déjà parlé dans la proposition 1.354 est strictement croissante.

138. Corps totalement ordonné, définition 1.318.

Démonstration. Nous supposons $q < l$ dans $\mathbb{Q}$. Nous devons montrer que $\bar{q} \leqslant \bar{l}$ dans $\mathbb{R}$, c'est-à-dire que $\bar{q} \leqslant \bar{l}$ et $\bar{q} \neq \bar{l}$.

Considérons la suite constante $x_n = l - q \in \mathbb{Q}$. Pour tout $\epsilon > 0$ dans $\mathbb{Q}$ nous avons

$$x_n = l - q > 0 > -\epsilon, \tag{1.536}$$

et donc $x_n \in \mathcal{E}^+$. Donc $\overline{l-q} \geqslant 0$. Cela signifie $\bar{q} \leqslant \bar{l}$.

D'autre part le lemme 1.355 dit que $\bar{q} = \bar{l}$ uniquement si $q = l$, ce qui est exclu parce que $q < l$. Donc $\bar{q} \neq \bar{l}$. □

Lemme 1.368.
Si $a, b \in \mathbb{R}^+$ satisfont $a + b = 0$, alors $a = b = 0$.

Voici une version dans $\mathbb{R}$ du lemme 1.105.

Lemme 1.369.
Soient $a > 0$ et $b > 1$ dans $\mathbb{R}$. Nous avons

$$ab > a. \tag{1.537}$$

Remarque 1.370.
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 1.371.
Si $x, y, z \in \mathbb{R}$ avec $x > 0$ sont tels que $z > y/x$ alors $zx > y$.

Démonstration. Nous savons que

$$z - \frac{y}{x} \in \mathcal{E}^+ \backslash \{0\} = \mathcal{E}^{++}. \tag{1.538}$$

Puisque $x \in \mathcal{E}^{++}$, multiplier par x fait rester dans $\mathcal{E}^{++}$:

$$zx - x\frac{y}{x} \in \mathcal{E}^{++}. \tag{1.539}$$

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 1.372.
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.

(i) **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 1.342(2)). Soit M un naturel strictement plus grand que m/q [139].
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 [140]. Par définition de la relation d'ordre, $M \geqslant a$.

(ii) **Seconde façon** La suite (a_n) est majorée par $\frac{m}{q}$, donc on a dans $\mathbb{Q}$ et pour tout n :

$$a_n \leqslant \frac{m}{q} = M \leqslant qM. \tag{1.540}$$

L'application $\varphi\colon \mathbb{Q} \to \mathbb{R}$ est croissante, donc

$$\varphi\big((a_n)\big) \leqslant \varphi(qM). \tag{1.541}$$

139. Lemme 1.333.
140. Et nous allons d'ailleurs arrêter de toujours préciser « la classe de » lorsque ce n'est pas nécessaire.

□

En corolaire, nous avons

Lemme 1.373.
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 $q = -p$. □

Théorème 1.374 ([15])**.**
Le corps $\mathbb{R}$ est archimédien[141].

Démonstration. La proposition 1.365 dit que $\mathbb{R}$ est totalement ordonné. Soient $x, y \in \mathbb{R}$ avec $x > 0$; posons $a = \frac{y}{x}$. Le lemme 1.372 nous donne un $p \in \mathbb{N}$ tel que $p > a$. Nous concluons alors avec le lemme 1.371 :

$$px > ax = \frac{y}{x}x = y. \tag{1.542}$$

□

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 1.375.
À propos de rationnels entre des réels.

(1) Si $x, y \in \mathbb{R}$ sont tels que $x < y$, alors il existe $s \in \mathbb{Q}$ tel que $x < s < y$.

(2) Si $\epsilon > 0$ dans $\mathbb{R}$, il existe $n \in \mathbb{N}$ tel que $\frac{1}{n} < \epsilon$.

Démonstration. Nous avons par hypothèse que $y - x > 0$ et donc le fait que $\mathbb{R}$ soit archimédien (théorème 1.374) nous donne $q \in \mathbb{N}$ tel que $q(y - x) > 1$. Soit

$$E = \{n \in \mathbb{Z} \text{ tel que } \frac{n}{q} \leqslant x\}. \tag{1.543}$$

Cet ensemble n'est pas vide à cause du lemme 1.373 ; de plus, comme $|x|q \leqslant n_0$ pour un certain n_0 (à cause du lemme 1.372), l'ensemble E est majoré par n_0. Donc E possède un plus grand élément[142] p qui vérifie

$$\frac{p}{q} \leqslant x < \frac{p+1}{q}. \tag{1.544}$$

De plus $(p + 1)/q < y$ parce que

$$\frac{p+1}{q} = \frac{p}{q} + \frac{1}{q} \leqslant x + \frac{1}{q} < x + y - x = y \tag{1.545}$$

où nous avons utilisé l'inégalité stricte $y - x > \frac{1}{q}$.

Nous avons donc

$$x < \frac{p+1}{q} < y, \tag{1.546}$$

et le nombre $(p + 1)/q$ convient comme s. Le point (1) est prouvé.

Pour le point (2), par le point (1) nous considérons $s \in \mathbb{Q}$ tel que $0 < s < \epsilon$. Si $s = p/q$ avec $p, q \in \mathbb{N}$ nous avons

$$0 < \frac{1}{q} \leqslant \frac{p}{q} < \epsilon. \tag{1.547}$$

□

141. Définition 1.69.
142. Lemme 1.106.

Remarque 1.376.
Le lemme 1.375 a également pour conséquence que des ensembles comme $[-1,1]$ ne sont pas bien ordonnés (définition 1.11). 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 $]0,1[$.

Tant que nous y sommes dans les encadrements de réels...

1.377.
Soit $q_0 \in \mathbb{Q}$ tel que $0 \leqslant q_0 < 1$. On définit alors $d_1 \in \{0,1\}$ comme valant 1 si $2q_0 \geqslant 1$ et 0 sinon. Puis on pose $q_1 = 2q_0 - d_1$.

Poursuivant de la sorte, on crée une suite $(d_n)_{n\geqslant 1}$: c'est le **développement dyadique** de q_0.

Lemme 1.378 ([1])**.**
Soit q, q' deux rationnels tels que $0 \leqslant q < q' < 1$. Il existe deux entiers naturels a et N tels que $q < \frac{a}{2^N} < q'$.

Démonstration. On crée les développements dyadiques de q et q', que l'on note respectivement $(d_n)_{n\geqslant 1}$ et $(d'_n)_{n\geqslant 1}$. Notons

$$E = \{n \in \mathbb{N} \text{ tel que } d_n \neq d'_n\}. \tag{1.548}$$

Comme $q \neq q'$, les développements dyadiques sont différents [143], l'ensemble E est non-vide, et il admet un plus petit élément N. Or, $q < q'$, et donc nécessairement $d_N < d'_N$. On construit alors $a = \sum_{i=1}^{N} d_i 2^i$. □

Corolaire 1.379.
Pour tous réels x, y tels que $0 \leqslant x < y \leqslant 1$, il existe un nombre de la forme $d = a/2^n$, avec $n \in \mathbb{N}$ et $a \in \mathbb{N}$, $a \leqslant 2^n$, tel que $x < d < y$.

La partie

$$D = \{\frac{a}{2^b} \text{ tel que } a \in \mathbb{Z}, b \in \mathbb{N}\} \tag{1.549}$$

est dense dans $\mathbb{R}$.

*Ce D est nommé l'ensemble des **fractions dyadiques**.*

Lemme 1.380 ([1])**.**
Soient des réels a, b, x, y tels que

$$a \leqslant x \leqslant b \tag{1.550}$$

et

$$a \leqslant y \leqslant b, \tag{1.551}$$

alors $|x - y| \leqslant |b - a|$.

Lemme 1.381.
Soient deux réels a, b tels que

(1) $a \geqslant 0$,
(2) $b \geqslant 0$
(3) $a + b = 0$.

Alors $a = 0$ et $b = 0$.

Lemme 1.382.
Si $a \in \mathbb{R}$, alors $a^2 \geqslant 0$ et $a^2 = 0$ si et seulement si $a = 0$.

Lemme 1.383.
Soit un réel strictement positif a. Si $b > 1$, alors $ab > a$.

Lemme 1.384.
Soient $a, b, x \in \mathbb{R}$ tels que $a + x = b + x$. Alors $a = b$.

143. À vérifier tout de même...

1.20.3 Complétude

Le théorème 17.136 donne une complétion de tout espace métrique en un espace complet. Il serait tentant de l'utiliser ici pour définir $\mathbb{R}$ à partir de $\mathbb{Q}$. Cette méthode ne fonctionne cependant pas parce que la démonstration de 17.136 utilise le fait que $\mathbb{R}$ est complet.

Lemme 1.385.
Nous avons

$$|\varphi(q)| = \varphi(|q|) \tag{1.552}$$

pour tout $q \in \mathbb{Q}$.

Démonstration. Soit $q \in \mathbb{Q}$. Si $q \geqslant 0$ alors nous avons d'une part $|q| = q$ dans $\mathbb{Q}$, et d'autre part $\varphi(q) \geqslant 0$ dans $\mathbb{R}$. Donc au final

$$|\varphi(q)| = \varphi(q) = \varphi(|q|). \tag{1.553}$$

Supposons au contraire que $q < 0$. Alors $|q| = -q$ dans $\mathbb{Q}$, mais aussi $\varphi(q) \leqslant 0$. Donc

$$|\varphi(q)| = -\varphi(q) = \varphi(-q) = \varphi(|q|). \tag{1.554}$$

□

Lemme 1.386 ([15, 1])**.**
Toute suite de Cauchy dans $\mathbb{Q}$ *converge dans* $\mathbb{R}$ *vers le réel qu'elle représente.*

Plus précisément, en suivant les notations de 1.353, si (x_k) *est une suite de Cauchy dans* $\mathbb{Q}$*, alors*

(1) $\varphi(x_k)$ *est une suite de Cauchy dans* $\mathbb{R}$.

(2) $\varphi(x_k) \xrightarrow{\mathbb{R}} \bar{x}$.

Ici $\bar{x}$ *est la classe de la suite* x*. C'est donc un élément de* $\mathbb{R}$.

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 $\lim_{n\to\infty} \varphi(x_n) = \bar{x}$.

(i) <u>$\varphi(x_k)$ **est de Cauchy**</u> Soit $\epsilon > 0$ dans $\mathbb{R}$. Nous considérons $\epsilon' \in \mathbb{Q}$ tel que $0 < \epsilon' < \epsilon$. Plus précisément tel que

$$0 < \varphi(\epsilon') < \epsilon. \tag{1.555}$$

Soit $N > 0$ tel que $|x_p - x_q| < \epsilon'$ pour tout $p, q \geqslant N$. Cela existe parce que (x_k) est dans Cauchy dans $\mathbb{Q}$. Nous avons alors

$$\begin{aligned}
|\varphi(x_p) - \varphi(x_q)| &= |\varphi(x_p - x_q)| && (1.556a)\\
&= \varphi\big(|x_p - w_q|\big) && (1.556b)\\
&\leqslant \varphi(\epsilon') && (1.556c)\\
&\leqslant \epsilon. && (1.556d)
\end{aligned}$$

Justifications :

— Pour (1.556b). C'est le lemme 1.385.

— Pour (1.556c). L'application φ est croissante, lemme 1.367.

Donc la suite $\varphi(x_k)$ est de Cauchy.

(ii) <u>$\varphi(x_k) \xrightarrow{\mathbb{R}} \bar{x}$</u> Nous devons prouver que pour tout $\epsilon \in \mathbb{R}$, il existe N tel que $n > N$ implique $\varphi(x_n) \in B\big(\bar{x}, \epsilon\big)$. Nous allons faire ça en deux parties. D'abord $\epsilon \in \mathbb{Q}$ et ensuite $\epsilon \in \mathbb{R}$.

(i) **Avec $\epsilon \in \mathbb{Q}$** Soit $\epsilon \in \mathbb{Q}$. Puisque x est une suite de Cauchy dans $\mathbb{Q}$, il existe N tel que si $p, n > N$ nous avons

$$x_p - \epsilon < x_n < x_p + \epsilon. \tag{1.557}$$

Ces inégalités sont dans $\mathbb{Q}$. Nous fixons p et nous commençons par écrire plus en détail la première inéquation :

$$x_p - \epsilon - x_n < 0. \tag{1.558}$$

Autrement dit, pour tout n nous avons

$$(\overline{x_p - \epsilon})_n < x_n. \tag{1.559}$$

Pour rappel, la suite $\overline{x_p - \epsilon}$ est la suite constante dans $\mathbb{Q}$. La suite

$$n \mapsto x_n - (\overline{x_p - \epsilon})_n \tag{1.560}$$

est dans $\mathcal{E}^+$. Donc, en vertu de la définition 1.364 nous avons

$$\bar{x} - \overline{x_p - \epsilon} \geqslant 0. \tag{1.561}$$

Nous pouvons aussi bien écrire

$$\bar{x} \geqslant \varphi(x_p) - \varphi(\epsilon). \tag{1.562}$$

En prenant l'autre inégalité de (1.557) nous trouvons de la même manière que

$$\bar{x} \leqslant \varphi(x_p) + \varphi(\epsilon). \tag{1.563}$$

Ces deux inégalités ensemble montrent que

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon)\big). \tag{1.564}$$

(ii) **Avec $\epsilon \in \mathbb{R}$** Nous considérons $\epsilon \in \mathbb{R}$ et $\epsilon' \in \mathbb{Q}$ tel que $\varphi(\epsilon') < \epsilon$. Par le point précédent, il existe N tel que $p > N$ implique

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon')\big). \tag{1.565}$$

Étant donné que $\varphi(\epsilon') < \epsilon$ nous avons

$$\varphi(x_p) \in B\big(\bar{x}, \varphi(\epsilon')\big) \subset B\big(\bar{x}, \epsilon\big). \tag{1.566}$$

□

Proposition 1.387.
Soit une suite convergente $x_k \xrightarrow{\mathbb{Q}} q$. Alors

$$\varphi(x_k) \xrightarrow{\mathbb{R}} \varphi(q) \tag{1.567}$$

où φ est la fonction qui à un rationnel fait correspondre la classe de la suite constante correspondante[144].

Démonstration. Le fait d'avoir une convergence $x_k \to q$ dans $\mathbb{Q}$ implique que la suite (x_k) est de Cauchy, par la proposition 1.342(1).

Le lemme 1.386 nous indique que $\varphi(x_k)$ est une suite dans $\mathbb{R}$ qui converge vers $\bar{q}$, la classe de la suite (x_k).

À prouver : $\varphi(x) = \bar{q}$. Autrement dit, nous devons prouver que la classe de la suite constante $a_k = q$ et la classe de la suite x sont les mêmes.

La suite $(x_k - q)$ est de Cauchy dans $\mathbb{Q}$ et converge vers zéro par hypothèse. Donc les suites x et (q) sont dans la même classe. □

144. Voir les notations en 1.353.

Proposition 1.388 ([1]).
Deux choses à propos de suites de rationnels convergeant vers un réel.

(1) Soit un réel x. Il existe une suite de rationnels strictement croissante qui converge vers x.

(2) Si de plus $x > 0$, alors la suite (toujours strictement croissante) peut être choisie parmi les rationnels strictement positifs.

Démonstration. Le lemme 1.375 nous sera d'une grande aide. Soit $x \in \mathbb{R}$. Il existe $q_0 \in \mathbb{Q}$ tel que $x - 1 < q_0 < x$. Ensuite nous construisons la suite par récurrence : q_k est choisi tel que $q_{k-1} < q_k < x$. Cela règle le point (1).

Pour (2). Il suffit de faire la même chose, en partant de $0 < q_0 < x$. □

Théorème 1.389 (Complétude de $\mathbb{R}$, critère de Cauchy[15]).
Nous avons :

(1) Le corps $\mathbb{R}$ est un corps complet (définition 1.318(5))

(2) Une suite dans $\mathbb{R}$ est convergente (définition 1.318(4)) si et seulement si elle est de Cauchy (définition 1.318(3)).

Notez la grande similitude entre ce théorème et le théorème 7.247. Ils ne sont pas équivalents, ne parlent pas exactement du même objet « $\mathbb{R}$ », ni des mêmes notions de suites de Cauchy et de complétude.

Démonstration. Soit (x_n) une suite de Cauchy dans $\mathbb{R}$. Pour chaque n, il existe par le lemme 1.375 un $y_n \in \mathbb{Q}$ tel que

$$x_n - \frac{1}{n} < y_n < x_n + \frac{1}{n}. \tag{1.568}$$

(i) **(y_n) est une suite de Cauchy dans $\mathbb{Q}$** Nous prouvons que (y_n) est une suite de Cauchy dans $\mathbb{Q}$ (définition 1.318(3)). Vu que (x_n) est de Cauchy pour le corps $\mathbb{R}$, si $\epsilon > 0$ dans $\mathbb{R}$ est donné, il existe n_ϵ tel que si $p, q \geqslant n_\epsilon$, alors $|x_p - x_q| < \epsilon$.

Nous avons :

$$|y_p - y_q| \leqslant |y_p - x_p| + |x_p - x_q| + |x_q - y_q| < \frac{1}{p} + \epsilon + \frac{1}{q}. \tag{1.569}$$

En choisissant $N_\epsilon > \max\{n_\epsilon, \frac{1}{\epsilon}\}$ (ce qui est possible par le lemme 1.372), et en prenant $p, q > N_\epsilon$, nous avons

$$|y_p - y_q| \leqslant 3\epsilon, \tag{1.570}$$

ce qui prouve que (y_p) est une suite de Cauchy dans $\mathbb{Q}$, pour la notion de suite de Cauchy dans $\mathbb{Q}$.

(ii) **Le réel représenté** Puisque (y_p) est de Cauchy dans $\mathbb{Q}$, elle représente un réel que nous notons $\bar{y}$.

(iii) **Convergence de (x_n)** Nous prouvons 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 1.386). 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

$$|\bar{y} - x_p| \leqslant |\bar{y} - y_p| + |y_p - x_p| \leqslant \epsilon + \frac{1}{p}. \tag{1.571}$$

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 un corps complet.

En ce qui concerne l'équivalence entre les suites convergentes et de Cauchy, nous venons de prouver que toute suite de Cauchy dans $\mathbb{R}$ est convergente. La réciproque est la proposition 1.341. □

Nous avons terminé avec la construction des réels. Les propriétés topologiques arrivent en la section 10.3. En particulier le théorème 7.247 pour la complétude de $\mathbb{R}$ en tant qu'espace métrique.

1.20.4 Intervalles

Nous avons déjà défini la notion d'intervalle pour un espace totalement ordonné en 1.20. Nous posons quelques notations dans $\mathbb{R}$.

Définition 1.390.
Soient $a \neq b$ dans $\mathbb{R}$. Nous définissons les parties suivantes de $\mathbb{R}$:

(1) $]a, b[= \{x \in \mathbb{R} \text{ tel que } a < x < b\}$
(2) $[a, b[= \{x \in \mathbb{R} \text{ tel que } a \leqslant x < b\}$
(3) $]a, b] = \{x \in \mathbb{R} \text{ tel que } a < x \leqslant b\}$
(4) $[a, b] = \{x \in \mathbb{R} \text{ tel que } a \leqslant x \leqslant b\}$
(5) $]-\infty, a] = \{x \in \mathbb{R} \text{ tel que } x \leqslant a\}$
(6) $]-\infty, a[= \{x \in \mathbb{R} \text{ tel que } x < a\}$
(7) $]a, \infty[= \{x \in \mathbb{R} \text{ tel que } x > a\}$
(8) $[a, \infty[= \{x \in \mathbb{R} \text{ tel que } x \geqslant a\}$.
(9) $]-\infty, \infty[= \mathbb{R}$.

La proposition 1.396 nous dira que tous les intervalles de $\mathbb{R}$ sont d'une de ces formes.

1.20.5 Maximum, supremum et compagnie

Ce n'est un secret pour personne que $\mathbb{R}$ est un ensemble totalement ordonné[145] : 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 $\mathbb{R}$.

Définition 1.391.
Soit A, une partie de $\mathbb{R}$.

(1) Un nombre M est un ***majorant*** *de A si M est plus grand que tous les éléments de A : pour tout $x \in A$, $M \geqslant x$.*
(2) Un nombre m est un ***minorant*** *de A si m est plus petit que tous les éléments de A : pour tout $x \in A$, $m \leqslant x$.*

Nous parlons de majorant ou de minorants stricts *lorsque les inégalités sont strictes.*

Nous insistons 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 $s + 1$, $s + 4$, et $s + \frac{3}{7}$ majorent également A.

Exemple 1.392.
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 majorant.

△

Proposition-Définition 1.393 (Least-upper-bound property[66])**.**
Soit A une partie majorée de $\mathbb{R}$. Il existe un unique élément $M \in \mathbb{R}$ tel que

145. Proposition 1.365.

(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* $x > M - \varepsilon$.

Cet élément est nommé ***supremum*** *de* A *et est noté* $\sup(A)$. *De la même façon,* ***l'infimum*** *de* A, *noté* $\inf A$ *est l'unique réel* $m \in \mathbb{R}$ *vérifiant*

(1) $m \leqslant x$ *pour tout* $x \in A$,

(2) pour tout $\epsilon > 0$, *le nombre* $m + \epsilon$ *n'est pas un minorant.*

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 $\infty \notin \mathbb{R}$.

Démonstration. Nous faisons la preuve pour l'infimum.

(i) **Unicité** En ce qui concerne l'unicité, soient m_1 et m_2, deux infimums de A. Supposons $m_1 > m_2$. Alors il existe $\epsilon > 0$ tel que $m_2 < m_2 + \epsilon < m_1$ (c'est le lemme 1.375). Cela prouve que $m_2 + \epsilon$ est un minorant de A et donc que m_2 n'est pas un infimum.

(ii) **Existence** 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) :

$$\begin{aligned} &x_0 \text{ est un élément de } A,\\ &x_1 \text{ est un minorant de } A,\\ &a_0 = x_0\\ &b_0 = x_1\\ &b_1 = x_1. \end{aligned} \tag{1.572}$$

Ensuite, nous faisons la récurrence suivante :

$$\begin{aligned} x_{n+1} &= \frac{a_n + b_n}{2},\\ a_{n+1} &= \begin{cases} a_n & \text{si } x_{n+1} \text{ minore } A\\ x_{n+1} & \text{sinon,} \end{cases}\\ b_{n+1} &= \begin{cases} x_{n+1} & \text{si } x_{n+1} \text{ minore } A\\ b_n & \text{sinon.} \end{cases} \end{aligned} \tag{1.573}$$

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

$$\begin{aligned} |a_n - b_n| &= |a_{n-1} - x_n|\\ &= \left| a_{n-1} - \frac{a_{n-1} + b_{n-1}}{2} \right|\\ &= \frac{1}{2}|a_{n-1} - b_{n-1}|, \end{aligned} \tag{1.574}$$

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

$$|a_n - a_{n-1}| = \begin{cases} 0 \\ \left|\frac{a_n - b_n}{2}\right| \end{cases} \leqslant \frac{1}{2n}. \tag{1.575}$$

Il en est de même pour la suite (b_n). Ce sont deux suites de Cauchy (donc convergentes par la proposition 1.341) 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. □

Lemme 1.394.
Soit une partie A de $\mathbb{R}$. Si M est un majorant de A, alors $M \geqslant \sup(A)$.

Démonstration. Si $M < \sup(A)$, alors en posant $\epsilon = \sup(A) - M$, le nombre $\sup(A) - \epsilon$ est encore un majorant de A, ce qui est impossible par définition d'un supremum. □

1.20.5.1 Intervalles

Lemme 1.395 ([1]).
Soit une partie A de $\mathbb{R}$.

(1) Nous supposons que A admette un supremum qui n'est pas dans A. Si x est un élément de A strictement plus petit que $\sup(A)$, alors il existe $y \in A$ tel que $x < y < \sup(A)$.

(2) Nous supposons que A admette un infimum qui n'est pas dans A. Si x est un élément de A strictement plus grand que $\inf(A)$, alors il existe $y \in A$ tel que $\inf(A) < y < x$.

Démonstration. Soit $\epsilon > 0$. Par définition 1.393 d'un supremum, le nombre $\sup(A) - \epsilon$ n'est pas un majorant de A. Autrement dit, il existe $y \in A$ tel que $y > \sup(A) - \epsilon$. Vu que $\sup(A)$ est plus grand que tous les éléments de A et qu'il n'est pas lui-même dans A, nous avons aussi $\sup(A) > y$. En mettant bout à bout :

$$x < \sup(A) - \epsilon < y < \sup(A). \tag{1.576}$$

□

Proposition 1.396 ([1]).
Tous les intervalles[146] *de $\mathbb{R}$ sont d'une des formes listées dans la définition 1.390.*

Démonstration. Il y a beaucoup de cas, et nous ne les feront pas tous[147].

(i) **Si I est borné vers le haut et vers le bas** Il y a 4 possibilités suivant que $\inf(I)$ et $\sup(I)$ soient ou non dans I.

(ii) **Si $\inf(I) \in I$ et $\sup(I) \in I$** Nous prouvons que $I = [\inf(I), \sup(II)]$.

  (i) **Dans un sens** Si $x \in I$ nous avons $\inf(I) \leqslant x \leqslant \sup(I)$ parce que $\inf(I)$ est un minorant de toute élément de I alors que $\sup(I)$ est un majorant de tout élément de I. Donc $x \in [\inf(I), \sup(I)]$.

  (ii) **Dans l'autre sens** Soit $\inf(I) \leqslant x \leqslant \sup(I)$. Vu que I est un intervalle et que $\sup(I)$ et $\inf(I)$ sont deux éléments de I, nous avons $x \in I$.

(iii) **$\inf(I) \in I$ et $\sup(I) \notin I$** Nous allons démontrer que $I = [\inf(I), \sup(I)[$.

  (i) **Dans un sens** Soit $x \in I$. Nous savons (par définition de l'infimum et du supremum) que $\inf(I) \leqslant x \leqslant \sup(x)$. Mais nous sommes dans un cas où $\sup(I) \neq x$ parce que $\sup(I)$ n'est pas dans I. Donc

$$\inf(I) \leqslant x < \sup(I), \tag{1.577}$$

  ce qui montre que $I \subset [\inf(I), \sup(I)[$.

  (ii) **Dans l'autre sens** Soit $x \in [\inf(I), \sup(I)[$. Nous utilisons le lemme 1.395(1) : il existe $y \in I$ tel que $\inf(I) \leqslant x < y < \sup(I)$. Vu que $\inf(I) \in I$, que $y \in I$ et que I est un intervalle, nous avons aussi $x \in I$ et donc $I \subset [\inf(I), \sup(I)[$.

146. Définition 1.20.
147. Si vous avez un doute, écrivez-moi.

(iv) $\inf(I) \notin I$ **et** $\sup(I) \in I$ C'est le même raisonnement, mais en utilisant l'autre partie du lemme.

(v) $\inf(I) \notin I$ **et** $\sup(I) \notin I$ Nous prouvons que $I =]\inf(I), \sup(I)[$.

(i) **Dans un sens** Nous avons

$$I \subset \{x \in \mathbb{R} \text{ tel que } \inf(I) \leqslant x \leqslant \sup(I)\}. \tag{1.578}$$

Mais comme $\inf(I)$ et $\sup(I)$ ne sont pas dans I, nous pouvons les enlever dans le membre de droite :

$$I \subset \{x \in \mathbb{R} \text{ tel que } \inf(I) < x < \sup(I)\} =]\inf(I), \sup(I)[. \tag{1.579}$$

(ii) **Dans l'autre sens** Si $x \in I$, nous avons $\inf(I) < x < \sup(I)$. En utilisant les deux parties du lemme, nous avons y_1 et y_2 dans I tels que

$$\inf(I) < y_1 < x < y_2 < \sup(I). \tag{1.580}$$

Vu que I est un intervalle, nous en déduisons que $x \in I$.

(vi) $\inf(I) = -\infty$, $\sup(I) \in I$ Nous prouvons que $I =]-\infty, \sup(I)]$.

(i) **Dans un sens** L'inclusion $I \subset \{x \in \mathbb{R} \text{ tel que } x \leqslant \sup(I)\}$ est automatique : $\sup(I)$ majore tous les éléments de I.

(ii) **Dans l'autre sens** Soit $x \leqslant \sup(I)$. Vu que I n'a pas d'infimum, il existe $y \in I$ tel que $y < x < \sup(I)$. Comme I est un intervalle nous déduisons que $x \in I$.

Je vous laisse voir les autres cas. □

1.20.5.2 Quelques exemples

En matière 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 1.397.
Exemples de différence entre majorant, supremum et maximum.

— Le nombre 10 est un supremum, majorant et maximum de l'intervalle fermé $[0, 10]$,

— 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 :

$$10 = \max[0, 10] = \sup[0, 10] \quad 10 = \sup[0, 10[\tag{1.581}$$

Exemple 1.398.
Si on dit qu'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 1.399.
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. △

Lemme 1.400.
À propos de bornes d'un intervalle dans $\mathbb{R}$.

(1) La borne inférieure[148] *d'un intervalle est son infimum,*

(2) la borne supérieure est le supremum.

(3) Si de plus l'intervalle est fermé, l'infimum est un minimum et le supremum est un maximum.

Démonstration. Soit I un des intervalles $[a,b]$, $[a,b[$, $]a,b]$, $]a,b[$. Nous allons montrer que dans tous ces cas, a est l'infimum de I.

Nous allons prouver que dans tous ces cas, $a = \int(I)$ en vérifiant les deux conditions de la définition 1.393. D'abord par définition d'un intervalle, pour tout $t \in I$ nous avons $t \geqslant a$. Première condition vérifiée. Ensuite, nous prenons $\epsilon > 0$. Si $a + \epsilon \geqslant b$, alors n'importe quel élément de I est plus petit que $a + \epsilon$. Si $a + \epsilon < b$, alors nous avons

$$a < a + \epsilon < b. \tag{1.582}$$

Dans ce cas, le lemme 1.375 donne un réel t tel que $a < t < a + \epsilon < b$. Donc $t \in I$ et $a + \epsilon$ n'est pas un minorant de I. □

Exemple 1.401.
Quelques exemples dans les intervalles.

(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.

Dans les deux cas, le nombre 53 est un majorant. △

Il existe évidemment de nombreux exemples plus vicieux.

Exemple 1.402.
Prenons $E = \{\frac{1}{n}$ tel que $n \in \mathbb{N}_0\}$, dont les premiers points sont indiqués sur la figure 1.1. 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. △

FIGURE 1.1 – Les premiers points du type $x_n = 1/n$.

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 10.23).

148. Ici par « borne inférieure » nous entendons le a dans les intervalles du type $]a,b]$, $[a,b]$, etc.

Exemple 1.403.
Les premiers points de l'ensemble $F = \{\frac{(-1)^n}{n}$ tel que $n \in \mathbb{N}_0\}$ sont représentés à la figure 1.2. 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 1.2 – Les quelques premiers points du type $(-1)^n/n$.

1.20.6 Racines

Dans cette section, nous définissons $\sqrt{x}$ pour $x \in \mathbb{Q}^+$. Vous notez que c'est fait de façon assez algébrique [149], ou en tout cas, en restant proche des définitions. Des définitions plus technologiques utilisant la continuité de $x \mapsto x^n$ et qui prouvent que l'application est bijective sur un domaine choisi avec prudence existent (voir la définition 12.386). Il est même expliqué dans [61] que la méthode décrite ici permet de définir $\sqrt[n]{x}$ pour tout n entier, et pas seulement pour $n = 2$.

Proposition 1.404.
Soit $q \in \mathbb{Q}^+$. Il existe un unique $r \in \mathbb{R}$ tel que $r^2 = q$.

Plus précisément, en termes des notations de 1.353, pour tout $q \in \mathbb{Q}^+$, il existe un unique $r \in \mathbb{R}$ tel que $r^2 = \varphi(q)$.

Démonstration. En deux parties : d'abord l'existence et ensuite l'unicité.

(i) **Existence** Si $q = 0$, c'est $r = 0$. Nous supposons $q > 0$. La suite (x_k) de la proposition 1.348 a la propriété d'être de Cauchy dans $\mathbb{Q}$. Donc il existe un réel r qui est la classe de cette suite. Nous posons donc

$$r = \bar{x}. \tag{1.583}$$

En ce qui concerne r^2, nous avons, par définition du produit dans $\mathbb{R}$,

$$r^2 = \bar{x}^2 = \overline{(x_k^2)}, \tag{1.584}$$

c'est la classe de la suite de Cauchy donnée par les x_k^2. Posons $y_k = x_k^2$; la relation (1.584) s'écrit

$$r^2 = \bar{y}. \tag{1.585}$$

La proposition 1.348 nous dit également que y est une suite de Cauchy et que

$$y_k \xrightarrow{\mathbb{Q}} q \tag{1.586}$$

149. Discutable parce que des limites sont utilisées.

La proposition 1.387 donne alors $\bar{y} = \bar{q}$, et finalement

$$r^2 = \bar{q} = \varphi(q). \tag{1.587}$$

Ici tout n'est pas encore terminé avec l'existence parce qu'il faut nous assurer que $r \geqslant 0$. Ce n'est pas très compliqué : si $r < 0$, alors nous pouvons faire le choix $-r$ qui convient tout aussi bien : $(-r)^2 = r^2$.

(ii) **Unicité** Supposons $r_1, r_2 \in \mathbb{R}$ tels que $r_1^2 = r_2^2$. La proposition 1.365 dit que $\mathbb{R}$ est totalement ordonné ; disons pour fixer les idées que $r_1 \leqslant r_2$. Cela signifie, par définition de l'ordre sur $\mathbb{R}$, que $r_2 - r_1 \geqslant 0$. En posant $s = r_2 - r_1$ nous avons $r_2 = r_1 + s$. Passons au carré ; la distribution dans le calcul suivant provient du fait que $\mathbb{R}$ est un corps :

$$r_2^2 = (r_1 + s)^2 = r_1^2 + 2r_1 s + s^2. \tag{1.588}$$

Vu que $r_1^2 = q = r_2^2$, nous avons $2r_1 s + s^2 = 0$ ou encore

$$s(2r_1 + s) = 0. \tag{1.589}$$

Puisque $\mathbb{R}$ est un corps, c'est un anneau intègre [150] et la règle du produit nul s'applique : soit $s = 0$, soit $2r_2 + s = 0$. Puisque $r_2 > 0$ et que $s \geqslant 0$, nous avons $2r_2 + s > 0$ et donc $s = 0$. Nous en déduisons que $r_1 = r_2$.

□

1.20.7 Corps valué

Définition 1.405 (Valeur absolue, corps valué[67, 68])**.**
*Soit un corps $\mathbb{K}$. Une **valeur absolue** sur $\mathbb{K}$ est une application $|.|\colon \mathbb{K} \to \mathbb{R}^+$ telle que*

(1) $|x| = 0$ si et seulement si $x = 0$,

(2) $|x + y| \leqslant |x| + |y|$

(3) $|xy| \leqslant |x||y|$.

*Un corps muni d'une valeur absolue est un **corps valué**.*

Un corps valué sera un espace topologique métrique dans la définition 7.158. Dans le cas d'un corps totalement ordonné, nous avons une valeur absolue donnée par 1.318(2) et les principales propriétés dans le lemme 1.322.

1.20.8 Partie entière, partie fractionnaire

Lemme-Définition 1.406 ([69])**.**
Pour tout réel x, il existe un unique entier n vérifiant

$$n \leqslant x < n + 1. \tag{1.590}$$

Dans ce cas nous avons $x - n \in [0, 1[$.

*Le nombre n ainsi défini est la **partie entière** de x, et il sera noté $\operatorname{int}(x)$. Le nombre $x - \operatorname{int}(x)$ est la **partie fractionnaire** de x que nous notons $\operatorname{frac}(x)$.*

Il est toutefois à noter que la partie fractionnaire de x n'est pas garantie d'être une fraction ; cette dénomination est donc un peu trompeuse.

Lemme 1.407.
Si $\lambda \in \mathbb{R}$ n'est pas un entier, alors $\operatorname{int}(\lambda) + 2 > \lambda$.

150. Lemme 1.242.

1.21 Les complexes

La notion de module d'un nombre complexe $|z|$ sera donnée beaucoup plus tard, dans le lemme 10.88. La raison est que le module demande la racine carrée.

Définition 1.408 (Nombres complexes[70]).
L'ensemble des ***nombres complexes*** $\mathbb{C}$ *est l'ensemble* $\mathbb{R}^2$ *muni des opérations suivantes :*

(1)

$$\begin{aligned} \times_{\mathbb{C}} \colon \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ \big((x,y),(x',y')\big) &\mapsto (xx'-yy',xy'+yx') \end{aligned} \tag{1.591}$$

(2)

$$\begin{aligned} +_{\mathbb{C}} \colon \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ \big((x,y),(x',y')\big) &\mapsto (x+x',y+y') \end{aligned} \tag{1.592}$$

(3)

$$\begin{aligned} \cdot \colon \mathbb{R} \times \mathbb{C} &\to \mathbb{C} \\ \big(\lambda,(x,y)\big) &\mapsto (\lambda x,\lambda y). \end{aligned} \tag{1.593}$$

Lemme 1.409.
Le triplet $(\mathbb{C},+_{\mathbb{C}},\times_{\mathbb{C}})$ *est un anneau*[151] *commutatif dont le neutre pour l'addition est* $(0,0)$ *et le neutre pour la multiplication est* $(1,0)$.

Démonstration. Ce ne sont que des calculs. Juste pour vous montrer, voici la première partie pour l'associativité :

$$\big((a,b)(x,y)\big)(s,t) = (ax-by,ay+bx)(s,t) \tag{1.594a}$$
$$= (axs-bys-ayt-bxt,axt-byt+ays+bxs). \tag{1.594b}$$

Nous avons utilisé la distributivité sur $\mathbb{R}$, provenant du fait que $\mathbb{R}$ est un corps par le théorème 1.352. □

Lemme 1.410.
L'anneau $\mathbb{C}$ *est un corps.*

Démonstration. Il suffit de trouver un inverse pour chaque élément non nul. Soit un élément non nul $(a,b) \in \mathbb{C}$. En combinant les lemmes 1.381 et 1.382 nous savons que $a^2+b^2>0$. En particulier, cet élément est inversible dans $\mathbb{R}$, et nous pouvons considérer l'élément suivant de $\mathbb{C}$:

$$z = \big(\frac{a}{a^2+b^2},-\frac{b}{a^2+b^2}\big). \tag{1.595}$$

Prouver que $z(a,b)=(1,0)$ est maintenant juste un calcul. □

Lemme 1.411.
L'application

$$\begin{aligned} \varphi \colon \mathbb{R} &\to \mathbb{C} \\ x &\mapsto (x,0) \end{aligned} \tag{1.596}$$

est un morphisme d'anneaux[152].

Démonstration. Simples calculs. Par exemple

$$\varphi(xx') = (xx',0) = (x,0)(x',0) = \varphi(x)\varphi(x'). \tag{1.597}$$

□

151. Définition 1.38.
152. Définition 1.39.

1.412.
Admirez ...

— Un nombre complexe est un couple de réels.
— Un réel est une classe d'équivalence de suites de Cauchy de rationnels.
— Une suite de Cauchy de rationnels est une application $\mathbb{N} \to \mathbb{Q}$ vérifiant certaines propriétés.
— Un rationnel est une classe d'équivalences d'éléments de $\mathbb{Z}$.
— Un élément de $\mathbb{Z}$ est une classe d'équivalence de couples de naturels.
— Un naturel sera ...là c'est plus compliqué. Une construction vraiment rigoureuse des naturels risque d'être en dehors du cadre du Frido.

Bref, les objets que nous manipulons sont d'une effroyable complexité.

1.413.
À partir de maintenant, lorsque nous parlons de $\mathbb{R}$, nous parlons en réalité de $\varphi(\mathbb{R}) \subset \mathbb{C}$.

Lemme 1.414.
Nous avons $(0,1)^2 = (-1,0)$. *Nous notons* $i = (0,1)$.

Démonstration. Calcul direct à partir de la définition 1.591. □

La proposition suivante donne la forme cartésienne des nombres complexes. Pour la forme trigonométrique, il faudra attendre la proposition 18.58.

Proposition 1.415 ([1])**.**
L'application

$$\begin{aligned} f\colon \mathbb{R}^2 &\to \mathbb{C} \\ (a,b) &\mapsto a\varphi(1) + bi \end{aligned} \tag{1.598}$$

est un isomorphisme de $\mathbb{R}$*-module*[153].

Cette proposition permet d'écrire tout nombre complexe sous la forme $a + bi$ pour des réels a et b.

Définition 1.416.
Si $z = a + bi$ *est un nombre complexe (avec* $a, b \in \mathbb{R}$*), son* ***complexe conjugué*** *est le nombre* $a - bi$.

L'étude de la série géométrique est reportée à (beaucoup) plus tard, à la proposition 11.119. Dans l'immédiat il nous est possible de calculer la somme partielle.

Lemme 1.417 (Somme partielle de la série géométrique)**.**
Soit $q \in \mathbb{C}$. *Nous avons*

$$\sum_{n=0}^{N} q^n = \frac{1 - q^{N+1}}{1 - q}. \tag{1.599}$$

Démonstration. Posons $S_N = 1 + q + \ldots + q^N$. Nous avons évidemment $S_N - qS_N = 1 - q^{N+1}$ et donc

$$S_N = \sum_{n=0}^{N} q^n = \frac{1 - q^{N+1}}{1 - q}. \tag{1.600}$$

□

153. Module, définition 1.274.

Chapitre 2

Théorie des groupes

Pour rappel, la notion de groupe est définie en 1.34.

2.1 Groupes

2.2 Groupe dérivé

Définition 2.1.
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.*

L'élément neutre est toujours un commutateur : pour $g = h$, $[g, g] = ggg^{-1}g^{-1} = e$.

Définition 2.2.
Le ***groupe dérivé*** *de G est le sous-groupe noté $D(G)$ ou $[G, G]$ engendré*[1] *par les commutateurs.*

Autrement dit, $D(G)$ est l'intersection de tous les sous-groupes de G contenant tous les commutateurs. Le groupe $D(G)$ contient toujours au moins le neutre parce que c'est un groupe.

En vertu du lemme 1.267, 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

$$D(G) = \bigcup_{m=1}^{\infty} S_m. \tag{2.1}$$

Lemme 2.3.
Le groupe dérivé est un sous-groupe caractéristique[2]*, et un sous-groupe normal*[3]*.*

Démonstration. Il est évident que si $\alpha \in \mathrm{Aut}(G)$ alors

$$\alpha\big([g, h]\big) = \big[\alpha(g), \alpha(h)\big], \tag{2.2}$$

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,

$$x(ghg^{-1}h^{-1})x^{-1} = (xgx^{-1})(xhx^{-1})(xg^{-1}x^{-1})(xh^{-1}x^{-1}) \tag{2.3a}$$
$$= (xgx^{-1})(xhx^{-1})(xgx^{-1})^{-1}(xhx^{-1})^{-1}. \tag{2.3b}$$

□

Proposition 2.4.
Le groupe quotient $G/D(G)$ est abélien.

1. Définition 1.262.
2. Définition 1.167.
3. Définition 1.166.

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

$$[g][h] = [gh] = [ghh^{-1}g^{-1}hg] = [hg] = [h][g]. \tag{2.4}$$

□

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 morphisme 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.

2.3 Théorèmes d'isomorphismes

Définition 2.5.
*Soient un groupe G, un ensemble X et une application $f\colon X \to G$. Le **noyau** de f est la partie*

$$\ker(f) = \{x \in X \text{ tel que } f(x) = e\} \tag{2.5}$$

où e est l'élément neutre de G.

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 morphisme surjectif de noyau N.

Théorème 2.6 (Premier théorème d'isomorphisme)**.**
Soit $\theta\colon G \to H$ un morphisme de groupe. Alors

(1) $\ker(\theta)$ *est normal dans G,*

(2) $\operatorname{Image}\theta$ *est un sous-groupe de H*

(3) nous avons un isomorphisme

$$\frac{G}{\ker(\theta)} \simeq \operatorname{Image}\theta \tag{2.6}$$

Démonstration. Point par point.

(1) Le fait que $\ker(\theta)$ soit un sous-groupe de G est clair ; montrons qu'il est normal. Si $g \in G$ et $u \in \ker(\theta)$, alors $\theta(g^{-1}ug) = \theta(g^{-1})\theta(u)\theta(g) = \big(\theta(g)\big)^{-1}\theta(g) = 1_H$, et donc $g^{-1}ug \in \ker(\theta)$.

(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 $G/\ker(\theta)$, l'isomorphisme est donné par $\varphi([g]) = \theta(g)$.

□

Théorème 2.7 (Deuxième théorème d'isomorphisme)**.**
Soient H et N deux sous-groupes de G et supposons que N soit normal[4]*. 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

$$\frac{HN}{N} \simeq \frac{H}{H \cap N}. \tag{2.7}$$

Démonstration. Point par point.

4. Si N n'est pas normal, il y aura la proposition 2.8.

(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,

$$nhn'h' = n\underbrace{hn'h^{-1}}_{\in N}hh' \in NH. \tag{2.8}$$

Cela prouve que NH est un groupe.

De la même façon, nous prouvons que HN est un groupe par

$$hnh'n' = hh'\underbrace{h'^{-1}nh'}_{\in N}n' \in HN \tag{2.9}$$

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

$$\begin{aligned} j\colon H &\to HN \\ h &\mapsto h \end{aligned} \tag{2.10}$$

et la surjection canonique

$$\sigma\colon HN \to HN/N \tag{2.11}$$

Nous considérons ensuite l'application composée

$$\begin{aligned} f\colon H &\to HN/N \\ h &\mapsto hN. \end{aligned} \tag{2.12}$$

(i) <u>f **est surjective**</u> L'application f est surjective parce que l'élément $hnN \in HN/N$ est l'image de h, étant donné que $hnN = hN$.

(ii) <u>$\ker(f) = H \cap N$</u> Si $a \in H \cap N$, nous avons $f(a) = aN = N$, et donc $H \cap N \subset \ker(f)$. D'autre part, si $h \in H$ vérifie $h \in \ker(f)$, alors $f(h) = hN = N$, ce qui est uniquement possible lorsque $h \in N$.

Le premier théorème d'isomorphisme implique alors que $H/\ker(f) \simeq \operatorname{Image} f$, c'est-à-dire

$$H/N \cap H \simeq HN/N. \tag{2.13}$$

□

Proposition 2.8 (Deuxième théorème d'isomorphisme (suite))**.**
Soient N et H des sous-groupes de G. Si H normalise N, c'est à dire si $hNh^{-1} \in N$ pour tout $h \in H$, alors nous avons l'isomorphisme

$$\frac{HN}{N} \simeq \frac{H}{H \cap N}. \tag{2.14}$$

Théorème 2.9 (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

$$(G/M)/(N/M) \simeq G/N. \tag{2.15}$$

Démonstration. Afin de montrer que N/M est normal dans G/M, nous considérons $g \in G$, $nM \in N/M$ et nous calculons

$$gnMg^{-1} = gng^{-1}\underbrace{gMg^{-1}}_{=M} = \underbrace{gng^{-1}}_{\in N} M \in N/M. \tag{2.16}$$

Pour prouver l'isomorphisme nous considérons le morphisme

$$\begin{aligned} \varphi\colon G/M &\to G/N \\ gM &\mapsto gN. \end{aligned} \tag{2.17}$$

Ce morphisme est surjectif et son 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

$$(G/M)/\ker(\varphi) \simeq \operatorname{Image}\varphi, \tag{2.18}$$

c'est-à-dire

$$(G/M)/(N/M) \simeq G/N. \tag{2.19}$$

□

2.4 Indice d'un sous-groupe et ordre des éléments

Lemme 2.10.
Lorsque H est normal dans G, alors la définition

$$[a] \cdot [b] = [ab] \tag{2.20}$$

définit une loi de groupe sur l'ensemble G/H.

Démonstration. Le neutre est $[e]$ et l'associativité ne pose pas plus de problème que l'existence d'un inverse. Le point à vérifier est que la formule (2.20) est une bonne définition : $[ah] \cdot [bh'] = [ab]$ pour tout $h, h' \in H$. Nous avons :

$$[ah] \cdot [ah'] = [ahah'] = [ahb]. \tag{2.21}$$

Pour montrer que c'est $[ab]$, l'astuce est d'introduire bb^{-1} à côté du a :

$$[ahb] = [abb^{-1}hb] = [ab] \tag{2.22}$$

parce que $b^{-1}hb \in H$ du fait que H soit normal dans G. □

Exemple 2.11 ([71])**.**
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

$$\begin{aligned} f\colon G &\to G/H \\ g &\mapsto [g] \end{aligned} \tag{2.23}$$

est un morphisme de groupes. △

Définition 2.12.
*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 6.57).

Théorème 2.13 (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

$$|G/H| = \frac{|G|}{|H|}. \tag{2.24}$$

Démonstration. Nous commençons par montrer que les classes de H ont toutes le même nombre d'éléments que H. En effet pour chaque $g \in G$ nous avons la bijection

$$\begin{aligned} \varphi\colon H &\to gH \\ h &\mapsto gh. \end{aligned} \tag{2.25}$$

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 :

$$|G| = |H| \cdot \text{ nombre de classes.} \tag{2.26}$$

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. □

Corolaire 2.14.
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

$$H = \{g^k \text{ tel que } k \in \mathbb{N}\}. \tag{2.27}$$

Par le théorème de Lagrange 2.13, 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. □

D'autres résultats à propos d'ordres et d'indices de groupes finis dans la proposition 3.31 et le lemme 3.33. En particulier le théorème de Cauchy 3.27 qui dit : si p divise l'ordre du groupe G, alors G contient au moins un élément d'ordre p.

2.5 Suite de composition

Définition 2.15 (Suite de composition)**.**
Soit un groupe G.

*(1) Une **suite de composition** dans G est une suite finie de sous-groupes $(G_i)_{i=0,\dots,n}$ telle que*

$$\{e\} = G_n \subseteq G_{n-1} \subseteq \dots \subseteq G_1 \subseteq G_0 = G \tag{2.28}$$

et telle que G_{i+1} est normal[5] *dans G_i.*

5. Nous rappelons au cas où, que « normal » signifie « distingué ».

(2) Les groupes G_i/G_{i+1} sont les ***quotients*** *de la suite de composition.*

(3) 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 2.21).

Lemme 2.16 (du papillon ou de Zassenhaus[72])**.**
Soient G un groupe et des sous-groupes A et B. Soient 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

$$\frac{A'(A\cap B)}{A'(A\cap B')}\simeq\frac{(A\cap B)B'}{(A'\cap B)B'}\simeq\frac{B'(A\cap B)}{B'(A'\cap B)}. \tag{2.29}$$

Démonstration. Nous n'allons pas démontrer chacun des points ; pour plus de détails, nous dirons simplement que « la preuve est très similaire dans les autres cas ».

Commençons par montrer que $A'(A\cap B')$ est un groupe. Si $a,b\in A'$ et $x,y\in A\cap B'$,

$$axby=xx^{-1}axbx^{-1}xy \tag{2.30}$$

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

$$x^{-1}a^{-1}=\underbrace{x^{-1}a^{-1}x}_{\in A'}x^{-1}. \tag{2.31}$$

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

$$\begin{aligned}(bf)^{-1}(ax)(bf)&=(bf)^{-1}(a\underbrace{xbx^{-1}}_{=c\in A'}xf) &(2.32a)\\ &=f^{-1}b^{-1}acxf &(2.32b)\\ &=f^{-1}b^{-1}acf\underbrace{f^{-1}xf}_{=y\in A\cap B'} &(2.32c)\\ &=\underbrace{f^{-1}b^{-1}acf}_{\in A'}y &(2.32d)\\ &\in A'(A\cap B'). &(2.32e)\end{aligned}$$

Pour prouver l'isomorphisme

$$\frac{A'(A\cap B)}{A'(A\cap B')}\simeq\frac{(A\cap B)B'}{(A'\cap B)B'}, \tag{2.33}$$

nous allons utiliser le deuxième théorème d'isomorphisme (2.8) 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

$$\frac{A'(A\cap B')(A\cap B)}{A'(A\cap B')}\simeq\frac{A\cap B}{A\cap B\cap A'(A\cap B')}. \tag{2.34}$$

Pour simplifier un peu cette expression nous prouvons d'abord que

$$(A\cap B)\cap A'(A\cap B')=(A'\cap B)(A\cap B'). \tag{2.35}$$

L'inclusion $\supset$ est facile. Pour l'autre sens, étant donné que $A'(A \cap B') \subset A$ nous avons

$$A \cap B \cap A'(A \cap B) = B \cap A'(A \cap B). \tag{2.36}$$

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 $x^{-1} \in A \cap B$. Par conséquent $a \in B$ et donc $a \in A' \cap B$. Donc un élément de $B \cap A'(A \cap B)$ s'écrit sous la forme ax avec $a \in A' \cap B$ et $x \in A \cap B$. Autrement dit

$$B \cap A'(A \cap B) \subset (A' \cap B)(A \cap B) \tag{2.37}$$

et nous avons

$$(A \cap B) \cap A'(A \cap B) = B \cap A'(A \cap B) \subset (A' \cap B)(A \cap B'), \tag{2.38}$$

et donc l'égalité (2.35). Toujours dans l'idée de simplifier (2.34) 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

$$\frac{A'(A \cap B)}{A'(A \cap B')} = \frac{A \cap B}{(A' \cap B)(A \cap B')}. \tag{2.39}$$

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

$$\frac{B'(A \cap B)}{B'(A' \cap B)} = \frac{A'(A \cap B)}{A'(A \cap B')}. \tag{2.40}$$

Nous devons encore justifier $B'(A \cap B) = (A \cap B)B'$ et $B'(A' \cap B) = (A' \cap B)B'$. Vérifions la première égalité, et laissons la seconde au lecteur. Si $b \in B'$ et $x \in A \cap B$, alors

$$bx = x \underbrace{x^{-1}bx}_{\in B'} \in (A \cap B)B'. \tag{2.41}$$

□

Proposition 2.17.
Si G est un groupe fini et si (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 2.13 s'applique et, à chaque pas de la suite de composition, nous avons :

$$\left| \frac{G_i}{G_{i+1}} \right| = \frac{|G_i|}{|G_{i+1}|}. \tag{2.42}$$

Il suffit maintenant d'écrire $|G|$ de façon télescopique :

$$|G| = \prod_{0 \leqslant i \leqslant n-1} \frac{|G_i|}{|G_{i+1}|} \tag{2.43}$$

□

Définition 2.18.
Nous disons que les deux suites de composition $(G_i)_{0 \leqslant i \leqslant r}$ et $(H_j)_{0 \leqslant j \leqslant s}$ sont ***équivalentes*** *si $r = s$ et si il existe une permutation $\sigma \in S_{r-1}$ telle que*

$$\frac{G_i}{G_{i+1}} \simeq \frac{H_{\sigma(i)}}{H_{\sigma(i)+1}}. \tag{2.44}$$

Proposition 2.19 (Schreider)**.**
Deux suites de composition d'un même groupe admettent des raffinements équivalents.

Démonstration. Soient les suites de composition

$$\{e\} = G_m \subseteq \ldots \subseteq G_1 \subseteq G_0 = G \tag{2.45a}$$
$$\{e\} = H_m \subseteq \ldots \subseteq H_1 \subseteq H_0 = G \tag{2.45b}$$

Nous raffinons la suite (G_i) en remplaçant $G_{i+1} \subseteq G_i$ par

$$G_{i+1} = G_{i+1}(G_i \cap H_n) \subset G_{i+1}(G_i \cap H_{n-1}) \subseteq \ldots \subseteq G_{i+1}(G_i \cap H_0) = G_i, \tag{2.46}$$

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 2.16 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 des 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

$$\frac{G_{i+1}(G_i \cap H_k)}{G_{i+1}(G_i \cap H_{k+1})} \simeq \frac{H_{k+1}(H_k \cap G_i)}{H_{k+1}(H_k \cap G_{i+1})} \tag{2.47}$$

en vertu du lemme du papillon (2.16). Le membre de droite de (2.47) est un des quotients du raffinement de (H_j). □

Lemme 2.20 (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 2.21 (Jordan-Hölder)**.**
À propos de suites de Jordan-Hölder dans un groupe fini.

(1) Tout groupe fini admet une suite de Jordan-Hölder.

(2) Toutes les suites de Jordan-Hölder dans un groupe fini sont équivalentes.

Démonstration. En deux parties.

(i) **Pour (1)**

(ii) **Pour (2)** 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 2.20. 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.

□

2.6 Groupes résolubles

Définition 2.22.
*Le groupe G est **résoluble** si il existe une suite finie de sous-groupes G_i*

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{2.48}$$

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[6] dont les quotients sont abéliens.

Lemme 2.23 ([73]).
Soit G un groupe et H un sous-groupe normal. Le groupe G/H est abélien si et seulement si[7] $D(G) \subset H$.

Démonstration. Les propositions suivantes sont équivalentes :

— Le groupe G/H est abélien
— pour tout $x, y \in G$, $[x][y] = [y][x]$
— $[x][y][x^{-1}][y^{-1}] = [e]$
— $[xyx^{-1}y^{-1}] = [e]$
— $[x, y] \in H$, voir la définition 2.1.
— $D(G) \subset H$.

□

Proposition 2.24 ([73]).
Un groupe est résoluble si et seulement si sa suite dérivée termine sur $\{e\}$.

Démonstration. Grâce au lemme 2.3 et à la proposition 2.4, si la suite dérivée termine sur $\{e\}$ alors la suite dérivée est une suite qui répond aux conditions de la définition 2.22 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

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{2.49}$$

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 :

$$D^i(G) \subset G_i \tag{2.50}$$

Par le lemme 2.23 nous avons aussi

$$D(G_i) \subset G_{i+1}. \tag{2.51}$$

En dérivant (2.50) et en tenant compte de (2.51), $D^{i+1}(G) \subset D(G_i) \subset G_{i+1}$. Donc par récurrence nous avons bien $D^k(G) \subset G_k$ pour tout k. En particulier pour $k = n$ nous avons $D^n(G) = G_n = \{e\}$. □

Proposition 2.25.
Soient des groupes G et H. Nous supposons que G est résoluble et nous considérons un morphisme $\psi\colon G \to H$. Alors $\psi(G)$ est résoluble.

Démonstration. Puisque G est résoluble, il existe une suite de sous-groupes G_i tels que

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{2.52}$$

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

$$\{e\} = \psi(G)_n \subset \psi(G)_{n-1} \subset \ldots \subset \psi(G)_1 \subset \psi(G)_0 = \psi(G). \tag{2.53}$$

La proposition 1.178 nous indique que $\psi(G)_i$ est normal dans $\psi(G)_{i+1}$, et que $\psi(G)_i/\psi(G)_{i+1}$ est abélien. □

6. Voir définition 2.15.
7. Ici $D(G)$ est le groupe dérivé de G, voir 2.2.

2.7 Action de groupes

Le concept d'action d'un groupe est donné par la définition 1.311

Lemme 2.26.
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 membres nous trouvons

$$(\phi_{g^{-1}}\phi_g)(x) = (\phi_{g^{-1}}\phi_g)(y), \tag{2.54}$$

ce qui donne $x = y$ parce que $\phi_{g^{-1}}\phi_g = \phi_{g^{-1}g} = \phi_e = \mathrm{Id}$.

Les trois dernières égalités écrites disent que $\phi_{g^{-1}}$ est l'inverse[8] 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 2.27 (Orbite)**.**
*Si G agit sur un ensemble E, nous notons $G \cdot x$ l'**orbite** de $x \in E$ sous l'action de G :*

$$G \cdot x = \{gx \text{ tel que } g \in G\}.$$

Définition 2.28 (Stabilisateur)**.**
*Si G est un groupe agissant sur l'ensemble E, et si $x \in E$, nous notons G_x ou $\mathrm{Stab}(x)$ le **stabilisateur** de x :*

$$\mathrm{Stab}(x) = G_x = \{g \in G \text{ tel que } g \cdot x = x\}. \tag{2.55}$$

Définition 2.29 (Fixateur)**.**
*Si G est un groupe agissant sur l'ensemble E, et si $g \in G$, nous notons enfin $\mathrm{Fix}(g)$ le **fixateur** de g :*

$$\mathrm{Fix}(g) = \{x \in E \text{ tel que } g \cdot x = x\}. \tag{2.56}$$

Définition 2.30.
*L'action de G sur E est **fidèle** si l'élément neutre est le seul élément de G à fixer tous les points de E, c'est-à-dire si*

$$(gx = x \quad \forall x \in E) \Rightarrow g = e. \tag{2.57}$$

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 23.95.

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 2.31.
*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

$$x \sim_g y \Leftrightarrow xh = y \tag{2.58}$$

8. Si vous décidez de dire ça à un jury dans un concours, soyez prêts à préciser les domaines.

pour un certain $h \in H$. Ensuite

$$x \sim_d y \Leftrightarrow hx = y \tag{2.59}$$

pour un certain $h \in H$.

Le lemme suivant est une généralisation du théorème de Lagrange 2.13.

Lemme 2.32.
L'ensemble $(G/H)_g$ est fini si et seulement si l'ensemble $(G/H)_d$ est fini. Si 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

$$\begin{aligned} f\colon (G/H)_g &\to (G/H)_d \\ [x]_g &\mapsto [x^{-1}]_d \end{aligned} \tag{2.60}$$

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

$$f([x]_g) = f([y]_g). \tag{2.61}$$

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

$$\begin{aligned} f\colon [x]_g &\to [y]_g \\ xh &\mapsto yh. \end{aligned} \tag{2.62}$$

(3) Le nombre d'éléments dans une classe est égal à $|H|$ par la bijection

$$\begin{aligned} g\colon [x]_g &\to H \\ xh &\mapsto h. \end{aligned} \tag{2.63}$$

Par conséquent

$$|G| = |H| \cdot \text{nombre de classes} = |H| \cdot \operatorname{Card}\big((G/H)_g\big), \tag{2.64}$$

et nous avons bien

$$\operatorname{Card}\big((G/H)_g\big) = \frac{|G|}{|H|} = |G : H|. \tag{2.65}$$

□

Proposition 2.33 (Orbite-stabilisateur[40])**.**
Soient 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 $\operatorname{Fix}(x)$ est finie si et seulement si $\operatorname{Fix}(x)$ est d'indice fini dans G. Dans ce cas nous avons

$$\operatorname{Card}(G \cdot x) = |G : \operatorname{Fix}(x)|. \tag{2.66}$$

Une autre façon d'écrire la même formule :

$$|G| = |\operatorname{Fix}(x)||\mathcal{O}_x|. \tag{2.67}$$

C'est la formule (2.66) qui est nommée formule des classes sur wikipédia.

Démonstration. En deux points.

(1) Soit l'application

$$\begin{aligned} \psi\colon G\cdot x &\to G/G_x \\ a\cdot x &\mapsto [a]. \end{aligned} \tag{2.68}$$

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 $\operatorname{Fix}(x)$, par conséquent $|A_y| = |\operatorname{Fix}(x)|$ pour tout $y\in\mathcal{O}_x$. Les A_y pour différents y sont disjoints et nous avons de plus

$$\bigcup_{y\in\mathcal{O}_x} A_y = G. \tag{2.69}$$

Les ensembles A_y divisent donc G en $|\mathcal{O}_x|$ paquets de $|\operatorname{Fix}(x)|$ éléments. D'où la formule (2.67).

□

Corolaire 2.34.
Soit C_g la classe de conjugaison d'un élément g du groupe fini G. Alors

$$\operatorname{Card}(C_g) = |G : Z_G(g)| \tag{2.70}$$

où $Z_G(g)$ est le centralisateur de g dans G[9] *de G.*

Démonstration. C'est une application de la proposition 2.33 (formule (2.66)) 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 :

$$\begin{aligned} \operatorname{Fix}(g) &= \{h\in G \text{ tel que } h\cdot g = g\} && (2.71a)\\ &= \{h\in G \text{ tel que } hgh^{-1} = g\} && (2.71b)\\ &= \{h\in G \text{ tel que } gh = hg\} && (2.71c)\\ &= Z_G(g). && (2.71d) \end{aligned}$$

Donc la formule $\operatorname{Card}(G\cdot g) = |G : G_g|$ devient, dans le cas de l'action adjointe de G sur lui-même : $\operatorname{Card}(C_g) = |G : Z_G(g)|$. □

Lemme 2.35.
Soit G un groupe agissant sur l'ensemble E. On définit $x\sim x'$ si et seulement si il existe $g\in G$ tel que $g\cdot x = x'$. Alors

(1) la relation $\sim$ est une relation d'équivalence.

(2) la classe $[x]$ est l'orbite $\mathcal{O}_x$ de x sous G.

Corolaire 2.36 (Équation des orbites)**.**
Soient 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) $\operatorname{Card}(E) = \sum_i \operatorname{Card}(\mathcal{O}_i)$.

Définition 2.37.
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.*

9. Définition 1.164.

Autrement dit, les images des éléments d'un domaine fondamental F forment une partition de l'ensemble :

$$E = \bigsqcup_{g \in G} g(F) \tag{2.72}$$

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 2.38 (Équation des classes[74])**.**
Soit G, un groupe fini opérant sur un ensemble E. Si E'' est un ensemble contenant exactement un élément de chaque orbite dans $E \backslash \operatorname{Fix}_G(E)$, alors

$$|G| = |\operatorname{Fix}_G(E)| + \sum_{x \in E''} \frac{|G|}{|\operatorname{Fix}_G(x)|}. \tag{2.73}$$

Si de plus G est un p-groupe, alors

$$|E| = |\operatorname{Fix}_G(E)| \mod p. \tag{2.74}$$

Démonstration. Par le corolaire 2.36, 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,

$$|G| = \operatorname{Card}(\operatorname{Fix}_G(E)) + \sum_{x \in E''} \frac{|G|}{|\operatorname{Fix}_G(x)|} \tag{2.75}$$

où nous avons utilisé le fait que $|G| = |\operatorname{Fix}_G(x)||\mathcal{O}_x|$.

Si G est un p-groupe alors si $x \in E''$, $\operatorname{Fix}_G(x)$ est un sous-groupe propre de G et donc $|\operatorname{Fix}_G(x)|$ est un diviseur propre de $|G|$. Du coup la fraction $|G|/|\operatorname{Fix}_G(x)|$ est une puissance non nulle de p et l'équation (2.73) devient immédiatement (2.74). □

Corolaire 2.39 (Équation des classes)**.**
Soient G un groupe, et $C_1, \ldots, C_l$ la liste de ses classes de conjugaison contenant plus d'un élément. Alors

$$\operatorname{Card}(G) = \operatorname{Card}\big(Z(G)\big) + \sum_i |G : Z_{g_i}| = \operatorname{Card}\big(Z(G)\big) + \sum_i \frac{\operatorname{Card}(G)}{\operatorname{Card}\big(\operatorname{Fix}(g_i)\big)} \tag{2.76}$$

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 qu'un seul élément sont celles des éléments de $Z(G)$. En ce qui concerne les autres orbites, $\operatorname{Card}(C_{g_i}) = |G : Z_{g_i}|$ par le théorème orbite-stabilisateur (proposition 2.33). □

Théorème 2.40 (Formule de Burnside)**.**
Si G est un groupe fini agissant sur l'ensemble fini E et si Ω est l'ensemble des orbites, alors

$$\operatorname{Card}(\Omega) = \frac{1}{|G|} \sum_{g \in G} \operatorname{Card}\big(\operatorname{Fix}(g)\big). \tag{2.77}$$

Démonstration. Nous considérons l'ensemble

$$A = \{(g, x) \in G \times E \text{ tel que } gx = x\}, \tag{2.78}$$

et nous en calculons le cardinal de deux façons. D'abord

$$\begin{aligned}
\mathrm{Card}(A) &= \sum_{x\in E} \mathrm{Card}\{g\in G \text{ tel que } gx=x\} && (2.79a)\\
&= \sum_{x\in E} \mathrm{Card}(\mathrm{Fix}(x)) && (2.79b)\\
&= \sum_{\omega\in\Omega}\sum_{x\in\omega} \mathrm{Card}(\mathrm{Fix}(x)) && (2.79c)\\
&= \sum_{\omega\in\Omega} \frac{|G|}{\mathrm{Card}(\omega)} && (2.79d)\\
&= |G|. && (2.79e)
\end{aligned}$$

Pour obtenir (2.79d) nous avons utilisé l'équation des classes (2.67). L'autre façon de calculer $\mathrm{Card}(A)$ est de regrouper ainsi :

$$\mathrm{Card}(A) = \sum_{g\in G} \mathrm{Card}\{x\in E \text{ tel que } gx=x\} = \sum_{g\in G} \mathrm{Card}(\mathrm{Fix}(g)). \tag{2.80}$$

En égalisant les deux expressions de $\mathrm{Card}(A)$ nous trouvons

$$|G|\,\mathrm{Card}(\Omega) = \sum_{g\in G} \mathrm{Card}(\mathrm{Fix}(g)). \tag{2.81}$$

□

Proposition 2.41.
Soient 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.

Proposition 2.42 ([75])**.**
L'indice du centre d'un groupe n'est jamais un nombre premier.

Lemme 2.43 ([76])**.**
Soient un groupe cyclique G et un sous groupe H.

(1) Il existe $r, n\in\mathbb{N}$ tels que $r\mid n$ et $H\simeq r\mathbb{Z}/n\mathbb{Z}$.

(2) $\mathrm{Card}(H) = n/r$

(3) H est cyclique.

En particulier, tout sous-groupe d'un groupe cyclique est cyclique.

Démonstration. En plusieurs parties.

(i) **Pour (1)** Soit un générateur a de G. L'application

$$\begin{aligned} f\colon \mathbb{Z} &\to G\\ k &\mapsto a^k \end{aligned} \tag{2.82}$$

est un morphisme de groupe surjectif.

La partie $f^{-1}(H)$ est un sous-groupe de $\mathbb{Z}$. La proposition 1.234(2) dit qu'il existe $r\in\mathbb{N}$ tel que $f^{-1}(H) = r\mathbb{Z}$. De même $\ker(f)$ est un sous-groupe de $\mathbb{Z}$, de telle sorte qu'il existe $n\in\mathbb{N}$ tel que $\ker(f) = n\mathbb{Z}$.

Nous avons

$$n\mathbb{Z} = \ker(f) = f^{-1}(e) \subset f^{-1}(H) = r\mathbb{Z}. \tag{2.83}$$

Étant donné que $n\mathbb{Z}\subset r\mathbb{Z}$, il existe $d\in\mathbb{N}$ tel que $n = dr$.

Nous considérons maintenant

$$\begin{aligned} g\colon r\mathbb{Z} &\to H \\ k &\mapsto a^k, \end{aligned} \tag{2.84}$$

qui est encore un morphisme de groupes. En utilisant le premier théorème d'isomorphismes 2.6(3), nous avons

$$\frac{r\mathbb{Z}}{\ker(f)} \simeq \text{Image}(g). \tag{2.85}$$

Mais $\ker(g) = n\mathbb{Z}$ et $\text{Image}(g) = H$. Donc nous avons bien $H \simeq r\mathbb{Z}/n\mathbb{Z}$.

(ii) **Pour (2)** En ce qui concerne le cardinal, vu que nous avons un isompshisme,

$$\text{Card}(H) = \text{Card}(r\mathbb{Z}/n\mathbb{Z}) = n/r \tag{2.86}$$

par le lemme 1.273(3).

(iii) **Pour (3)** Le fait que H soit cyclique est le lemme 1.273(2).

□

Théorème 2.44 (Sous-groupe d'un groupe cyclique[76])**.**
Soit G un groupe cyclique[10] *d'ordre n. Pour chaque $d \in \mathbb{N}$ nous posons*

$$H_d = \{x \in G \text{ tel que } x^d = e\}. \tag{2.87}$$

(1) Ces H_d sont des sous-groupes de G.

(2) $\text{Card}(H_d) = d$.

(3) L'ensemble des sous-groupes de G est $\{H_d \text{ tel que } d \mid n\}$.

(4) Si g est un générateur de G, alors H_d peut être décrit des façons suivantes :

$$H_d = \{x \in G \text{ tel que } x^d = e\} = \text{gr}(g^{n/d}). \tag{2.88}$$

Démonstration. Point par point.

(i) **H_d est un sous-groupe** Si $g, h \in H_d$, alors $(gh)^d = g^d h^d = ee = e$. De plus $(g^{-1})^d = (g^d)^{-1} = e^{-1} = e$.

(ii) **Tout sous-groupe est un H_d avec $d \mid n$** Soit un sous-groupe H de G. Nous notons $\text{Card}(H) = d$. Le lemme 2.43(3) nous indique que H est cyclique. Soit un générateur $a \in H$; son ordre est le cardinal de H : $\text{ord}(a) = \text{Card}(H) = d$.

Tout élément de H vérifie $x^{\text{ord}(a)} = (a^k)^{\text{ord}(a)} = (a^{\text{ord}(a)})^k = e$. Donc $H \subset H_d$. En particulier $\text{Card}(H) \leqslant \text{Card}(H_d)$.

Tout sous-groupe de G est cyclique. En particulier H_d est cyclique, et nous pouvons considérer un générateur : il existe $a \in H_d$ tel que $\text{ord}(a) = \text{Card}(H_d)$. Étant donné que $a \in H_d$, il vérifie $a^d = e$, et donc

$$\text{ord}(a) \mid d. \tag{2.89}$$

Vu que $\text{ord}(a) = \text{Card}(H_d)$ et que $d = \text{Card}(H)$, nous avons $\text{Card}(H_d) \mid \text{Card}(H)$ et en particulier $\text{Card}(H_d) \leqslant \text{Card}(H)$. Donc $\text{Card}(H) = \text{Card}(H_d)$ et au final $H = H_d$.

Nous avons prouvé au passage que $\text{Card}(H_d) = d$. Vu que le cardinal d'un sous-groupe divise le cardinal du groupe[11], nous avons $d \mid n$.

(iii) $\text{Card}(H_d) = d$ C'est contenu dans ce que nous avons fait dans la partie ii.

(iv) **Si $d \mid n$ alors H_d est un sous-groupe de G** C'est un cas particulier de la partie i.

10. Définition 1.270.
11. Théorème de Lagrange 2.13(1).

(v) **Note** Le point (3) est prouvé. Cela ne contredit pas le fait que H_d soit un sous-groupe de G pour tout d, et pas seulement pour les d qui divisent n. En effet si d ne divise pas n, il existe d' divisant n tel que $H_d = H_{d'}$. Pas besoin de chercher une preuve de ce fait : la preuve est déjà faite. Il s'agit de dire que H_d est un sous-groupe de G pour tout d (partie i), et que tous les sous-groupes de G sont de la forme H_d avec $d \mid n$.

(vi) **Pour (4)** Soit un générateur g de G. Nous avons d'une part $g^{n/d} \in H_d$ parce que $(g^{n/d})^d = g^n = e$. D'autre part, les éléments

$$g^{n/d}, g^{2n/d}, \dots, g^{(d-1)n/d}, g^n = e \tag{2.90}$$

sont distincts. En effet si $g^{kn/d} = g^{ln/d}$ avec $0 < l < k < d$, alors $g^{(k-l)n/d} = e$. Comme $k - l < d$, nous avons $(k-l)n/d < n$ et donc $g^{(k-l)n/d} \neq e$ parce que g est générateur.

Bref, les $g^{kn/d}$ avec $k = 1, \dots, d$ sont d éléments distincts, tous dans H_d dont le cardinal est d. Donc

$$H_d = \{g^{kn/d}\}_{1,\dots,d}. \tag{2.91}$$

□

Définition 2.45.
*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 $g = g'$.*

2.8 Produit semi-direct de groupes

Définition 2.46.
*Une **suite exacte** est une suite d'applications comme suit :*

$$\dots \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \xrightarrow{f_{i+2}} \dots \tag{2.92}$$

où pour chaque i, les applications f_i et f_{i+1} vérifient $\ker(f_{i+1}) = \mathrm{Image}(f_i)$. Lorsque les ensembles A_i sont des groupes, alors nous demandons de plus que les f_i soient des morphismes.

Très souvent nous sommes confrontés à des suites exactes de la forme

$$1 \longrightarrow A \xrightarrow{f} G \xrightarrow{g} B \longrightarrow 1 \tag{2.93}$$

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 2.47.
*Soient N et H deux groupes et un morphisme de groupes $\phi\colon H \to \mathrm{Aut}(N)$. 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)*

$$(n, h) \cdot (n', h') = (n\phi_h(n'), hh'). \tag{2.94}$$

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 2.31.

Le théorème suivant permet de reconnaitre un produit semi-direct lorsqu'on en voit un.

Théorème 2.48 ([9]).
Soit une suite exacte de groupes

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{s} H \longrightarrow 1 \tag{2.95}$$

Si il existe un sous-groupe $\tilde{H}$ de G à partir duquel s est un isomorphisme, alors

$$G \simeq i(N) \times_\sigma \tilde{H} \tag{2.96}$$

où σ est l'action adjointe[12] *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

$$\begin{aligned} \phi\colon G &\to \tilde{N} \times \tilde{H} \\ nh &\mapsto (n, h) \end{aligned} \tag{2.97}$$

est une bijection. C'est une conséquence des points (3) et (4).

(6) Si sur $\tilde{N} \times \tilde{H}$ nous mettons le produit

$$(n, h) \cdot (n', h') = (n\sigma_h n', hh') \tag{2.98}$$

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

$$\begin{aligned} \phi(nhn'h') &= \phi(n \underbrace{hn'h^{-1}}_{\in \tilde{N}} hh') && \text{(2.99a)} \\ &= \phi\big((nhn'h^{-1})(hh')\big) && \text{(2.99b)} \\ &= (nhn'h^{-1}, hh') && \text{(2.99c)} \\ &= (n, h) \cdot (n', h') && \text{(2.99d)} \\ &= \phi(nh)\phi(n'h'). && \text{(2.99e)} \end{aligned}$$

□

Corolaire 2.49.
Soit G un groupe, et N, H des sous-groupes de G tels que

(1) H normalise N (c'est-à-dire que $\sigma_h(n) = hnh^{-1} \in N$ pour tout $h \in H$ et $n \in N$[13]*),*

(2) $H \cap N = \{e\}$,

(3) $HN = G$.

12. Le fait que H agisse sur $i(N)$ fait partie du théorème.
13. Ou encore que H agit sur N par automorphismes internes.

Alors l'application

$$\begin{aligned}\psi\colon N\times_\sigma H &\to G\\ (n,h)&\mapsto nh\end{aligned}\tag{2.100}$$

est un isomorphisme de groupes.

Démonstration. En plusieurs parties.

(i) **ψ est un morphisme** Simple calcul en utilisant la formule (2.94) du produit dans $N\times_\sigma H$:

$$\begin{aligned}\psi\big((n,h)(n',h')\big)&=\psi\big(n\sigma_h(n'),hh'\big)&(2.101a)\\ &=nhn'h^{-1}hh'&(2.101b)\\ &=nhn'h'&(2.101c)\\ &=\psi(n,h)\psi(n',h').&(2.101d)\end{aligned}$$

(ii) **ψ est injective** Supposons que $\psi(n,h)=\psi(n',h')$. Alors $nh=n'h'$ et

$$n=n'h'h^{-1}.\tag{2.102}$$

En multipliant par $(n')^{-1}$ nous trouvons

$$h'h^{-1}=(n')^{-1}n\in N.\tag{2.103}$$

Donc $h'h^{-1}\in H\cap N=\{e\}$. Nous en déduisons que $h=h'$ et donc que $n=n'$.

(iii) **ψ est surjective** Nous savons que $HN=G$. Soit $g\in G$. Il existe $h\in H$ et $n\in N$ tel que $g^{-1}=h^{-1}n^{-1}$. Avec ces éléments nous avons

$$g=nh=\psi(n,h).\tag{2.104}$$

□

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.

2.9 Groupe de torsion

Soit G un groupe. Un élément $g\in G$ est un **élément de torsion** si 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 2.50.
Le groupe additif $\mathbb{Q}/\mathbb{Z}$ est un groupe de torsion parce que si $[x]=[p/q]$, alors $q[x]=[p]=[0]$. △

2.10 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 $\{i\in I$ 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 $I=\mathbb{N}$.

Chapitre 3

Anneaux

Attention aux conventions. Dans le Frido, un corps peut être réduit à $\{0\}$ et un idéal premier ne peut pas être $\{0\}$. Ces conventions ont une série de conséquences un peu partout, par exemple dans la proposition 3.78 où nous parlons d'idéal maximum propre. Comparez par exemple avec [77]. Soyez attentifs.

En cas de doutes, nous suivons les conventions de Wikipédia.

3.1 Inversibles et nilpotents

Le concept d'anneau est la définition 1.38.

Lemme 3.1.
Si a et b commutent, nous avons, pour tout $r \in \mathbb{N}$ et $r > 0$, la formule

$$a^{r+1} - b^{r+1} = (a-b)\left(\sum_{k=0}^{r} a^{r-k}b^k\right). \tag{3.1}$$

Démonstration. Démontrons cela par récurrence. Le cas $r = 0$ est évident. Pour un r donné, si (3.1) 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 3.2.
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 (3.1) nous avons

$$1 = 1 - a^n = (1-a)(1 + a + \cdots + a^{n-1}) = (1 + a + \cdots + a^{n-1})(1-a). \tag{3.2}$$

La somme $1 + a + \cdots + a^{n-1}$ est donc un inverse de $(1-a)$. □

3.2 PGCD, PPCM et éléments inversibles

La définition de pgcd et ppcm dans un anneau commutatif est la définition 1.215. Dans la plus grande tradition, elle a été introduite sans motivation, et utilisée par-ci par-là. Nous revenons maintenant dessus.

Commençons par donner une autre vision de la divisibilité dans les anneaux intègres.

Proposition 3.3.
Dans un anneau intègre[1] *A, on a l'équivalence suivante concernant deux éléments $a, b \in A$:*

$$a \mid b \Leftrightarrow (b) \subset (a). \tag{3.3}$$

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 $x \in S$.

Nous rappelons également la définition 1.39 de morphisme d'anneaux. Remarquons que si f est un morphisme, nous avons $f(0) = 0$ et $f(x)^{-1} = f(x^{-1})$.

Lemme 3.4 ([78])**.**
Les éléments inversibles d'un anneau sont diviseurs de tous les éléments.

Démonstration. Soit k inversible d'inverse k' : $kk' = 1$; soit aussi $a \in A$. Alors $a = k(k'a)$, ce qui montre que k divise a. □

Lemme 3.5 ([78])**.**
Dans un anneau, 1 est un pgcd de a et b si et seulement si les seuls diviseurs communs sont les inversibles.

Démonstration. Supposons pour commencer que 1 est un pgcd de a et b. Un diviseur commun de a et b doit donc diviser 1. Or un diviseur de 1 est forcément inversible.

Dans l'autre sens, les diviseurs communs de a et b sont tous inversibles et donc diviseurs de 1. Donc 1 est un pgcd de a et b. □

3.2.1 Calcul effectif du PGCD et théorème de Bézout

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 (19.2).

Cela se base sur le lemme suivant.

Lemme 3.6.
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

$$\frac{a}{s} = \frac{qb}{s} + \frac{r}{s}$$

1. Définition 1.221.

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 3.7.
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 $\operatorname{pgcd}(a, b)$ consiste à écrire des divisions euclidiennes successives de la manière suivante :

$$a = q_2 b + r_2 \qquad r_2 < b \tag{3.4a}$$
$$b = q_3 r_2 + r_3 \qquad r_3 < r_2 \tag{3.4b}$$
$$\vdots \tag{3.4c}$$

en remarquant que $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(b, r_2) = \operatorname{pgcd}(r_2, r_3)$. 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 $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(r_{n-1}, r_n) = r_n$.

3.2.1.1 Algorithme d'Euclide pour le PGCD

Écrivons l'algorithme[79] en détail (parce que Bézout, ça va être la même chose en cinq fois plus compliqué). On pose

$$r_0 = a \tag{3.5a}$$
$$r_1 = b \tag{3.5b}$$

(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 :

$$r_2 = r_0 - q_2 r_1. \tag{3.6}$$

Ensuite, sachant r_2 nous pouvons continuer :

$$r_1 = q_3 r_2 + r_3 \tag{3.7}$$

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

$$\begin{aligned} r_0 &= a \\ r_1 &= b \\ r_k &= q_{k+2} r_{k+1} + r_{k+2} \end{aligned} \tag{3.8}$$

où la troisième équation 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 3.6 et le principe de l'algorithme d'Euclide nous donnent

$$\begin{cases} \operatorname{pgcd}(r_k, r_{k+1}) = \operatorname{pgcd}(r_{k+1}, r_{k+2}) = \operatorname{pgcd}(a, b) & (3.9a) \\ 0 \leqslant r_{k+1} < r_k. & (3.9b) \end{cases}$$

La suite étant strictement décroissante, nous prenons r_n, le dernier non nul : $r_{n+1} = 0$. Dans ce cas, en prenant $k = n - 1$ dans la la dernière équation (3.8) devient :

$$r_{n-1} = q_{n+1} r_n \tag{3.10}$$

avec $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(r_n, r_{n-1}) = r_n$.

Exemple 3.8.
Calculons le PGCD de 18 et 231. Pour cela nous écrivons les divisions euclidiennes en chaine :

$$231 = 18 \cdot 12 + 15 \tag{3.11a}$$
$$18 = 1 \cdot 15 + 3 \tag{3.11b}$$
$$15 = 5 \cdot 5 + 0. \tag{3.11c}$$

Donc le PGCD est 3 (le dernier reste non nul). △

3.2.1.2 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 connaitre la liste complète des r_k jusqu'à $r_n = \text{pgcd}(a, b)$, ainsi que la liste complète des divisions euclidiennes

$$r_k = q_{k+2}r_{k+1} + r_{k+2}. \tag{3.12}$$

Nous voulons trouver les couples (u_k, v_k) de telle façon à avoir à chaque étape

$$r_n = u_k r_k + v_k r_{k-1}. \tag{3.13}$$

Notons que c'est à chaque fois r_n que nous construisons. La première équation de type Bézout à résoudre est

$$r_n = u_n r_n + v_n r_{n-1}, \tag{3.14}$$

sachant que $r_{n-1} = q_{n+1}r_n$. 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 :

$$r_n = u_k r_k + v_k r_{k-1} = u_{k-1}r_{k-1} + v_{k-1}r_{k-2}; \tag{3.15}$$

dans laquelle nous substituons $r_{k-2} = q_k r_{k-1} + r_k$:

$$u_k r_k + v_k r_{k-1} = u_{k-1}r_{k-1} + v_{k-1}(q_k r_{k-1} + r_k) \tag{3.16}$$
$$= (u_{k-1} + q_k v_{k-1})r_{k-1} + v_{k-1}r_k \tag{3.17}$$

et nous égalons les coefficients de r_k et r_{k-1} pour obtenir

$$\begin{cases} v_{k-1} = u_k & (3.18a) \\ u_{k-1} = v_k - v_{k-1}q_k. & (3.18b) \end{cases}$$

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

$$r_n = u_1 r_1 + v_1 r_0, \tag{3.19}$$

c'est-à-dire

$$\text{pgcd}(a, b) = u_1 b + v_1 a. \tag{3.20}$$

Nous avons ainsi trouvé des coefficients de Bézout pour a et b.

3.2.2 Générateurs

3.9.
Les éléments de $\mathbb{Z}/n\mathbb{Z}$ ne sont pas des éléments de $\mathbb{Z}$; ce sont des parties de $\mathbb{Z}$. Pour rappel :

$$[a]_n = \{a + kn \text{ tel que } k \in \mathbb{Z}\}. \tag{3.21}$$

Pour écrire les éléments de $\mathbb{Z}/n\mathbb{Z}$, nous pouvons écrire

$$\{[i]_n\}_{i=1,\dots,n} \tag{3.22}$$

ou

$$\{[i]_n \text{ tel que } 1 \leqslant i \leqslant n\}. \tag{3.23}$$

Mais attention : l'ensemble

$$\bigcup_{i=1}^{n}[i]_n \tag{3.24}$$

est très différent. Ce dernier ensemble est $\mathbb{Z}$.

Proposition 3.10 ([80]).
Soit $n \geqslant 2$.

(1) L'anneau $\mathbb{Z}/n\mathbb{Z}$ contient n éléments

(2) Nous avons $\mathbb{Z}/n\mathbb{Z} = \{[k]_n\}_{k=0,\dots,n-1}$.

Proposition 3.11 ([80]).
Soient $n \geqslant 2$, a premier avec n, $b \in \mathbb{N}$, et $m_1, \dots, m_n \in \mathbb{N}$. Alors

(1) Nous avons $\mathbb{Z}/n\mathbb{Z} = \{[am_i + b]_n\}_{i=1,\dots,n}$,

(2) Nous avons $\mathbb{Z}/n\mathbb{Z} = \{[b + k]_n\}_{k=0,\dots,n-1}$

(3) Nous avons $\mathbb{Z}/n\mathbb{Z} = \{[ka]_{k=0,\dots,n-1}\}$.

Proposition 3.12 ([80]).
Soit $n \geqslant 2$ et $m \in \mathbb{Z}$. Nous avons équivalence entre

(1) $\operatorname{pgcd}(n, m) = 1$,

(2) $[m]_n$ engendre le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$

(3) $[m]_n$ est inversible dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \big)$.

3.2.3 Générateurs pour le groupe multiplicatif

Proposition 3.13 ([81]).
Le groupe $\big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \big)$ est cyclique.

La proposition suivante est un pas important dans l'algorithme de Shor permettant aux ordinateurs quantiques de factoriser rapidement des grands nombres.

Proposition 3.14 ([82]).
Soient $A, B \in \mathbb{N}$ tels que $\operatorname{pgcd}(A, B) = 1$. Alors il existe $p, m \in \mathbb{N}$ tels que $A^p = mB + 1$.

3.2.4 Décomposition en facteurs premiers

Lemme 3.15 (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. Puisque a est premier avec c, nous avons $\operatorname{pgcd}(a, c) = 1$ et le théorème de Bézout 1.228 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 3.16 (Lemme d'Euclide[83]).
Soient $a, b \in \mathbb{Z}$. Si le nombre premier p divise le produit ab, alors p divise a ou b.

Démonstration. Comme p est premier, si il ne divise pas a c'est que $\operatorname{pgcd}(a, p) = 1$. Dans ce cas le lemme de Gauss 3.15 implique que p divise b. □

Le théorème fondamental de l'arithmétique permet de décomposer des nombres en facteurs premiers.

Théorème 3.17 (Décomposition en facteurs premiers[84])**.**
Tout entier strictement positif peut être écrit comme un produit de nombres premiers d'une unique façon, à l'ordre près des facteurs.

En d'autres termes, pour tout entier $n > 1$, il existe une unique suite finie unique (p_1, k_1),...(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) $n = \prod_{i=1}^{r} p_i^{k_i}$.

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 [2].

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 3.16, 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. □

Lemme 3.18 ([1])**.**
Nous notons $\mathcal{P}$ l'ensemble des nombres premiers dans $\mathbb{N}$. Soient des suites finies $(a_p)_{p\in\mathcal{P}}$ et $(b_p)_{p\in\mathcal{P}}$. Nous posons

$$a = \prod_{p\in\mathcal{P}} p^{a_p} \quad \text{et} \quad b = \prod_{p\in\mathcal{P}} p^{b_p}. \tag{3.25}$$

Alors $a \mid b$ si et seulement si $a_p \leqslant b_p$ pour tout p.

Démonstration. Dire que $a \mid b$ signifie qu'il existe $s \in \mathbb{N}$ tel que $as = b$; le théorème 3.17 nous permet de décomposer s en $s = \prod_{p\in\mathcal{P}} p^{s_p}$. Puisque le produit dans $\mathbb{N}$ est commutatif et associatif,

$$b = as = \prod_{p\in\mathcal{P}} p^{s_p+a_p}. \tag{3.26}$$

Par unicité de la décomposition de b (toujours le théorème 3.17), nous en déduisons que $b_p = s_p + a_p \geqslant a_p$.

Dans l'autre sens, l'hypothèse $a_p \leqslant b_p$ implique l'existence de $s_p \geqslant 0$ tels que $b_p = a_p + s_p$. En posant $s = \prod_{p\in\mathcal{P}} p^{s_p}$, nous avons

$$as = \prod_{p\in\mathcal{P}} p^{s_p+a_p} = \prod_{p\in\mathcal{P}} p^{b_p} = b. \tag{3.27}$$

Donc $a \mid b$. □

2. Voir 1.260.

Proposition 3.19.
Soient $x \in \mathbb{Z}$ ainsi que $p, q \in \mathbb{N}$. Si q divise x^p, alors q divise x.

Lemme 3.20.
Soient un nombre premier q ainsi que $a \in \mathbb{Z}$. Soit un entier $n \geqslant 1$. Le nombre q divise a si et seulement si il divise a^n.

Démonstration. Nous numérotons les nombres premiers p_i pour que p_1 soit q. La décomposition en nombre premiers du théorème 3.17 nous dit que

$$a = q^{a_1} \prod_{i \neq 1} p_i^{a_i} \tag{3.28}$$

et

$$a^n = q^{na_1} \prod_{i \neq 1} p_i^{na_i} \tag{3.29}$$

Nous avons équivalence entre les énoncés suivants :

— q divise a
— $a_1 \neq 0$
— $na_1 \neq 0$ (parce que $n \neq 0$)
— q divise a^n.

□

Lemme 3.21 ([1])**.**
Dans $\mathbb{N}$, le pgcd[3] et le ppcm sont uniques.

Démonstration. Supposons que δ_1 et δ_2 soient des pgcd de la partie S. Puisque $\delta_1 \mid S$, nous avons $\delta_1 \mid \delta_2$ parce que δ_2 est un pgcd. Le même raisonnement, inversant δ_1 et δ_2 montre que $\delta_2 \mid \delta_1$. Si (a_p) sont les éléments de la décomposition de δ_1 et (b_p) ceux de δ_2, alors le lemme 3.18 nous indique que $a_p \leqslant b_p$ et $b_p \leqslant a_p$, ce qui implique que $a_p = b_p$.

La démonstration pour le ppcm s'effectue selon le même principe. □

Lemme 3.22.
Soit une partie S de $\mathbb{N}$.

(1) Le pgcd de S est le plus grand élément de $\mathbb{N}$ divisant tous les éléments de S.
(2) Le ppcm de S est le plus petit élément de $\mathbb{N}$ que tous les éléments de S divisent.

Lemme 3.23 ([1])**.**
Soient $a, b \in \mathbb{Z}$ et $k \in \mathbb{N}$ tels que $ab = q^k$ et $\operatorname{pgcd}(a, b) = 1$. Alors il existe $\alpha, \beta \in \mathbb{Z}$ tels que $a = \alpha^k$ et $b = \beta^k$.

Démonstration. Nous décomposons a, b et q en facteurs premiers suivant le théorème 3.17 :

$$a = \prod_i p_i^{a_i} \tag{3.30a}$$

$$b = \prod_i p_i^{b_i} \tag{3.30b}$$

$$q = \prod_i p_i^{q_i}. \tag{3.30c}$$

D'un part, en utilisant la commutativité et l'associativité du produit,

$$ab = \prod_i p_i^{a_i + b_i}. \tag{3.31}$$

3. Le pgcd et ppcm sont définis dans 1.215.

D'autre part, puisque $ab = q^k$, nous avons

$$ab = (\prod_i p_i^{q_i})^k = \prod_i p_i^{kq_i}. \tag{3.32}$$

En vertu de l'unicité de la décomposition en facteurs premiers, pour chaque i nous avons

$$a_i + b_i = kq_i. \tag{3.33}$$

Comme a et b sont premiers entre eux, si $a_i \neq 0$ alors $b_i = 0$ et inversement. Prenons un i tel que $a_i \neq 0$. Alors $b_i = 0$ et nous avons $a_i = kq_i$. Idem pour les b_i.

Donc tous les a_i et les b_i qui sont non nuls sont des multiples de k. Nous posons $a_i = ks_i$ et nous reportons dans (3.30a) :

$$a = \prod_i p_i^{ks_i} = (\prod_i p_i^{s_i})^k, \tag{3.34}$$

de telle sorte que a soit une puissance k^{e}. La même chose tient pour b. □

Proposition 3.24.
Soient $a, b \in \mathbb{Z}\backslash\{0\}$ décomposés en $a = \prod_{p\in\mathcal{P}} p^{a_p}$ et $b = \prod_{p\in\mathcal{P}} p^{b_p}$. En posant

$$m_p = \min\{a_p, b_p\} \tag{3.35a}$$
$$M_p = \max\{a_p, b_p\}, \tag{3.35b}$$

nous avons

$$\text{pgcd}(a, b) = \prod_{p\in\mathcal{P}} p^{m_p} \tag{3.36a}$$
$$\text{ppcm}(a, b) = \prod_{p\in\mathcal{P}} p^{M_p}. \tag{3.36b}$$

Pour rappel, la définition de pgcd et ppcm sont dans 1.215.

Démonstration. Nous commençons par le pgcd. Nous notons $\delta = \prod_{p\in\mathcal{P}} p^{m_p}$ et nous prouvons que δ est un pgcd de $\{a, b\}$. Il y a deux propriétés à vérifier.

(i) **δ divise a et b** Puisque $m_p = \min\{a_p, b_p\}$, nous avons $m_p \leqslant a_p$ et $m_p \leqslant b_p$. Le lemme 3.18 nous dit alors que $\delta \mid a$ et $\delta \mid b$.

(ii) **Si d divise a et b** De même, si $s \mid a$ et $s \mid b$, nous avons $s_p \leqslant a_p$ et $s_p \leqslant b_p$, ce qui montre que $s_p \leqslant m_p$ et donc que $s \mid \delta$.

Pour le ppcm, nous posons $\mu = \prod_{p\in\mathcal{P}} p^{M_p}$, et nous prouvons que μ est un ppcm de $\{a, b\}$.

(i) **a et b divisent μ** Pour tout p, nous avons $M_p \geqslant a_p$. Le lemme 3.18 implique que $a \mid \mu$. Idem pour b, donc tous les éléments de $\{a, b\}$ divisent μ.

(ii) **Si a et b divisent r** Supposons que a et b divisent un certain nombre r. Alors $a_p \leqslant r_p$ et $b_p \leqslant r_p$. Donc $r_p \geqslant \max\{a_p, b_p\} = M_p$. Nous en déduisons que $\mu \mid r$.

Puisque les pgcd et ppcm sont uniques (lemme 3.21), nous avons prouvé que δ et μ sont les nombres recherchés. □

Corolaire 3.25.
Soit un nombre premier p. Un élément $m \in \mathbb{Z}^*$ vérifie $\text{pgcd}(m, p^n) \neq 1$ si et seulement si $m = qp$ pour un certain q.

Démonstration. Nous considérons les décompositions en facteurs premiers $m = \prod_{s\in\mathcal{P}} s^{a_s}$ et $p^n = \prod_{s\in\mathcal{P}} s^{b_s}$. Par la partie unicité du théorème 3.17, nous savons que $b_s = 0$ pour $s \neq p$ et $b_p = n$.

Nous prenons ensuite l'expression du pgcd donné par la proposition 3.24 :

$$\text{pgcd}(m, p^n) = \prod_{s\in\mathcal{P}} s^{\min\{a_s, b_s\}}. \tag{3.37}$$

Le minimum est toujours zéro lorsque $s \neq p$, donc $\text{pgcd}(m, p^n) = p^{\min\{a_s, n\}}$.

Nous avons donc $\text{pgcd}(m, p^n) \neq 1$ si et seulement si $a_p \neq 0$. Mais $a_p \neq 0$ si et seulement si m est un multiple de p. □

Lemme 3.26.
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[4] et nous séparons les puissances paires des puissances impaires :

$$n = \prod_{i=1}^{r} p_i^{2\alpha_i} \prod_{j=1}^{s} q_j^{2\beta_j+1} \tag{3.38a}$$

$$= \underbrace{\left(\prod_{i=1}^{r} p_i^{2\alpha_i} \prod_{j=1}^{s} q^{2\beta_j}\right)}_{m^2} \underbrace{\prod_{j=1}^{s} q_j}_{q} . \tag{3.38b}$$

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

$$n = q_1 d^2 k_1^2 = q_2 d^2 k_2^2, \tag{3.39}$$

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 $m_2 = lm_1$ alors l'équation (3.39) se réduit à $n = q_1 m_1^2 = q_2 l^2 m_1^2$ et donc

$$q_1 = q_2 l^2, \tag{3.40}$$

ce qui signifie $l = 1$ et donc $m_1 = m_2$. □

Les nombres premiers ne sont pas si rares que ça dans $\mathbb{N}$. Nous allons voir dans 15.111 que la somme des inverses des nombres premiers diverge. Pour comparaison, la somme des inverses des carrés converge par la proposition 11.120. Il y a donc, dans un certains sens, plus de nombres premiers que de carrés ; dans un autre sens, il y en a autant : une infinité dénombrable.

3.2.5 Ordre d'un élément dans un groupe fini

Voir plus d'informations dans la partie 5.2 sur les groupes monogènes.

Théorème 3.27 (Théorème de Cauchy[85])**.**
Soit un groupe fini d'ordre[5] 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 3.28 ([86])**.**
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 comme r et s sont premiers entre eux, l'ordre de ab s'écrit sous la forme $r_1 s_1$ avec $r_1 \mid r$ et $s_1 \mid s$. Nous avons

$$a^{r_1 s_1} b^{r_1 s_1} = (ab)^{r_1 s_1} = 1, \tag{3.41}$$

4. Théorème 3.17.
5. Définition 1.172.

que nous élevons à la puissance r_2 où r_2 est défini en posant $r = r_1 r_2$:

$$a^{rs_1} b^{rs_1} = 1. \tag{3.42}$$

Et comme $a^{rs_1} = 1$, nous concluons que $b^{rs_1} = 1$. Donc $s \mid rs_1$. Par le lemme de Gauss 3.15, nous avons en fait $s \mid s_1$. Puisqu'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 $r_1 s_1 = rs$. □

Lemme 3.29 ([40])**.**
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 2.13) 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 1.170. □

Lemme 3.30 ([87])**.**
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 3.31 ([87])**.**
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.

Démonstration. Soit H' un sous-groupe d'ordre n. Si $h \in H'$ alors $h^n = 1$ et $h^m \in H$ par le lemme 3.30. Étant donné que m et n sont premiers entre eux, par le théorème de Bézout 1.228, il existe $a, b \in \mathbb{Z}$ tels que

$$am + bn = 1. \tag{3.43}$$

Et donc, $h = h^1 = (h^m)^a (h^n)^b$. 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 $|H'| = |H| = |G|/m$. □

3.32.
Notons que cette proposition ne dit pas qu'il existe un sous-groupe d'ordre n et d'indice m. Il dit juste que si il y en a un et si il est normal, alors il n'y en a pas d'autre.

Lemme 3.33.
L'ensemble des ordres d'un groupe commutatif est stable par PPCM[6]*.*

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 $\operatorname{ppcm}(r, s)$.

Démonstration. Soit $m = \operatorname{ppcm}(r, s)$. Afin d'écrire m sous une forme pratique, nous considérons les décompositions en facteurs premiers de r et s :

$$r = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{3.44a}$$

$$s = \prod_{i=1}^{k} p_i^{\beta_i} \tag{3.44b}$$

6. Définition 1.215.

où $\{p_i\}_{i=1,\dots,k}$ est l'ensemble des nombres premiers arrivant dans les décompositions de r et de s. Si nous posons

$$r' = \prod_{\substack{i=1\\ \alpha_i>\beta_i}}^{k} p_i^{\alpha_i} \tag{3.45a}$$

$$s' = \prod_{\substack{i=1\\ \alpha_i\leqslant\beta_i}}^{k} p_i^{\beta_i}, \tag{3.45b}$$

alors $\operatorname{ppcm}(r,s) = r's'$ 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 3.28 pour conclure que l'ordre de $x^{r/r'}y^{s/s'}$ est $r's' = m$. □

3.2.6 Écriture des fractions

Théorème 3.34.
Tout élément de $\mathbb{Q}^+$ s'écrit de façon unique comme quotient de deux entiers premiers entre eux.

Démonstration. En deux parties[7]

(i) **Unicité** Supposons avoir $\frac{a}{b} = \frac{c}{d}$ avec $\operatorname{pgcd}(a,b) = \operatorname{pgcd}(c,d) = 1$. Nous avons

$$ad = bc \tag{3.46}$$

donc

(1) a divise bc mais est premier avec b donc a divise c par le lemme de Gauss 3.15.

(2) c divise ad mais est premier avec d donc c divise a par le lemme de Gauss 3.15.

En conclusion a divise c et c divise a, ergo $a = c$. L'égalité $b = d$ est alors immédiate.

(ii) **Existence** Soit le quotient $\frac{a}{b}$. Nous avons

$$\frac{a}{b} = \frac{a/\operatorname{pgcd}(a,b)}{b/\operatorname{pgcd}(a,b)}, \tag{3.47}$$

qui est encore un quotient d'entiers parce que $\operatorname{pgcd}(a,b)$ divise aussi bien a que b. Il faut montrer que les nombres $a/\operatorname{pgcd}(a,b)$ et $b/\operatorname{pgcd}(a,b)$ sont premiers entre eux. Pour cela nous supposons que k est un diviseur commun. En particulier, les nombres $a/k\operatorname{pgcd}(a,b)$ et $b/k\operatorname{pgcd}(a,b)$ sont des entiers, ce qui fait que $k\operatorname{pgcd}(a,b)$ est un diviseur commun de a et b. Étant donné que $\operatorname{pgcd}(a,b)$ est le plus grand tel diviseur, nous devons avoir $k\operatorname{pgcd}(a,b) = \operatorname{pgcd}(a,b)$ c'est-à-dire que $k = 1$. Donc les nombres $a/\operatorname{pgcd}(a,b)$ et $b/\operatorname{pgcd}(a,b)$ sont premiers entre eux.

□

Proposition 3.35.
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 1.228 il existe $a, b \in \mathbb{Z}$ tels que

$$ap^2 + bq^2 = 1 \tag{3.48}$$

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 $\operatorname{pgcd}(p,q) = k \neq 1$. 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. □

7. Définitions des pgcd et ppcm en 1.215.

Une des conséquences de ces résultats sera le fait que $\sqrt{n}$ est irrationnelle dès que n n'est pas un carré parfait, théorème 3.36.

Nous avons déjà vu dans la proposition 1.344 que $\sqrt{2}$ était irrationnel. En fait le théorème suivant va nous montrer que le nombre $\sqrt{n}$ est soit entier, soit irrationnel.

Théorème 3.36.
Soit $n \in \mathbb{N}$. Le nombre $\sqrt{n}$ est rationnel si et seulement si n est un carré parfait.

Démonstration. Supposons que $\sqrt{n}$ soit rationnel. Le théorème 3.34 nous donne $p, q \in \mathbb{N}$ premiers entre eux tels que $\sqrt{n} = p/q$. La proposition 3.35 nous enseigne de plus que p^2 et q^2 sont premiers entre eux. Nous avons

$$p^2 = nq^2. \tag{3.49}$$

Le nombre q est alors un diviseur commun de q^2 et de p. Donc $q = 1$ et $n = p^2$ est un carré parfait. □

3.2.7 Équation diophantienne linéaire à deux inconnues

Soient a, b et c des entiers naturels donnés. Nous considérons l'équation

$$ax + by = c \tag{3.50}$$

à résoudre[88] pour $(x, y) \in \mathbb{N}^2$.

Si a ou b est nul, c'est facile ; nous supposons donc que a et b sont tous 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

$$d(a'x + b'y) = c, \tag{3.51}$$

et donc si c n'est pas un multiple de d, il n'y a pas de solution [8]. Si par contre c est un multiple de d alors nous écrivons $c = c'd$ et l'équation devient

$$a'x + b'y = c' \tag{3.52}$$

C'est maintenant que nous utilisons le théorème de Bézout 1.228 : puisque a' et b' sont premiers entre eux, nous avons la relation $a'u + b'v = 1$ pour certains [9] 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 :

$$a'(x - c'u) = b'(c'v - y). \tag{3.53}$$

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 [10] diviser $c'v - y$. Disons $c'v - y = ka'$. Alors $a'(x - c'u) = b'ka'$ et donc

$$x = b'k + c'u. \tag{3.54}$$

Nous trouvons alors une expression pour y en injectant ce résultat dans $a'x + b'y = c'$ et en utilisant le théorème de Bézout : $a'c'u = (1 - b'v)c'$. Au final nous avons prouvé que toutes les solutions sont de la forme

$$\begin{cases} x = b'k + c'u & \text{(3.55a)} \\ y = vc' - a'k & \text{(3.55b)} \end{cases}$$

avec $k \in \mathbb{Z}$. Si nous ne voulons réellement que les 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 (3.55) sont solutions.

8. Exemple : $8x + 2y = 9$. Le membre de gauche est certainement un nombre pair et il n'y a donc pas de solution.
9. Nous avons décrit un algorithme pour les trouver en démontrant l'équation 3.20.
10. C'est Gauss 3.15 qui le dit, et vous savez que lorsque Gauss dit, c'est *forcément* vrai.

Exemple 3.37.
Résoudre l'équation $2x + 6y = 52$.

Nous pouvons factoriser 2 dans le membre de gauche et simplifier alors toute l'équation par 2 :

$$x + 3y = 26. \tag{3.56}$$

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

$$x + 3y = 26 \times (-5 + 3 \times 2), \tag{3.57}$$

et donc $x + 5 \times 26 = -3(y - 26 \times 2)$, et en posant $k = y - 26 \times 2$ nous avons

$$x = -3k - 130. \tag{3.58}$$

En injectant nous trouvons l'équation $3y - 3k - 130 = 26$ et donc

$$y = 52 + k. \tag{3.59}$$

△

3.2.8 Quotients

Nous écrivons $a = b \mod p$ essentiellement si 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 3.38.
Soit $n \in \mathbb{N}$. Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est monogène. Si $n \neq 0$, il est cyclique d'ordre n.

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 $\mathbb{Z}/n\mathbb{Z} = \mathrm{gr}\big(\mu(1)\big)$ 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 1.224 donne l'existence de q et r avec $0 \leqslant r < n$ et $q \geqslant 0$ tels que

$$x_0 = nq + r. \tag{3.60}$$

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

$$\mathbb{Z}/n\mathbb{Z} = \mu(\mathbb{Z}) = \mu(\mathbb{N}_{n-1}). \tag{3.61}$$

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 $\mathbb{Z}/n\mathbb{Z} = \mathrm{gr}(\mu(1))$. Il est donc cyclique. □

Lemme 3.39 ([89]).
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 1.224, 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

$$q^n = (q^d)^a q^b \in [1]_{q^d-1} q^b = [q^b]_{q^d-1}. \tag{3.62}$$

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

$$q^n \in [1]_{q^d-1}. \tag{3.63}$$

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

$$[1]_{q^d-1} = [q^b]_{q^d-1}, \tag{3.64}$$

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. □

3.3 Binôme de Newton et morphisme de Frobenius

Proposition 3.40 ([90]).
Pour tout x, $y \in \mathbb{R}$ et $n \in \mathbb{N}$, nous avons

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \tag{3.65}$$

où

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{3.66}$$

sont les ***coefficients binomiaux****.*

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 (3.65) soit vraie pour $n \geqslant 1$, et prouvons la pour $n+1$. Nous avons

$$\begin{aligned}
(x+y)^{n+1} &= (x+y) \cdot \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n-1} \binom{n}{k} x^{n-k} y^{k+1} + y^{n+1}.
\end{aligned} \tag{3.67}$$

La seconde grande somme peut être transformée en posant $i = k+1$:

$$\sum_{k=0}^{n-1} \binom{n}{k} x^{n-k} y^{k+1} = \sum_{i=1}^{n} \binom{n}{i-1} x^{n-(i-1)} y^{i-1+1}, \tag{3.68}$$

dans lequel nous pouvons immédiatement renommer i par k. En remplaçant dans la dernière expression de (3.67), nous trouvons

$$(x+y)^{n+1} = x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] x^{n-k+1} y^k. \tag{3.69}$$

La thèse découle maintenant de la formule

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k} \tag{3.70}$$

qui est vraie parce que

$$\frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} = \frac{n!(n-k+1) + n!k}{k!(n-k+1)!} = \frac{n!(n+1)}{k!(n+1-k)!}, \tag{3.71}$$

par simple mise au même dénominateur. □

Lemme 3.41 ([1]).
Si $n \geqslant k$ nous avons

$$\frac{n!}{k!(n-k)!} \leqslant \frac{n^{k-1}}{k!}. \tag{3.72}$$

Démonstration. Nous décomposons le produit définissant $n!$ en les facteurs entre 1 et $n-k$ et ceux entre $n-k+1$ et n :

$$n! = (n-k)! \prod_{i=n-k+1}^{n} i \leqslant n^{k-1}(n-k)!. \tag{3.73}$$

Donc

$$\frac{n!}{k!(n-k)!} \leqslant \frac{n^{k-1}}{k!}. \tag{3.74}$$

□

Tant que nous sommes à démontrer des égalités, en voici une.

Lemme 3.42 ([91]).
Pour $a, b \in \mathbb{R}$ et $n \in \mathbb{N}$ nous avons

$$a^n + (-1)^{n-1}b^n = (a+b)\sum_{k=0}^{n-1}(-1)^k a^{n-1-k}b^k. \tag{3.75}$$

Démonstration. C'est un simple calcul :

$$(a+b)\sum_{k=0}^{n-1}(-1)^k a^{n-1-k}b^k = \sum_{k=0}^{n-1}(-1)^k a^{n-k}b^k + \sum_{k=0}^{n-1}(-1)^k a^{n-k-1}b^{k+1} \tag{3.76a}$$

$$= a^n + \sum_{k=1}^{n-1}(-1)^k a^{n-k}b^k + \sum_{k=0}^{n-2}(-1)^k a^{n-k-1}b^{k+1} + (-1)^{n-1}b^n \tag{3.76b}$$

$$= a^n + (-1)^{n-1}b^n \tag{3.76c}$$

Justifications.

— Pour (3.76b). Dans la première somme, nous avons séparé le terme $k = 0$ et dans la seconde nous avons séparé le terme $k = n-1$

— Pour (3.76c). Dans la seconde somme, décaler les termes pour sommer de 1 à $n-1$ et remarquer que ce qu'on obtient annule la première somme.

□

Proposition 3.43.
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

$$(a+b)^p = a^p + b^p \tag{3.77}$$

pour tout $a, b \in A$.

Démonstration. Nous utilisons la formule du binôme de la proposition 3.40 et le fait que les coefficients binomiaux non extrêmes sont divisibles par p et donc nuls. □

3.4 Idéal dans un anneau

La définition d'un idéal dans un anneau est la définition 1.213.

Définition 3.44 (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 3.45.
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 3.46.
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 3.47.
L'ensemble $2\mathbb{Z}$ est un idéal de $\mathbb{Z}$. On peut aussi le noter (2). △

Proposition 3.48 (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

$$\tilde{f}\colon A/\ker f \to f(A) \tag{3.78}$$

tel que $f = j \circ \tilde{f} \circ \phi$.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\phi"'] & B \\ A/\ker f \arrow[r, "\tilde{f}"'] & f(A) \subset B \arrow[u, "j"'] \end{tikzcd} \tag{3.79}$$

Proposition 3.49.
Soient I un idéal de A et la projection canonique

$$\phi\colon A \to A/I. \tag{3.80}$$

C'est une bijection entre les idéaux de A contenant I et les idéaux de A/I.

Dit de façon imagée :

$$\{\text{idéaux de } A \text{ contenant } I\} \simeq \{\text{idéaux de } A/I\}. \tag{3.81}$$

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

$$\phi(i)\phi(j) = \phi(ij) \in \phi(J). \tag{3.82}$$

Soit maintenant K un idéal dans A/I et 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 $I \subset \phi^{-1}(K)$. □

Proposition 3.50 ([1])**.**
Si A est un anneau, nous avons les équivalences

(1) *A est un corps*[11].

(2) *A est non nul et ses seuls*[12] *idéaux à gauche sont $\{0\}$ et A.*

(3) *A est non nul et ses seuls idéaux à droite sont $\{0\}$ et A.*

Démonstration. Nous allons montrer que le point (1) est équivalent aux deux autres.

11. Définition 1.237.
12. Je vous laisse vous poser de grandes questions sur le fait que le vide est un idéal ou non.

(i) **(3) implique (2)** Si I est un idéal à gauche différent de $\{0\}$, alors il contient un certain $a \neq 0$. Puisque A est un corps, il contient un inverse a^{-1}, et comme I est un idéal, $a^{-1}I \subset I$. En particulier $a^{-1}a \in I$. Donc $1 \in I$ et $I = A$.

(ii) **(2) implique (1)** Supposons que les seuls idéaux de A soient $\{0\}$ et A. Soit $a \in A$. Si a est non nul, alors $aA = A$, en particulier, $1 \in aA$, c'est-à-dire qu'il existe $b \in A$ tel que $ab = 1$. L'élément a est donc inversible.

□

Définition 3.51.
*Un idéal I dans un anneau A est dit **idéal maximal** ou idéal maximum si tout idéal J vérifiant $I \subset J \subset A$ est soit I, soit A.*

Proposition 3.52 (Thème 16)**.**
Un idéal I dans un anneau A est maximum si et seulement si A/I est un corps.

Démonstration. Soit un idéal maximum $I \subset A$. Alors les idéaux contenant I sont A et I. L'application ϕ de la proposition 3.49 est une bijection, donc l'ensemble des idéaux de A/I ne contient que deux éléments. Les seuls idéaux de A/I sont donc $\{0\}$ et A/I ; donc A/I est un corps par la proposition 3.50.

Dans l'autre sens, c'est la même chose : si A/I est un corps, il possède exactement deux idéaux, donc A ne contient que deux idéaux contenant I. Donc I est un idéal maximum. □

3.5 Polynômes de plusieurs variables

Définition 3.53.
*L'ensemble des **polynôme de n variables** sur l'anneau A est $A^{(\mathbb{N}^n)}$, c'est-à-dire l'ensemble des suites indexées par $\mathbb{N}^n$ et dont seulement une quantité finie de coefficients sont non nuls.*

Le produit sur $A[X_1, \dots, X_n]$ est défini par

$$(PQ)(k_1, \dots, k_n) = \sum_{\substack{(l_1,\dots,l_n),(m_1,\dots,m_n)\in\mathbb{N}^n\times\mathbb{N}^n \\ l_i+m_i=k_i}} P_{l_1,\dots,l_n} Q_{m_1,\dots,m_n}. \tag{3.83}$$

3.54.
Dans $A[X_1, \dots, X_n]$, la multiplication n'est pas la multiplication de fonctions $\mathbb{N}^n \to \mathbb{K}$, parce que le but est d'obtenir la multiplication usuelle au niveau des évaluations.

Définition 3.55.
*Si P est un polynôme de n variables sur A, et si $(x_1, \dots, x_n) \in A^n$, **l'évaluation** de P sur $(x_1, \dots, x_n)$ est*

$$P(x_1, \dots, x_n) = \sum_{(k_1,\dots,k_n)\in\mathbb{N}^n} P_{k_1,\dots,k_n} x_1^{k_1} \dots x_n^{k_n}. \tag{3.84}$$

Notez que la somme, bien que sur $\mathbb{N}^n$, est une somme finie.

3.56.
Comme dans le cas des polynômes d'une seule variable, les X_i dans la notation $A[X_1, \dots, X_n]$ sont à prendre à la légère. L'anneau des polynômes de n variables sur A aurait mieux fait d'être noté par exemple par $\mathcal{P}_n(A)$.

Le fait est que nous avons les polynômes élémentaires définis par

$$X_1(k_1, \dots, k_n) = \begin{cases} 1 & \text{si } (k_1, \dots, k_n) = (1, 0 \dots, 0) \\ 0 & \text{sinon.} \end{cases} \tag{3.85}$$

et que l'anneau des polynômes peut être vu comme A (les polynômes constants) étendus par les X_i.

Quoi qu'il en soit, les X_i dans la notation $A[X_1, \dots, X_n]$ sont des indices muets. L'anneau $A[X_1, \dots, X_n]$ est exactement le même que $A[T_1, \dots, T_n]$.

3.5.1 Divisibilité et classes d'association

Définition 3.57.
On dit de deux éléments $a, b \in A$ qu'ils sont ***associés*** *si 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.*

Lemme 3.58.
Si A est un anneau intègre et si $a, b \in A$ sont deux éléments vérifiant $a \mid b$ et $b \mid a$, alors ils sont associés, c'est à dire qu'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 $x, y \in A$. Alors,

$$b(1 - yx) = 0. \tag{3.86}$$

Étant donné que 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. □

Lemme 3.59.
Si A est un anneau et si $a \in A$, la classe d'association de a est $aU(A)$ où $U(A)$ est l'ensemble des éléments inversibles de A.

Exemple 3.60.
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 :

$$(2 + i)(4 + i) = 7 + 6i. \tag{3.87}$$

△

Exemple 3.61.
Si $\mathbb{K}$ est un corps, l'élément XY de $\mathbb{K}[X, Y]$ n'est pas premier parce que $XY \mid X^2Y^2$ alors que XY ne divise ni X^2 ni Y^2. △

¡¡ Avertissement/question à la lectrice !! 3.62
Est-ce que quelqu'un connaît un anneau contenant $\mathbb{Z}$ dans lequel 7 est divisible en 2 ?

Peut-être $\mathbb{Z}$ étendu par tous les $1/2^n$?

3.5.2 PGCD et PPCM

Pour le théorème de Bézout et autres opérations avec des modulo, voir le thème 3. Le pgcd et le ppcm sont définis en 1.215.

Lemme 3.63.
Soient 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. Soient 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 $u \in A$ tel que $\delta' = u\delta$. Comme δ' divise x pour tout $x \in S$, nous avons $\delta' \mid \delta$, et symétriquement nous trouvons $\delta \mid \delta'$. Par conséquent (lemme 3.58), 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 3.64.
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.

3.5.3 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 3.65.
Un anneau fini intègre est un corps.

Démonstration. Soit A un tel anneau. Soit $a \neq 0$. Les applications

$$l_a \colon x \to ax \tag{3.88a}$$
$$r_a \colon x \to xa \tag{3.88b}$$

sont injectives. En tant que applications injectives entre ensembles finis, elles sont surjectives. Il existe donc b et c tels que $1 = ba = ac$. Nous trouvons que b et c sont égaux parce que [13]

$$b = b(ac) = (ba)c = c. \tag{3.89}$$

Par conséquent b est un inverse de a. □

Proposition 3.66.
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 corolaire 1.248. Le fait que $\mathbb{Z}/n\mathbb{Z}$ soit un corps lorsque $\mathbb{Z}/n\mathbb{Z}$ est intègre est la proposition 3.65. 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 1.242). □

3.5.4 Élément irréductible

Définition 3.67 (Élément irréductible[92])**.**
*Un élément d'un anneau commutatif est **irréductible** si il n'est ni inversible, ni le produit de deux éléments non inversibles.*

3.68.
Nous allons voir dans la section 3.8 que le concept d'élément irréductible n'est vraiment utile que dans le cas des anneaux intègres.

13. Il faut être un peu souple avec les notations communément employées dans les ouvrages mathématiques, et que nous reprenons telles quelles. Dans l'équation qui suit, $b(ac)$ est le produit de b par l'élément ac, et non quelque chose comme le produit de b avec l'idéal (ac) par exemple.

Exemple 3.69.
Un corps n'a pas d'élément irréductible 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 3.70.
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 $b = \pm 1$. △

3.6 Anneau factoriel

Définition 3.71 (Anneau factoriel)**.**
*Un anneau commutatif A est **factoriel** si 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[14] $\sigma \in S_k$ telle que p_i et $q_{\sigma(i)}$ soient associés[15].

Un anneau factoriel permet de caractériser le pgcd et le ppcm de nombres.

Proposition 3.72.
Soit une famille $\{a_n\}$ d'éléments de A qui se décomposent en irréductibles comme

$$a_i = \prod_k p_k^{\alpha_{k,i}}. \tag{3.90}$$

Alors

$$\text{pgcd}\{a_n\} = \prod_k p_k^{\min_i\{\alpha_{k,i}\}}. \tag{3.91}$$

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[16].

De la même manière,

$$\text{ppcm}\{a_n\} = \prod_k p_k^{\max_i\{\alpha_{k,i}\}}. \tag{3.92}$$

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[17] : $a < b$ si et seulement si $(b) \subset (a)$.

Exemple 3.73.
L'anneau $\mathbb{Z}[i\sqrt{3}]$ n'est pas factoriel parce que 4 a au moins deux décompositions distinctes en irréductibles :

$$4 = 2 \cdot 2, \tag{3.93}$$

et

$$4 = (1 + i\sqrt{3})(1 - i\sqrt{3}). \tag{3.94}$$

△

Nous allons voir dans l'exemple 3.110 que $\mathbb{Z}[i\sqrt{2}]$ est factoriel parce qu'il sera euclidien.

14. Définition 1.179.
15. Définition 3.57.
16. 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.
17. Voir proposition 3.3.

3.7 Anneau principal et idéal premier

Définition 3.74.
*Un idéal I dans A est **principal à gauche** si il existe $a \in I$ tel que $I = Aa$. Il est **principal à droite** si il existe $a \in I$ tel que $I = aA$. Nous disons qu'il est **principal** si il est principal à gauche et à droite.*

Définition 3.75.
*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 3.105) et nous utilisons la proposition 3.107 qui dit qu'un anneau euclidien est principal.

Une manière de prouver qu'un anneau n'est pas principal est de prouver qu'il n'est pas factoriel, théorème 3.98.

Définition 3.76.
*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 $b \in I$.*

Lemme 3.77.
L'idéal nul (réduit à $\{0\}$) est premier si et seulement si A est intègre.

Démonstration. En deux sens.

(i) **Si $\{0\}$ est premier** Alors $A \neq \{0\}$ parce que $I = \{0\}$ est propre (définition d'idéal premier). De plus, si $ab = 0$, alors $ab \in I$ qui est un idéal premier. Donc soit a soit b est dans I, c'est-à-dire que soit a soit b est nul. Donc A est intègre.

(ii) **Si A est intègre** Alors $A \neq \{0\}$ et l'idéal $I = \{0\}$ est strictement inclus dans A. Si $ab \in I$, alors $ab = 0$ et comme A est intègre, soit a soit b est nul, c'est-à-dire appartient à I.

□

Proposition 3.78 ([77]).
Soit un anneau commutatif[18] *et un idéal I dans A.*

(1) I est un idéal premier si et seulement si A/I est un anneau intègre.

(2) I est un idéal maximal si et seulement si A/I est un corps.

(3) Tout idéal maximal propre est premier.

Démonstration. En plein d'étapes.

(i) **I premier implique A/I intègre** Évacuons le cas trivial pour être sûr. Si $I = \{0\}$ alors A est intègre par le lemme 3.77. Donc $A/I = A/\{0\} = A$ est intègre également.

Soient $a, b \in A$ tels que $[a][b] = [0]$. Donc $[ab] = [0]$, c'est-à-dire $ab \in I$. Puisque I est un idéal premier nous avons $a \in I$ ou $b \in I$, c'est-à-dire $[a] = 0$ ou $[b] = 0$; nous en déduisons que A/I est un anneau intègre.

(ii) **A/I intègre implique I premier** Soit $ab \in I$. Alors $[ab] = 0$, ce qui signifie que $[a][b] = 0$ donc que $[a] = 0$ ou que $[b] = 0$ parce que A/I est intègre. Mais la condition $[a] = 0$ signifie $a \in I$, et $[b] = 0$ signifie $b \in I$. Nous avons donc prouvé que soit a soit b est dans I, c'est-à-dire que I est premier.

(iii) **Si I est un idéal maximum** Nous devons montrer que tout élément non nul de A/I est inversible. Un élément non nul de A/I est $[x]$ avec $x \in A \backslash I$.

Nous considérons $J = Ax + I$, qui est un idéal parce que pour tout $a \in A$, $aAx + aI \in Ax + I$. Mais comme I est maximal, $J = I$ ou $J = A$.

18. Tous les anneaux du Frido sont commutatifs

Si $J = I$, nous aurions que pour tout $a \in A$ et pour tout $i \in I$, $ax + i \in I$. En particulier pour $a = 1$ et $i = 0$ nous aurions $x \in I$, ce qui est contraire à l'hypothèse faite sur x.

Donc $J = A$. En particulier, $1 \in J$, c'est-à-dire qu'il existe $a \in A$ et $i \in I$ tels que $ax + i = 1$. En passant aux classes, $[ax] = 1$, c'est-à-dire $[a][x] = 1$ qui signifie que $[a]$ est un inverse de $[x]$ dans A/I.

(iv) **Si A/I est un corps** Si $x \in A \backslash I$, il faut prouver que tout idéal contenant I et x est A. Un idéal contenant I et x doit contenir l'idéal $J = Ax + I$. Comme $x \notin I$, nous avons $[x] \neq 0$ dans A/I. Donc $[x]$ est inversible et il existe $a \in A$ tel que $[ax] = [A]$. C'est-à-dire que $ax - 1 \in I$. Nous avons alors

$$1 = ax + \underbrace{(1 - ax)}_{\in I}. \tag{3.95}$$

C'est-à-dire que $1 \in Ax + I$ et donc $Ax + I = A$.

Enfin nous prouvons que tout idéal maximal propre est premier.

Si I est maximal, A/I est un corps par le point (2), et vu que I est propre, le corps A/I n'est pas réduit à $\{0\}$. Donc le lemme 1.242 dit que A/I est un anneau intègre. Le point (1) dit alors que I est un idéal premier. □

Remarque 3.79.
Puisqu'un corps peut être réduit à $\{0\}$, dans (2), l'idéal peut être A. Mais pas dans (3), parce qu'un idéal premier est propre, ça fait partie de la définition 3.76.

Proposition 3.80 ([93])**.**
Si A est un anneau commutatif intègre, alors un idéal I dans A est premier si et seulement si A/I est intègre.

Démonstration. Supposons que I soit un idéal premier. Si $[a], [b] \in A/I$ sont tels que $[a][b] = 0$, alors $[ab] = 0$, ce qui signifie que $ab \in I$. Mais alors, puisque I est premier, soit a soit b est dans I. Cela signifie que soit $[a]$ soit $[b]$ est nul dans A/I. Cela prouve que A/I est un anneau intègre.

Dans l'autre sens, nous supposons que A/I est intègre. Cela implique immédiatement que $I \neq A$ parce que A/A n'est pas un anneau intègre (tout le monde est évidemment diviseur de zéro).

Soient donc $a, b \in A$ tels que $ab \in I$. Alors $[a][b] = [ab] = 0$ dans A/I, mais comme A/I est intègre, cela implique que soit $[a]$ soit $[b]$ est nul. Autrement dit, soit a soit b est dans I. □

Proposition 3.81 (Thème 16, [1])**.**
Soit A un anneau principal[19] *qui n'est pas un corps. Pour un idéal propre I de A, les conditions suivantes sont équivalentes :*

(1) I est un idéal maximal[20] *;*

(2) I est un idéal premier non nul[21] *;*

(3) il existe p irréductible[22] *dans A tel que $I = (p)$.*

Démonstration. En plusieurs implications.

(i) **(1) implique (2)** Par hypothèse, I est un idéal propre, de plus il n'est pas égal à $\{0\}$, parce que lorsque A et $\{0\}$ sont les seuls idéaux, nous avons un corps (proposition 3.50). Étant donné que I est un idéal maximal, le quotient A/I est un corps par la proposition 3.52.

Soient maintenant, pour entrer dans le vif du sujet, des éléments $a, b \in A$ tels que $ab \in I$. Dans le corps A/I nous avons $[ab] = 0$, et par définition du produit dans le quotient, $[a][b] = 0$. Par intégrité de l'anneau A/I (un corps est un anneau intègre, lemme 1.245) nous avons soit $[a] = 0$, soit $[b] = 0$, soit les deux en même temps. Dans tous les cas, soit a soit b est dans I.

19. Définition 3.75.
20. Définition 3.51.
21. Définition 3.76.
22. Définition 3.67.

(ii) **(2) implique (3)** Maintenant I est un idéal premier non réduit à $\{0\}$. Puisque A est un anneau principal, il existe $x \in A$ tel que $I = (x)$. Nous devons prouver que x peut être choisi irréductible ; et nous allons faire plus : nous allons prouver que x ne peut être que irréductible [23].

Supposons que x ne soit pas irréductible. Alors il existe $a, b \in A$ non inversibles tels que $x = ab$. Si $a \in (x)$ alors il existe $k \in A$ tel que $a = xk$, et en particulier, $a = abk$, c'est-à-dire $1 = bk$ (parce que A est principal et donc intègre). Cela signifie que b est inversible alors que nous avions dit qu'il ne l'était pas. Nous en déduisons que a n'est pas dans (x). On montre de manière similaire que b n'est pas dans (x) non plus.

Nous nous retrouvons donc avec $a, b \in A$ tel que $ab \in I$ sans que ni a ni b ne soient dans I. Cela contredit le fait que I soit un idéal premier. En conclusion, x est irréductible.

(iii) **(3) implique (1)** Nous avons $I = (p)$ avec p irréductible dans A. Supposons que J est un idéal différent de A contenant I. Comme A est principal, il existe $y \in A$ tel que $J = (y)$. En particulier $p \in J$, donc $p = ay$ pour un certain $a \in A$. Mais p est irréductible, donc soit a est inversible, soit y est inversible. Si y est inversible, alors $J = A$, ce qui est exclu. Si a est inversible, alors $(y) = (p)$, et $I = J$.

□

3.82.
Dans la proposition 3.81, l'hypothèse d'idéal propre est importante. En effet dans le cas $I = A$, nous avons évidemment que I est un idéal maximum. Mais A n'est d'abord pas un idéal premier parce qu'un idéal premier doit être strictement inclus dans l'anneau. Et ensuite, A est en général loin d'être garanti d'être égal à (p) pour un de ses éléments p.

Proposition 3.83.
Soit A un anneau principal, et soit $p \in A$ un élément irréductible. Alors

(1) (p) est un idéal maximum.

(2) $A/(p)$ est un corps.

Démonstration. Nous notons $I = (p)$. Soit un idéal J contenant I. Comme A est principal, J est monogène : $J = (q)$. Mais comme p est dans I qui est dans J, il existe $a \in A$ tel que $p = qa$.

Puisque p est irréductible, soit q, soit a est inversible. Si q est inversible, alors $J = A$. Si a est inversible, alors nous avons $p = qa$, donc $q = pa^{-1}$, ce qui signifie que $q \in (p)$ et donc que $J = I$.

Cela prouve que (p) est un idéal maximum.

Le fait que $A/(p)$ soit un corps est maintenant la proposition 3.52. □

Exemple 3.84.
L'anneau $\mathbb{Z}$ est principal parce qu'il est intègre et que ses seuls idéaux sont les $n\mathbb{Z}$ qui sont principaux : $n\mathbb{Z}$ est engendré par n. △

Exemple 3.85 (Les idéaux de $\mathbb{Z}/n\mathbb{Z}$)**.**
Les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont principaux, mais l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'est pas principal lorsque n n'est pas premier. Nous allons voir ça.

(i) **Les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont principaux** Soit un idéal S dans $\mathbb{Z}/n\mathbb{Z}$. Nous considérons la projection canonique $\phi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. La proposition 3.49 dit que $S = \phi(J)$ où J est un idéal de $\mathbb{Z}$ contenant $n\mathbb{Z}$. Mais le corolaire 1.235 nous dit qu'alors $J = m\mathbb{Z}$ pour un certain m. Pour que $m\mathbb{Z}$ contienne $n\mathbb{Z}$, il faut que m divise n.

Bref, $S = \phi(m\mathbb{Z})$ avec $m \mid n$. Nous montrons maintenant que S est engendré par $[m]_n$. D'abord, l'élément $[m]_n$ est bien dans $\phi(m\mathbb{Z})$. Ensuite un élément de $\phi(m\mathbb{Z})$ est de la forme

$$[km]_n = k[m]_n \in ([m]_n). \tag{3.96}$$

23. ça me semble un peu trop facile. Lisez attentivement, et écrivez-moi pour dire si vous êtes d'accord ou pas.

Donc $S \subset ([m]_n)$. Et l'inclusion dans l'autre sens est tout aussi immédiate : un élément de $([m]_n)$ est de la forme

$$k[m]_n = [km]_n = \phi(km) \in \phi(m\mathbb{Z}). \tag{3.97}$$

(ii) **Si n n'est pas premier, $\mathbb{Z}/n\mathbb{Z}$ n'est pas principal** Le fait est que lorsque n n'est pas premier, $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre (corolaire 1.248).

(iii) **Moralité** Un anneau comme $\mathbb{Z}/6\mathbb{Z}$ est un anneau dont tous les idéaux sont principaux, mais qui n'est pas principal.

△

Exemple 3.86.
Nous verrons dans la proposition 26.23 que l'anneau des fonctions holomorphes sur un compact de $\mathbb{C}$ est principal. △

Définition 3.87.
Nous disons que deux éléments d'un anneau principal sont ***premiers entre eux*** *si leur PGCD est* 1.

Théorème 3.88.
Si A est un anneau principal et si p et q sont premiers entre eux dans A, alors on a l'isomorphisme d'anneaux

$$A/pqA \simeq A/pA \times A/qA. \tag{3.98}$$

3.7.1 Autour du théorème de Bézout

Rappel de notations : si A est un anneau et si $p \in A$, nous notons (p) l'idéal engendré par p, c'est à dire pA. C'est la définition 3.44.

Lemme 3.89 ([94])**.**
Soient un anneau principal[24] *A ainsi que $a, b \in A$.*

(1) d est PGCD de $\{a, b\}$ si et seulement si $(d) = (a) + (b)$.

(2) m est PPCM de $\{a, b\}$ si et seulement si $(m) = (a) \cap (b)$.

Théorème 3.90 ([95, 94])**.**
Soient un anneau principal et $S \subset A$ non vide.

(1) d est un PGCD de S si et seulement si il génère l'idéal engendré par S.

(2) m est PPCM de S si et seulement si $(m) = \bigcap_{s\in S}(s)$.

Toute partie d'un anneau principal admet un PGCD et un PPCM[25].

¡¡ Avertissement/question au lecteur !! 3.91
La preuve est à revoir.

Démonstration. Puisque 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

$$(\delta) = \sum_{s\in S}(s) \tag{3.99a}$$

$$(\mu) = \bigcap_{s\in S}(s) \tag{3.99b}$$

(i) **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

$$\sum_{x\in S}(x) \subset (d), \tag{3.100}$$

24. Définition 3.75.
25. Ce n'est pas aussi trivial que ça parce qu'il faut encore qu'il existe des éléments vérifiant les formules données.

puis $(\delta) \subset (d)$ et finalement $d \mid \delta$.

(ii) **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 $\mu \mid m$.

□

Corolaire 3.92 (Théorème de Bézout[95])**.**
Soit un anneau principal A. Deux éléments $a, b \in A$ sont premiers entre eux si et seulement si il existe un couple $(u, v) \in A^2$ tel que

$$ua + vb = 1. \tag{3.101}$$

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

$$1 \in \operatorname{pgcd}(a, b) = \sum_{x=a,b} (x) = (a) + (b). \tag{3.102}$$

À l'inverse, si nous avons $ua + vb = 1$, alors $1 \in (a) + (b)$, mais puisque $(a) + (b)$ est un idéal principal, $(1) = (a) + (b)$ et donc $1 \in \operatorname{pgcd}(a, b)$. □

Le lemme de Gauss est une application immédiate du théorème de Bézout. Il y aura aussi un lemme de Gauss à propos de polynômes (lemme 6.48), et une généralisation directe au théorème de Gauss, théorème 6.47.

Lemme 3.93 (Lemme de Gauss[96])**.**
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. Comme a est premier avec c, nous avons $\operatorname{pgcd}(a, c) = 1$ et Bézout (1.228) nous donne donc $s, t \in A$ tels que $sa + tc = 1$. En multipliant par b,

$$sab + tbc = b. \tag{3.103}$$

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 $A = \mathbb{N}^*$.

3.7.2 Idéal premier

Proposition 3.94 ([97, 1], thème 21)**.**
Soit un anneau principal A et un élément $p \neq 0$ dans A. Nous avons équivalence de :

(1) (p) est un idéal premier,

(2) p est un élément premier,

(3) p est un élément irréductible,

(4) (p) est un idéal maximum propre[26].

Démonstration. En plusieurs implications.

(i) **(1) implique (2)** En plusieurs points.

26. Ce « propre » n'est pas dans l'énoncé sur Wikipédia. Je ne comprends pas pourquoi, et j'ai posé la question sur la page de discussion.
https://fr.wikipedia.org/wiki/Discussion:Idéal_premier

— La condition $p \neq 0$ est dans les hypothèses de la proposition.

— Si p était inversible, nous aurions $(p) = A$ et donc pas que (p) est un idéal premier.

— Soient $a, b \in A$ tels que $p \mid ab$. En particulier, $(ab) \in (p)$. Mais comme (p) est un idéal premier, cela implique soit $a \in (p)$ soit $b \in (p)$. Donc soit p divise a soit p divise b.

(ii) **(2) implique (3)** Un anneau principal est intègre ; c'est dans la définition 3.75. Dans un anneau intègre, tout élément premier est irréductible, c'est la proposition 1.251.

(iii) **(3) implique (4)** Soit un idéal I contenant (p). Puisque A est principal, I est engendré par un seul élément ; soit $I = (a)$. Vu que $p \in I$, l'élément a divise p. Mais comme p est un élément premier, $a \mid p$ implique $a = p$ ou $a = 1$. Dans le premier cas, $I = (a) = (p)$, et dans le second cas, $I = (a) = (1) = A$. Donc (p) est bien un idéal maximum.

De plus l'idéal (p) est propre. En effet avoir $(p) = A$ dirait en particulier que $1 \in (p)$, c'est-à-dire qu'il existe $x \in A$ tel que $xp = 1$. Or p étant irréductible, il est non inversible.

(iv) **(4) implique (1)** C'est la proposition 3.78(3).

□

Un exemple d'élément premier non irréductible est $[4]_6$ dans l'anneau non principal $\mathbb{Z}/6\mathbb{Z}$. Voir 3.103 et le lemme 3.104.

3.7.3 Anneau noethérien

Définition 3.95.
*Un anneau est dit **noethérien** si toute suite croissante d'idéaux est stationnaire (à partir d'un certain rang).*

Montrer que tout anneau principal est noethérien est le premier pas pour montrer que tout anneau principal est factoriel.

Lemme 3.96.
Tout anneau principal[27] *est noethérien.*

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

$$J \subset J_N \subset J_n \subset J. \tag{3.104}$$

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. □

Exemple 3.97.
Il y a moyen d'avoir un anneau noetherien non principal. C'est le cas de $\mathbb{Z}/6\mathbb{Z}$ dont nous parlerons dans 3.104. △

Théorème 3.98 ([98])**.**
Tout anneau principal est factoriel.

Exemple 3.99 ($\mathbb{Z}[i\sqrt{5}]$ n'est ni factoriel ni principal)**.**
Puisque $(i\sqrt{5})^2 = -5$, les éléments de $\mathbb{Z}[i\sqrt{5}]$ sont les éléments de $\mathbb{C}$ de la forme $a + bi\sqrt{5}$ avec $a, b \in \mathbb{Z}$. Nous définissons la **norme** sur $\mathbb{Z}[i\sqrt{5}]$ par[28]

$$\begin{aligned} N\colon \mathbb{Z}[i\sqrt{5}] &\to \mathbb{N} \\ z &\mapsto z\bar{z}. \end{aligned} \tag{3.105}$$

27. Définition 3.75.
28. C'est le carré de la norme usuelle, mais c'est l'usage dans le milieu.

Le fait que ce soit à valeurs dans $\mathbb{N}$ est un simple calcul :

$$N(x+iy\sqrt{5}) = (x+iy\sqrt{5})(x-iy\sqrt{5}) = x^2+5y^2. \tag{3.106}$$

De plus N est multiplicative : $N(z_1z_2) = N(z_1)N(z_2)$.

Nous pouvons maintenant déterminer les inversibles de $\mathbb{Z}[i\sqrt{5}]$. Si α est inversible, alors il existe β tel que $\alpha\beta = 1$. Au niveau de la norme,

$$N(\alpha)N(\beta) = 1, \tag{3.107}$$

ce qui implique que $N(\alpha) = 1$. Or l'équation $x^2+5y^2 = 1$ dans $\mathbb{N}$ donne $y = 0$, $x = \pm 1$.

Au final, les inversibles de $\mathbb{Z}[i\sqrt{5}]$ sont ± 1.

L'anneau $\mathbb{Z}[i\sqrt{5}]$ n'est alors pas factoriel (définition 3.71) parce que

$$2\times 3 = (1+i\sqrt{5})(1-i\sqrt{5}). \tag{3.108}$$

Cela donne deux décompositions du nombre 6 en produit d'éléments non associés [29] (2 n'est associé qu'à 2 et -2) parce que les inversibles sont 1 et -1.

Le fait que $\mathbb{Z}[i\sqrt{5}]$ ne soit pas factoriel implique qu'il ne soit pas principal, théorème 3.98. △

3.8 Anneau $\mathbb{Z}/6\mathbb{Z}$

Nous allons donner quelques propriétés de cet anneau, et en particulier voir que dans cet anneau non intègre, la notion d'élément irréductible n'est pas très intéressante.

Voici pour commencer un calcul de la table de multiplication de $A = \mathbb{Z}/6\mathbb{Z}$. Pour les multiples de (par exemple) $[4]_6$ nous écrivons

$$1\times [4]_6 = [4_6] \tag{3.109}$$

et ensuite

$$2\times [4]_6 = [8]_6 = [2]_6, \tag{3.110}$$

puis

$$3\times [4]_6 = [2+4]_6 = [6]_6 = [0]_6, \tag{3.111}$$

et caetera. Le résultat est :

$\times$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[0]_6$	0	0	0	0	0	0
$[1]_6$	0	1	2	3	4	5
$[2]_6$	0	2	4	0	2	4
$[3]_6$	0	3	0	3	0	3
$[4]_6$	0	4	2	0	4	2
$[5]_6$	0	5	4	3	2	1

(3.112)

Pour ne pas alourdir, nous n'avons pas écrit $[x]_6$ partout au lieu de x.

3.100 (Inversibles)**.**
Les éléments inversibles de $\mathbb{Z}/6\mathbb{Z}$ sont ceux qui ont un $[1]_6$ dans leur table de multiplication. Ce sont donc

$$U(\mathbb{Z}/6\mathbb{Z}) = \{[1]_6, [5]_6\}. \tag{3.113}$$

3.101 (Diviseurs de zéro)**.**
Les diviseurs de zéro sont ceux qui ont un $[0]_6$ dans leur table de multiplication, c'est-à-dire

$$\{[2]_6, [3]_6, [4]_6\}. \tag{3.114}$$

29. Définition 3.57.

3.102 (Irréductibles)**.**
Les irréductibles sont ceux qui ne sont ni inversibles ni produit de deux éléments non inversibles. Les non inversibles sont :

$$\{[0]_6, [2]_6, [3]_6, [4]_6\}. \tag{3.115}$$

Ils sont candidats à être irréductibles. Les produits de ces éléments (on oublie les crochets) sont :

$$2 \times 2 = 4 \tag{3.116a}$$
$$2 \times 3 = 0 \tag{3.116b}$$
$$2 \times 4 = 2 \tag{3.116c}$$
$$3 \times 3 = 3 \tag{3.116d}$$
$$3 \times 4 = 0 \tag{3.116e}$$
$$4 \times 4 = 4. \tag{3.116f}$$

Donc $[0]_6$, $[2]_6$, $[3]_6$ et $[4]_6$ ne sont plus candidats à être irréductible. Bref, il ne reste aucun candidats.

L'anneau $\mathbb{Z}/6\mathbb{Z}$ n'a aucun élément irréductible.

3.103 (Éléments premiers)**.**
Les éléments non nuls et non inversibles sont 2, 3 et 4.

(i) **Pour 2** L'élément $[2]_6$ divise 2, 4 et 0.
- Les (a, b) tels que $ab = 2$ sont : $(1, 2)$, $(2, 4)$ et $(5, 4)$. L'élément 2 divise donc toujours a ou b.
- Les (a, b) tels que $ab = 4$ sont : $(1, 4)$, $(2, 5)$ et $(4, 4)$. L'élément 2 divise donc toujours a ou b.
- Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 2 divise donc toujours a ou b. En particulier, $[2]_6$ divise $[0]_6$; c'est important.

Donc $[2]_6$ est un élément premier.

(ii) **Pour 3** L'élément $[3]_6$ divise 3 et 0.
- Les (a, b) tels que $ab = 3$ sont : $(1, 3)$ et $(3, 5)$. L'élément 3 divise donc toujours a ou b.
- Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 3 divise donc toujours a ou b.

Donc $[3]_6$ est un élément premier. L'élément $[4]_6$ divise 4, 2 et 0.
- Les (a, b) tels que $ab = 4$ sont : $(1, 4)$, $(2, 5)$ et $(4, 4)$. L'élément 4 divise donc toujours a ou b.
- Les (a, b) tels que $ab = 2$ sont : $(1, 2)$, $(2, 4)$ et $(5, 4)$. L'élément 4 divise donc toujours a ou b.
- Les (a, b) tels que $ab = 0$ sont : $(0, x)$, $(3, 2)$ et $(4, 3)$. L'élément 4 divise donc toujours a ou b.

Donc $[4]_6$ est un élément premier.

Au final, les éléments premiers dans $\mathbb{Z}/6\mathbb{Z}$ sont

$$\{[2]_6, [3]_6, [4]_6\}. \tag{3.117}$$

Vous noterez que $\mathbb{Z}/6\mathbb{Z}$ a des éléments premiers non irréductibles. Cela est un contre-exemple à la proposition 3.94 dans le cas d'un anneau non-intègre.

Lemme 3.104 ([1])**.**
L'anneau $\mathbb{Z}/6\mathbb{Z}$ est noetherien, mais ni intègre ni principal[30].

Démonstration. Comme c'est un anneau fini, toute suite croissante de quoi que ce soit devient stationnaire ; donc $\mathbb{Z}/6\mathbb{Z}$ est noetherien.

Puisque $\mathbb{Z}/6\mathbb{Z}$ a des diviseurs de zéro, il n'est pas intègre. Et s'il n'est pas intègre, il n'est pas factoriel non plus. □

30. Toutes les définitions dans le thème 21.

3.9 Anneau euclidien

Définition 3.105 (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

$$a = qb + r \tag{3.118}$$

et $\alpha(r) < \alpha(b)$.

(2) Pour tout $a, b \in A\backslash\{0\}$, $\alpha(b) \leqslant \alpha(ab)$.

Un anneau est ***euclidien*** *si 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 3.106.
Le stathme de $\mathbb{N}$ pour la division euclidienne usuelle est $\alpha(n) = n$. Si $a, b \in \mathbb{N}$ nous écrivons

$$a = qb + r \tag{3.119}$$

où q est l'entier le plus proche *inférieur* à a/b (on veut que le reste soit positif) et $r = a - qb$. Nous avons donc

$$r - b = a - b(q + 1) < a - b\frac{a}{b} = 0, \tag{3.120}$$

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.

Proposition 3.107 ([99]).
Tout 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[31] I.

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. □

Proposition 3.108.
L'anneau $\mathbb{Z}$ est principal et euclidien.

Démonstration. Nous allons seulement montrer que $\alpha(x) = |x|$ est un stathme euclidien. Ainsi $\mathbb{Z}$ sera euclidien et donc principal par la proposition 3.107.

D'abord $\mathbb{Z}$ est intègre, c'est l'exemple 1.246.

La condition $\alpha(b) \leqslant \alpha(ab)$ est immédiate : $|a| \leqslant |ab|$ pour tout $a, b \in \mathbb{Z}$.

Soient maintenant $a, b \in \mathbb{Z}$. Nous définissons $q_0, r_0 \in \mathbb{N}$ tels que

$$|a| = q_0|b| + r_0 \tag{3.121}$$

avec $r_0 < |b|$. Cela existe parce que $\alpha(x) = x$ est un stathme sur $\mathbb{N}$ par l'exemple 3.106.

(i) **Si $a > 0$ et $b > 0$** Alors $a = q_0b + r_0$ et le couple (q_0, r_0) vérifie les conditions de la définition 3.105(1).

31. Un tel élément existe...

(ii) **Si $a > 0$ et $b < 0$** Alors $a = -q_0 b + r_0$, et le couple $(-q_0, r_0)$ vérifie les conditions de la définition 3.105(1).

(iii) **Si $a < 0$ et $b > 0$** Alors $a = -q_0 b - r_0$, et le couple $(-q_0, -r_0)$ vérifie les conditions de la définition 3.105(1) parce que

$$\alpha(-r_0) = r_0 < |b| = \alpha(b). \tag{3.122}$$

(iv) **Si $a < 0$ et $b < 0$** Alors $a = q_0 b - r_0$, et le couple $(q_0, -r_0)$ vérifie les conditions de la définition 3.105(1).

□

Nous venons de voir que $\mathbb{Z}$ est principal ; le lemme suivant nous dit que $\mathbb{Z}[X]$ n'est lui, pas principal.

Lemme 3.109 ([100]).
Si A est un anneau intègre[32] *qui n'est pas un corps, alors $A[X]$ n'est pas principal.*

Démonstration. Soit un élément non nul $a \in A$.

(i) **Un idéal principal contenant a et X est $\mathbf{A[X]}$** Soit (P) un idéal principal contenant a et X. Puisque $a \in (P)$, il existe Q tel que $a = QP$. Donc P divise a dans $\mathbb{Z}[X]$. L'égalité des degrés indique que P est un polynôme constant, c'est-à-dire en réalité un élément de A. Soit $P = k \in A$.

Comme P divise X, nous avons aussi $X = kQ$ pour un certain $Q \in \mathbb{Z}[X]$. L'égalité des degrés dit qu'il existe $k' \in A$ tel que $Q = k'X$ et donc $Q = k'X = k'kQ$, ce qui implique que $kk' = 1$. L'idéal engendré par k contient donc en particulier $kk' = 1$ et donc contient $A[X]$ en entier :

$$1 = k'k \in k'(P) = (P). \tag{3.123}$$

(ii) **Si $(a, X) = A[X]$ alors a est inversible** Si $(a, X) = A[X]$, en particulier, $1 \in (a, X)$, ce qui signifie qu'il existe des polynômes $U, V \in A[X]$ tels que

$$1 = UX + Va. \tag{3.124}$$

Nous évaluons cette égalité en 0 : comme $(UX)(0) = 0$, nous avons $1 = V(0)a$, ce qui signifie que $V(0)$ est un inverse de A. Donc a est inversible.

(iii) **Si a n'est pas inversible alors (a, X) n'est pas principal** Si (a, X) était principal, alors nous aurions, par ce qui est dit plus haut, $(a, X) = A[X]$. Mais cette dernière égalité impliquerait que a soit inversible.

En conclusion, si A n'est pas un corps, il possède un élément ni nul ni inversible. Dans ce cas, l'idéal (a, X) n'est pas principal dans $A[X]$ et nous en déduisons que $A[X]$ n'est pas un anneau principal. □

Nous verrons dans le lemme 3.132 que si $\mathbb{K}$ est un corps, alors $\mathbb{K}[X]$ est principal.

Exemple 3.110.
Prouvons que $\mathbb{Z}[i\sqrt{2}]$ est un anneau euclidien. Pour cela nous démontrons que

$$\begin{aligned} N\colon \mathbb{Z}[i\sqrt{2}] &\to \mathbb{N} \\ a + bi\sqrt{2} &\mapsto a^2 + 2b^2 \end{aligned} \tag{3.125}$$

est un stathme euclidien.

32. Définition 1.221.

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

$$\frac{z}{t} = \alpha + \beta i\sqrt{2}. \tag{3.126}$$

Nous désignons par $\alpha + \epsilon_1$ et $\beta + \epsilon_2$ les entiers les plus proches de α et β. Nous avons $|\epsilon_1|, |\epsilon_2| \leqslant \frac{1}{2}$. Nous posons alors naturellement

$$q = (\alpha + \epsilon_1) + (\beta + \epsilon_2)i\sqrt{2} \tag{3.127}$$

et nous calculons $r = z - qt$:

$$2b'\epsilon_2 - a'\epsilon_1 + i\sqrt{2}\big(\epsilon_1 b' - a'\epsilon_2\big). \tag{3.128}$$

Nous trouvons

$$N(r) = a'^2\epsilon_1^2 + 4b'^2\epsilon_2^2 + 2a'^2\epsilon_1^2 + 2b'^2\epsilon_2^2 \leqslant \frac{3}{4}a'^2 + \frac{3}{2}b'^2. \tag{3.129}$$

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

$$N(zt) = (a^2 + 2b^2)(a'^2 + 2b'^2), \tag{3.130}$$

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 3.111 (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 3.57) si et seulement si $a = \pm b$.

De plus si p est irréductible[33], 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 parmi chaque paire. Tout élément x de $\mathbb{Z}[i\sqrt{2}]$ peut alors être écrit $x = \pm p_1^{\alpha_1} \dots p_n^{\alpha_n}$. Cette écriture va être pratique pour comparer des décompositions en facteurs irréductibles d'éléments. △

Le lemme suivant fait en pratique partie de l'exemple 3.114, mais nous l'isolons pour plus de clarté[34].

Lemme 3.112.
Si a et b sont deux éléments premiers entre eux de $\mathbb{Z}[i\sqrt{2}]$, et si 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 3.111 nous pouvons écrire

$$y = \pm p_1^{\sigma_1} \dots p_n^{\sigma_n} \tag{3.131a}$$
$$a = \pm p_1^{\alpha_1} \dots p_n^{\alpha_n} \tag{3.131b}$$
$$b = \pm p_1^{\beta_1} \dots p_n^{\beta_n} \tag{3.131c}$$

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 $\beta_i = 3\sigma_i$ (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. □

33. Définition 3.67
34. Merci à Marvoir pour m'avoir souligné le manque.

3.9.1 Équations diophantiennes

Exemple 3.113.
L'équation diophantienne

$$x^2 = 3y^2 + 8 \tag{3.132}$$

n'a pas de solution. En effet si nous prenons l'équation modulo 3 nous obtenons

$$[x^2]_3 = [3y^2 + 8]_3 = [8]_3 = [2]_3. \tag{3.133}$$

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 3.114.
Résolvons l'équation diophantienne

$$x^2 + 2 = y^3. \tag{3.134}$$

Une première remarque est que x doit être impair. En effet si $x = 2k$, nous devons avoir y^3 pair. Mais un cube pair est divisible par 8. Donc $y^3 = 8l$ pour un certain l. L'équation devient $4k^2 + 2 = 8l$, c'est-à-dire $2k^2 + 1 = 4l$. 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 (3.125).

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 $(-i\sqrt{2})$. Alors nous devrions avoir $d = (i\sqrt{2})^\beta$ et donc

$$x = (i\sqrt{2})^\beta q \tag{3.135}$$

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 $d = 1$.

Comme 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 3.112). 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

$$\begin{cases} x = a^3 - 6ab^2 & (3.136a) \\ 1 = 3a^2b - 2b^3 & (3.136b) \end{cases}$$

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

$$(x, z) = (-5, 1 + i\sqrt{2}) \tag{3.137a}$$
$$(x, z) = (5, -1 + i\sqrt{2}) \tag{3.137b}$$
$$\tag{3.137c}$$

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 $(-5, 3)$. △

3.9.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$

Définition 3.115.
*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 3.116.
*On dit qu'un triplet pythagoricien est **primitif** si les trois nombres sont premiers dans leur ensemble*[35].

Remarquons que c'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[36], donc les nombres x, y et z sont premiers deux à deux.

Lemme 3.117.
Dans un triplet pythagoricien primitif (x, y, z), on a toujours z impair et :

- *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 3.118 (Triplets pythagoriciens[101, 102])**.**
Un triplet $(x, y, z) \in (\mathbb{N}^*)^3$ est solution de $x^2 + y^2 = z^2$ si et seulement si il existe $d \in \mathbb{N}$ et $u, v \in \mathbb{N}^*$ premiers entre eux tels que

$$\begin{cases} x = d(u^2 - v^2) & (3.138a) \\ y = 2duv & (3.138b) \\ z = d(u^2 + v^2) & (3.138c) \end{cases}$$

ou

$$\begin{cases} x = 2duv & (3.139a) \\ y = d(u^2 - v^2) & (3.139b) \\ z = d(u^2 + v^2) & (3.139c) \end{cases}$$

La différence entre les deux est seulement d'inverser les rôles de x et y.

35. Définition 1.227.
36. Parce que k et k^2 ont les mêmes facteurs premiers.

Démonstration. Montrons d'abord que les formules proposées sont bien des solutions ; nous vérifions (3.138) :

$$x^2 + y^2 = d^2(u^2 - v^2)^2 + 4d^2u^2v^2 = d^2(u^2 + v^2)^2, \tag{3.140}$$

qui correspond bien au z^2 proposé.

Une vérification du même style fonctionne pour (3.139).

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 [37], grâce au lemme 3.117, nous pouvons supposer x pair tandis que y et z seront impairs. Comme $x^2 = (z + y)(z - y)$, nous avons

$$\left(\frac{x}{2}\right)^2 = \left(\frac{z + y}{4}\right)\left(\frac{z - y}{4}\right). \tag{3.141}$$

Puisque z et y sont premiers entre eux, les nombres $z - y$ et $z + y$ sont également premiers entre eux [38]. Donc les facteurs premiers (qui arrivent au moins au carré) de $(x/2)$ sont chacun soit dans $(z + y)/4$ soit dans $(z - y)/4$. Nous en déduisons que ces derniers sont des carrés d'entiers. Nous posons

$$\frac{z - y}{4} = u^2, \quad \frac{z + y}{4} = v^2. \tag{3.142}$$

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 :

$$\begin{cases} x = 2uv & (3.143a) \\ y = v^2 - u^2 & (3.143b) \\ z = u^2 + v^2, & (3.143c) \end{cases}$$

ce qu'il fallait. □

Remarque 3.119.
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 3.120 ([101])**.**
Les équations $x^4 + y^4 = z^2$ et $x^2 + y^4 = z^4$ n'ont pas de solution dans $(\mathbb{N}^*)^3$.

Démonstration. Si la première équation n'a pas de solution, alors la seconde n'en 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 [39]). Du coup, les trois nombres x, y et z sont premiers dans leur ensemble parce qu'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 3.118, il existe $u, v \in \mathbb{N}^*$ premiers entre eux tels que, sans perte de généralité [40], on ait

$$\begin{cases} \bar{x} = 2uv & (3.144a) \\ \bar{y} = u^2 - v^2 & (3.144b) \\ z = u^2 + v^2. & (3.144c) \end{cases}$$

37. En échangeant les rôles de x et y ici, nous obtenons à la fin la seconde forme des solutions.
38. Si $z - y = kn$ et $z + y = km$, faisant la somme et la différence on voit que y et z sont divisibles par k.
39. Voir quelque chose comme le lemme 1.136.
40. En inversant les rôles de x et y au besoin.

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 3.117 implique alors que v soit pair.

De l'équation 3.144b, 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 appliquons de nouveau la proposition 3.118, en se souvenant que v est pair : il existe donc deux nombres r et s premiers entre eux tels que

$$\begin{cases} v=2rs & (3.145a)\\ y=r^2-s^2 & (3.145b)\\ u=r^2+s^2. & (3.145c)\end{cases}$$

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 :

$$\begin{cases} r=\alpha^2 & (3.146a)\\ s=\beta^2 & (3.146b)\\ r^2+s^2=\gamma^2. & (3.146c)\end{cases}$$

En mettant les deux premiers dans la troisième équation, nous avons $\alpha^4+\beta^4=\gamma^2$. Donc $(\alpha^2,\beta^2,\gamma)$ est une solution. Nous allons montrer que $\gamma<z$, ce qui terminera la preuve, puisque z était supposé minimal. Nous avons :

$$\begin{aligned} z&=u^2+v^2 && \text{par 3.144c}\\ &=r^2+s^2+4r^2s^2 && \text{par 3.145}\\ &=\gamma^2+4r^2s^2\\ &>\gamma^2, \end{aligned}$$

et a fortiori $\gamma<z$. □

3.10 Polynômes à coefficients dans un anneau commutatif

Lemme 3.121.
Nous considérons un polynôme $P\in A[X]$, et le quotient $A[X]/(P)$. Pour tout polynôme $Q\in A[X]$ nous avons les égalités

$$Q(\bar{X})=\overline{Q(X)}=\bar{Q}. \tag{3.147}$$

Démonstration. Si $Q=\sum_k a_kX^k$, alors par la linéarité de la prise de classes,

$$\bar{Q}=\sum_k a_k\overline{X^k}. \tag{3.148}$$

Nous insistons sur le fait que cette égalité n'est rien d'autre que l'itération de la définition de la somme dans l'espace quotient : $\bar{x}+\bar{y}=\overline{x+y}$ ainsi que du produit $k\bar{x}=\overline{kx}$ (définition 1.214). Toujours par définition du produit appliqué à l'élément $\bar{X}$ nous avons $(\bar{X})^2=\overline{X^2}$; par récurrence $\overline{X^k}=\bar{X}^k$, et

$$\bar{Q}=\sum_k a_k\bar{X}^k=Q(\bar{X}). \tag{3.149}$$

Le fait que $\bar{Q}=\overline{Q(X)}$ n'est rien d'autre que le fait que dans $A[X]$ nous avons $Q=Q(X)$, comme expliqué dans le lemme 1.310. □

3.10.1 Monômes

3.122.
Les éléments de la forme λX^k avec $\lambda \in A$ et $k \in \mathbb{N}$ sont des **monômes**.

Nous allons aussi considérer

$$A_n[X] = \{P \in A[X] \text{ tel que } \deg(P) \leqslant n\}. \tag{3.150}$$

Cet ensemble est un sous-module libre.

3.10.2 Évaluation

Soit $P \in A[X]$. À priori, P n'est qu'une suite dans A indexée par $\mathbb{N}$.

Nous avons déjà défini son évaluation sur un élément $\alpha \in A$ dans la définition 1.306 :

$$P(\alpha) = \sum_k a_k \alpha^k. \tag{3.151}$$

Cette somme est toujours finie.

3.123.
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 à 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 6.104 et l'exemple 19.40.

Définition 3.124.
Soient un anneau A et un anneau B qui contient A (comme sous-anneau). Pour $\alpha \in B$ nous définissons $A[\alpha]_B$ comme étant l'intersection de tous les sous-anneaux de B contenant A.

Comme dit plus haut, nous nous permettons d'écrire $A[\alpha]$ sans préciser B lorsque ce dernier sera clair dans le contexte.

Proposition 3.125.
Soient un anneau A et un anneau B qui contient A (comme sous-anneau). Pour tout $\alpha \in B$ nous avons

$$A[\alpha] = \{P(\alpha) \text{ tel que } P \in A[X]\} \tag{3.152}$$

où encore une fois, $P(\alpha)$ est calculé dans B ; le contexte est clair là-dessus.

Démonstration. Si A' est un sous-anneau de B contenant A et α, alors A' contient tous les $P(\alpha)$ avec $P \in A[X]$. Nous avons donc

$$\{P(\alpha) \text{ tel que } P \in A[X]\} \subset A[\alpha]. \tag{3.153}$$

Par ailleurs, $\{P(\alpha) \text{ tel que } P \in A[X]\}$ est un sous-anneau de B contenant A et α. Donc $A[\alpha]$ y est inclus. □

3.10.3 Polynômes sur un anneau intègre

Théorème 3.126.
L'anneau A est intègre si et seulement si $A[X]$ est intègre.

Démonstration. Soient P et Q des éléments non nuls de $A[X]$. Puisque l'anneau A est intègre, nous avons

$$\deg(PQ) = \deg(P) + \deg(Q) \tag{3.154}$$

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 considéré comme sous-anneau de $A[X]$. □

3.127.
Si A n'est pas intègre, soient $\alpha, \beta \in A$ non nuls tels que $\alpha\beta = 0$. Le produit des polynômes $X \mapsto \alpha X$ et $X \mapsto \beta$ est $(\alpha X) \cdot (\beta) = 0$; 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.

Corolaire 3.128.
Si A est intègre, les inversibles de $A[X]$ sont les éléments inversibles de 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 degré zéro et il faut que $P, Q \in A$. Donc Q est un inversible de A. □

3.10.4 Division euclidienne

Le théorème suivant établit la **division euclidienne** dans $A[X]$ du polynôme P par un polynôme D.

Théorème 3.129.
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

$$P = QD + R \tag{3.155}$$

avec $\deg(R) < \deg(D)$.

Les polynômes Q et R sont déterminés de façon univoque par cette condition.

Définition 3.130.
*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 si il existe Q tel que $P = QD$.* [41]

3.131.
Le théorème 3.129 nous incite à utiliser le degré comme stathme euclidien sur $A[X]$ dès que A est un anneau intègre. Or cela ne fonctionne pas en général, parce que très peu de polynômes ont à priori un coefficient dominant inversible.

Lemme 3.132 (Thème 16)**.**
Si $\mathbb{K}$ est un corps [42]*, alors l'anneau $\mathbb{K}[X]$ est euclidien et principal.*

Démonstration. Puisque $\mathbb{K}$ est un corps, tous les éléments sont inversibles et le degré donne un stathme par le théorème 3.129. L'anneau $\mathbb{K}[X]$ est donc euclidien et par conséquent principal (proposition 3.107). □

Dans le théorème 6.40 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 3.133 ([103])**.**
*Soit un anneau A. Deux polynômes P et Q dans $A[X]$ sont dits **étrangers** entre eux si* 1 *est un pgcd* [43] *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.*

*Les polynômes P et Q sont **premiers entre eux** si les seuls diviseurs communs de P et Q sont les inversibles.*

41. Ceci se rapproche tout naturellement des notions générales de divisibilité dans un anneau intègre, vues en sous-section 3.5.1.
42. Définition 1.237.
43. Définition 1.215.

Les notions de polynômes étrangers entre eux ou de polynômes premiers entre eux ne sont pas identiques, comme le montre l'exemple suivant.

Exemple 3.134 ([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). △

3.10.5 Polynôme primitif

Définition 3.135.
*Le **contenu** du polynôme $P = \sum_i a_i X^i \in \mathbb{K}[X]$ est le pgcd de ses coefficients : $c(P) = \operatorname{pgcd}(a_i)$.*

Définition 3.136 (Ordre d'un polynôme)**.**
*Soit P un polynôme irréductible de degré n sur $\mathbb{F}_p[X]$. L'**ordre** de P est*

$$\min\{k \text{ tel que } P \mid X^k - 1\}. \tag{3.156}$$

Définition 3.137 (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 3.138 (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.*

3.139.
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 3.87 pour avoir des nombres premiers entre eux.

(2) Le contenu de ses coefficients est 1. Parce que le contenu est précisément le pgcd, définition 3.135.

La notion de polynôme primitif au sens du pgcd est particulière aux polynômes à coefficients dans un anneau comme le montre le lemme suivant.

Lemme 3.140.
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 par exemple le cas pour le corolaire 3.152.

3.10.6 Racines des polynômes

Définition 3.141.
*Soient 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 3.142.
Soient 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 6.94. 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[44] de P par μ : il existe des polynômes Q et R tels que

$$P = Q\mu + R \tag{3.157}$$

avec $\deg(R) < \deg(\mu)$. Donc R est une constante, élément de A : appelons-le a. En évaluant (3.157) en α, il vient

$$0 = P(\alpha) = Q(\alpha)\mu(\alpha) + a, \tag{3.158}$$

et nous en déduisons que $a = 0$, ce qui montre que $P = Q\mu$ et que μ divise P. □

Définition 3.143 (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.*

3.144.
Pour une définition générale d'une racine simple de l'équation $f(x) = 0$, voir la définition 34.50. La proposition 3.142 nous indique que toute racine est de multiplicité au moins égale à 1.

Proposition 3.145.
Soient un anneau A, un polynôme $P \in A[X]$ de degré n, ainsi qu'une racine $\alpha \in A$ de P. Alors il existe un polynôme $Q \in A[X]$ de degré $n - 1$ tel que $P = (X - \alpha)Q$.

Proposition 3.146.
L'élément $\alpha \in A$ est une racine de multiplicité h du polynôme P si et seulement si il existe $Q \in A[X]$ tel que $P = (X - \alpha)^h Q$ avec $Q(\alpha) \neq 0$.

Lemme 3.147.
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.

Dans le théorème suivant, la partie importante en pratique est le point (2) parce qu'il dit que, lorsque nous cherchons les racines d'un polynôme, nous pouvons nous arrêter lorsque nous en avons trouvé autant que le degré, multiplicité comprise.

Théorème 3.148.
Soit A un anneau intègre et $P \in A[X]\backslash\{0\}$, un polynôme de degré n.

(1) Si $\alpha_1, \dots, \alpha_p \in A$ sont des racines deux à deux distinctes de multiplicités $k_1, \dots, k_p$, 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.

(2) 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 $Q \in A[X]$ tel que $P = Q(X - \alpha)^k$. Il reste à voir que $Q(\alpha) \neq 0$. Cela est une conséquence de la proposition 3.142 : 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.

44. Théorème 3.129.

Nous supposons que $p \geqslant 2$ et nous effectuons une récurrence sur p. Nous considérons donc les $p-1$ premières racines $\alpha_1, \dots, \alpha_{p-1}$ et un polynôme $R \in A[X]$ tel que $R(\alpha_i) \neq 0$ pour $i = 1, \dots, p-1$ et

$$P = \underbrace{(X-\alpha_1)^{k_1} \dots (X-\alpha_{p-1})^{k_{p-1}}}_{S} R. \tag{3.159}$$

Par hypothèse $P(\alpha_p) = S(\alpha_p)R(\alpha_p) = 0$. L'anneau 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 3.147 afin de dire que le degré de α_p pour $P = SR$ est k_p. Par conséquent

$$R = (X-\alpha_p)^{k_p} T \tag{3.160}$$

avec $T(\alpha_p) \neq 0$ et enfin

$$P = \prod_{i=1}^{p} (X-\alpha_i) T. \tag{3.161}$$

De plus $T(\alpha_i) \neq 0$, sinon $R(\alpha_i)$ serait nul. □

Corolaire 3.149.
Un polynôme de degré n sur un anneau intègre possède au maximum n racines distinctes.

Démonstration. Le théorème 3.148(2) dit que la somme des multiplicités des racines de P est au maximum n. Mais la proposition 3.142 dit que toutes les racines ont une multiplicité au moins égale à un. Donc il ne peut pas y en avoir plus de n. □

Proposition 3.150 ([104, 105])**.**
Soit un anneau intègre et $n > 0$. Les racines n^e de l'unité forment un groupe cyclique dont l'ordre divise n.

Lemme 3.151.
Soit le polynôme $P = X^2 - 1$ sur l'anneau intègre[45] *A. Si $1 \neq -1$*[46]*, alors les racines de P sont ± 1.*

Démonstration. Le fait que ± 1 sont racines est le lemme 1.211(1)(4). Puisque par hypothèse $1 \neq -1$, le corolaire 3.149 termine la preuve. □

Corolaire 3.152 (Conséquence du lemme de Gauss[106])**.**
Soient A un anneau factoriel et $\mathrm{Frac}(A)$ *son corps des fractions. Un polynôme non constant $P \in A[X]$ est irréductible (sur A) si et seulement si il est irréductible et primitif au sens du pgcd*[47] *sur* $\mathrm{Frac}(A)[X]$.

Exemple 3.153.
Il ne faudrait pas croire qu'être irréductible dans un anneau A implique d'être irréductible dans le corps des fractions. En effet soit $A = \mathbb{Z}[\sqrt{5}]$ et $P = X^2 - X - 1$. Nous savons que sa factorisation est

$$P = \left(X - \frac{1+\sqrt{5}}{2}\right)\left(X - \frac{1-\sqrt{5}}{2}\right). \tag{3.162}$$

Si vous ne le saviez pas, faites juste le calcul pour vous en assurer.

Ce polynôme est irréductible sur $\mathbb{Z}[\sqrt{5}]$ mais pas irréductible sur $\mathrm{Frac}\left(\mathbb{Z}[\sqrt{5}]\right)$. △

45. Définition 1.221.
46. À mon avis cette hypothèse n'est pas nécessaire. Et d'ailleurs j'ai un peu du mal à voir des exemples exotiques d'anneaux intègres dans lesquels $1 = -1$. Il y aurait $[0,1]$ modulo 2, certaines de ses parties, comme $\mathbb{Z}/2\mathbb{Z}$. Et quoi d'autre ?
47. Définition 3.138.

3.10.7 Quelques identités

Lemme 3.154 ([1, 107])**.**
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. Plusieurs points.

(i) **Pour (1)** Nous allons utiliser le lemme 3.1 avec $a = 1$, $b = -X$ et $r = n - 1$. Notez que n étant impair, $r = n - 1$ est encore positif. En ce qui concerne le membre de gauche de (3.1), nous remarquons que $(-X)^n = (-1)^n X^n = -X^n$. Nous avons donc :

$$1 + X^n = (a - b)\sum_{k=0}^{n-1} a^{n-1-kb^k} = (1 + X)(\cdots). \tag{3.163}$$

Cela prouve que $1 + X$ divise $1 + X^n$.

(ii) **Pour (2)** Posons $P = X + \ldots + X^{n-1}$. Nous avons $XP = P - X + X^n$ et donc

$$(X - 1)(1 + P) = X + XP - 1 - P = X + (P - X + X^n) - 1 - P = X^n - 1. \tag{3.164}$$

(iii) **Pour (3)** Déjà fait dans (3.1).

□

Chapitre 4

Espaces vectoriels (début)

4.1 Parties libres, génératrices, bases et dimension

Nous avons déjà défini (dans 1.276) un espace vectoriel comme étant un module sur un corps commutatif. En explicitant un peu, cela donne ceci[108].

Un espace vectoriel sur le corps $\mathbb{K}$ est un ensemble E muni de deux opérations :

— une loi de composition interne $+ \colon E \times E \to E$,

— une loi de composition externe $\cdot \colon \mathbb{K} \times E \to E$

telles que

(1) $(E, +)$ soit un groupe abélien,

(2) pour tout $u, v \in E$ et pour tout $k, k' \in \mathbb{K}$,

(2a) $k(u + v) = (ku) + (kv)$

(2b) $(kk')u = k(k'u)$

(2c) $(k + k')u = (ku) + (k'u)$

(2d) $1u = u$

où 1 est le neutre de $\mathbb{K}$ et où nous avons directement adopté la notation ku pour $k \cdot u$.

Si $u \in E$, nous notons $-u$ l'inverse de u dans le groupe $(E, +)$.

Définition 4.1 (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 $\{u_i\}_{i=1,\dots,n}$ de A, l'égalité*

$$a_1 u_1 + \cdots + a_n u_n = 0 \tag{4.1}$$

implique $a_i = 0$ pour tout i (ici les a_i sont dans le corps de base).

Remarque 4.2.
Notons que le vecteur nul n'est dans aucune partie libre, ne fût-ce que parce que $a0 = 0$ n'implique pas $a = 0$.

Si A est une partie de l'espace vectoriel E, nous notons $\mathrm{Span}(A)$ 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 4.3 (Partie génératrice)**.**
Une partie B d'un espace vectoriel E est ***génératrice*** *si* $\mathrm{Span}(B) = E$.

Remarque 4.4.
Ces définitions demandent des commentaires en dimension infinie [1].

1. Nous n'avons pas encore défini le concept de dimension, mais nous nous adressons au lecteur trop pressé.

(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 infinie (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 4.5 (Base)**.**
*Une **base** de l'espace vectoriel E est une partie à la fois génératrice et libre.*

Nous prouvons à présent que tout élément non nul d'un espace vectoriel possédant une base [2] se décompose de façon unique en combinaison linéaire finie d'éléments d'une base.

Proposition 4.6 ([1])**.**
Soient un espace vectoriel E sur $\mathbb{K}$ muni d'une base $\{e_i\}_{i\in I}$. Si $v \in E$, alors il existe un unique couple (J, c) où

(1) J est une partie finie de I ;
(2) $c\colon J \to \mathbb{K}$ est une application
(3) $v = \sum_{j\in J} c(j)e_j$.

Les éléments $c(i)$ seront notés v_i et sont les composantes de v dans la base. Sous-entendu, on prolonge c de J vers I par zéro sur $I\backslash J$.

Démonstration. Soit un espace vectoriel E et une base $\{e_i\}_{i\in I}$ où I est un ensemble à priori quelconque. Soit $v \in E$. Puisque $E = \mathrm{Span}\{e_i\}_{i\in I}$, il existe une partie finie J de I et des coefficients $\{v_j\}_{j\in J}$ dans $\mathbb{K}$ tels que

$$v = \sum_{j\in J} v_j e_j. \tag{4.2}$$

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

$$v = \sum_{j\in J} v_j e_j = \sum_{k\in K} w_k e_k. \tag{4.3}$$

Nous posons $L = J \cup K$. Remarquez que les unions suivantes sont des unions disjointes :

$$L = (J\backslash K) \cup (J \cap K) \cup (K\backslash J) \tag{4.4a}$$
$$J = (J\backslash K) \cup (J \cap K) \tag{4.4b}$$
$$K = (K \cap J) \cup (K\backslash J). \tag{4.4c}$$

Écrivons $0 = v - v$ en utilisant les expressions de v de (4.3) et en décomposant les sommes :

$$0 = \sum_{j\in J} v_j e_j - \sum_{k\in K} w_k e_k \tag{4.5a}$$
$$= \sum_{j\in J\backslash K} v_j e_j + \sum_{j\in J\cap K} v_j e_j - \sum_{k\in K\cap J} w_k e_k - \sum_{k\in K\backslash J} w_k e_k \tag{4.5b}$$
$$= \sum_{j\in J\backslash K} v_j e_j + \sum_{j\in J\cap K} (v_j - w_j) e_j - \sum_{k\in K\backslash J} w_k e_k. \tag{4.5c}$$

Là-dessus, nous posons

$$\alpha_l = \begin{cases} v_l & \text{si } l \in J\backslash K \\ v_l - w_l & \text{si } l \in J \cap K \\ -w_l & \text{si } l \in K\backslash J. \end{cases} \tag{4.6}$$

2. Nous n'avons pas démontré que tout espace vectoriel possède une base. Donc à notre niveau, il est possible que ce théorème soit sans objet pour beaucoup d'espaces.

Nous avons alors

$$\sum_{l\in L}\alpha_l e_l = 0. \tag{4.7}$$

Vu que $\{e_i\}_{i\in I}$ est libre, la partie $\{e_l\}_{l\in L}$ est également libre. Donc l'équation (4.7) implique que $\alpha_l = 0$ pour tout $l \in L$.

Nous devons prouver que

$$\{j \in J \text{ tel que } v_j \neq 0\} = \{k \in K \text{ tel que } w_k \neq 0\} \tag{4.8}$$

et que pour l dans cet ensemble, $v_l = w_l$.

Soit $j \in J$; il y a deux possibilités : soit $j \in J\backslash K$, soit $j \in J \cap K$. 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 $k \in J \cap K$. □

Lemme 4.7 ([1])**.**
Soit un espace vectoriel admettant des bases. Un endomorphisme est une bijection si et seulement si il change toute base en une base.

Démonstration. En deux parties. Soit un espace vectoriel E possédant des bases et un endomorphisme $f\colon E \to E$.

(i) **Si f est bijective** Soit une base $\{v_i\}_{i\in I}$; nous devons voir que $\{f(v_i)\}_{i\in I}$ est une base.

(i) **Libre** Si J est une partie finie de I et si les λ_j sont des scalaires tels que $\sum_{j\in J}\lambda_j f(v_j) = 0$, alors

$$0 = \sum_{j\in J}\lambda_j f(v_j) = f\big(\sum_{j\in J}\lambda_j v_j\big). \tag{4.9}$$

Mais comme f est bijective, cela implique que $\sum_{j\in J}\lambda_j v_j = 0$. En retour, parce que $\{v_i\}$ est une base, cela implique que $\lambda_j = 0$ pour tout j.

(ii) **Générateur** Soit $x \in E$. Puisque f est bijective, il existe un unique $y \in E$ tel que $x = f(y)$. Comme $\{v_i\}_{i\in I}$ est une base, il existe une partie finie $J \subset I$ et des scalaires $\{\lambda_j\}_{j\in J}$ tels que

$$y = \sum_{j\in J}\lambda_j v_j. \tag{4.10}$$

Nous avons alors

$$x = f(y) = \sum_{j\in J}\lambda_j f(v_j), \tag{4.11}$$

qui montre que $\{f(v_i)\}_{i\in I}$ est bien génératrice de E

(ii) **Si f change les bases en bases** Soit un endomorphisme changeant toute base en une base. Nous devons prouver qu'il est bijectif.

(i) **Injective** Nous considérons une base $\{v_i\}_{i\in I}$. La partie $\{f(v_i)\}_{i\in I}$ est par hypothèse également une base.

Soient $x, y \in E$ tels que $f(x) = f(y)$. Il existe J et K finis dans I qui permettent de décomposer x et y respectivement dans la base $\{f(v_i)\}_{i\in I}$. Quitte à poser $J' = J \cup K$, nous supposons que J suffit[3]. Il existe donc des scalaires $\{\lambda_j\}_{j\in J}$ et $\{\mu_j\}_{j\in J}$ tels que $x = \sum_{j\in J}\lambda_j f(v_j)$ et $y = \sum_{j\in J}\mu_j f(v_j)$.

La relation $f(x) = f(y)$ donne immédiatement, par la linéarité de f,

$$\sum_{j\in J}(\lambda_j - \mu_j)f(v_j) = 0. \tag{4.12}$$

Du fait que $\{f(v_i)\}_{i\in I}$ soit une base, nous déduisons que $\lambda_j - \mu_j = 0$ pour tout j. Donc $x = y$, et f est injective.

3. Nous utilisons le fait que l'union de deux parties finies d'un ensemble est finie (lemme 1.122(2)).

(ii) **Surjective** Soit $x \in E$. Puisque $\{f(v_i)\}_{i\in I}$ est une base, il existe des scalaires λ_j tels que

$$x = \sum_{j\in J} \lambda_j f(v_j) = f\big(\sum_{j\in J} \lambda_j v_j\big). \tag{4.13}$$

Donc f est surjective.

□

Lemme 4.8 ([1])**.**
Bases dans $\mathbb{R}^n$.

(1) Si v et w ne sont pas colinéaire dans $\mathbb{R}^2$, alors $\{v, w\}$ est une base de $\mathbb{R}^2$.

(2) Toute partie libre de $\mathbb{R}^n$ contenant n élément est une base.

Définition 4.9.
*Un espace vectoriel est **de type fini** si il contient une partie génératrice finie.*

Nous verrons dans les résultats qui suivent que cette définition est en réalité inutile parce qu'un espace vectoriel sera de type fini si et seulement si il est de dimension finie.

Lemme 4.10.
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. Pour $n = 1$, nous avons $E = \mathrm{Span}(e)$ et donc si $v_1, v_2 \in E$ nous avons $v_1 = \lambda_1 e$, $v_2 = \lambda_2 e$ pour certains éléments non nuls λ_1, λ_2 du corps de base. Nous avons donc $\lambda_2 v_1 - \lambda_1 v_2 = 0$. Cela prouve que $\{v_1, v_2\}$ est liée.

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 $\lambda_{i,k}$ par

$$v_i = \sum_{k=1}^{n} \lambda_{i,k} e_k \tag{4.14}$$

Puisque

$$v_{n+1} = \sum_{k=1}^{n} \lambda_{n+1,k} e_k \neq 0, \tag{4.15}$$

quitte à changer la numérotation des e_i, nous pouvons supposer que $\lambda_{n+1,n} \neq 0$. Nous considérons les vecteurs

$$w_i = \lambda_{n+1,n} v_i - \lambda_{i,n} v_{n+1}. \tag{4.16}$$

En calculant un peu,

$$w_i = \lambda_{n+1,n} \sum_{k=1}^{n} \lambda_{i,k} e_k - \lambda_{i,n} \sum_{k=1}^{n} \lambda_{n+1,k} e_k \tag{4.17a}$$

$$= \sum_{k=1}^{n-1} \big(\lambda_{n+1,n}\lambda_{i,k} - \lambda_{i,n}\lambda_{n+1,k}\big) e_k \tag{4.17b}$$

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 $\mathrm{Span}\{e_1, \dots, e_{n-1}\}$; 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

$$0 = \sum_{i=1}^{n} \alpha_i w_i = \sum_{i=1}^{n} \alpha_i \lambda_{n+1,n} v_i - \sum_{i=1}^{n} \alpha_i \lambda_{i,n} v_{n+1}. \tag{4.18}$$

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 (4.18) est une combinaison linéaire nulle non triviale des vecteurs de $\{v_1, \dots, v_{n+1}\}$. Cette partie est donc liée. □

Lemme 4.11.
Soient L une partie libre et G une partie génératrice d'un espace vectoriel E. Si l'ensemble des parties libres L' telles que $L \subset L' \subset G$ possède un élément maximum[4]*, alors cet élément est une base.*

Qu'entend-on par « maximale » ? La partie B candidate, 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

$$\sum_{i=1}^{l} \lambda_i b_i + \lambda_x x = 0. \tag{4.19}$$

Si $\lambda_x = 0$ alors un de λ_i doit être non nul et l'équation (4.19) 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

$$x = -\frac{1}{\lambda_x} \sum_{i=1}^{l} \lambda_i b_i. \tag{4.20}$$

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 4.12 (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.

(4) Si V est un sous-espace vectoriel de E, et si L est une base de V, alors il existe une base de E qui contient L.

Démonstration. Point par point.

(1) Comme 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 4.10. 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 4.11.

(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 4.10 indique que $\mathrm{Card}(B') \leqslant \mathrm{Card}(B)$. Par symétrie on a l'inégalité inverse. Donc $\mathrm{Card}(B) = \mathrm{Card}(B')$.

(4) La partie L étant une base de V, elle est en particulier libre dans E. Par le point (2), L peut être étendue en une base.

4. Encore une fois, à part quelques cas triviaux, il n'est pas clair à ce point que ce maximum existe.

□

Remarque 4.13.
Le théorème de la base incomplète 4.12(2) est ce qui permet de construire une base d'un 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 4.14.
*La **dimension** d'un espace vectoriel de type fini est le cardinal*[5] *d'une*[6] *de ses bases.*

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.

Exemple 4.15.
La base de **canonique** de $\mathbb{R}^m$ est la partie $\{e_1, \dots, e_m\}$, où le vecteur e_j est

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{j-ème} \\ \\ \\ \\ \end{array} .$$

La composante numéro j de e_i est 1 si $i = j$ et 0 si $i \neq j$. Cela s'écrit $(e_i)_j = \delta_{i,j}$ où δ est le **symbole de Kronecker** défini par

$$\delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{4.21}$$

Les éléments de la base canonique de $\mathbb{R}^m$ peuvent donc être écrits $e_i = \sum_{k=1}^m \delta_{i,k} e_k$. △

Le théorème suivant est essentiellement une reformulation du théorème 4.12.

Théorème 4.16.
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.

Proposition 4.17.
Si E est un espace vectoriel de dimension finie n, alors

(1) toute partie contenant $n+1$ éléments est liée.

(2) toute partie libre contenant n éléments est une base,

(3) toute partie génératrice contenant n éléments est une base.

Démonstration. Soit une partie M contenant $n+1$ éléments. L'espace E possède une partie génératrice contenant n éléments (n'importe quelle base). Donc M est liée par le lemme 4.10.

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 4.10.

5. Définition 1.120.
6. Le théorème de la base incomplète 4.12(3) montre que cette définition ne souffre d'aucune ambiguïté.

Pour la dernière assertion, soit une partie génératrice $\{v_i\}_{i\in I}$ où I contient n éléments. Par le théorème 4.16(1) nous pouvons en extraire une base : il existe $J \subset I$ tel que $\{v_j\}_{j\in J}$ soit une base. Si l'inclusion $J \subset I$ était stricte, alors la base $\{v_j\}_{j\in J}$ contiendrait moins de n éléments, ce qui serait en contradiction avec le théorème 4.12(3). □

Définition 4.18.
Soit F un sous-espace vectoriel de l'espace vectoriel E. La ***codimension*** *de F dans E est*

$$\operatorname{codim}_E(F) = \dim(E/F). \tag{4.22}$$

¡¡ Avertissement/question à la lectrice !! 4.19
Voir que E/F a une structure vectorielle, expliciter sa dimension en fonction de celles de E et F.

4.1.1 Et en dimension infinie

Dans ZFC, en dimension infinie, il existe aussi une base pour tout espace vectoriel ainsi qu'un théorème de la base incomplète. Nous ne parlerons pas de ce qu'il se passe lorsque nous ne considérons que ZF [7].

Lemme 4.20 ([109])**.**
Soient un $\mathbb{K}$-espace vectoriel E et un sous-espace vectoriel V de E. Soient encore deux sous-espaces vectoriels W_1 et W_2 tels que

(1) $V \cap W_1 = \{0\}$;

(2) $V + W_2 = E$.

Alors il existe un supplémentaire W de V tel que $W_1 \subset W \subset W_2$.

Juste une remarque : dans le Frido le symbole « $\subset$ » ne signifie pas une inclusion stricte.

Démonstration. Nous divisons en petits morceaux.

(i) **Un gros ensemble** Soit $\mathcal{A}$ l'ensemble des sous-espaces vectoriels S de E tels que $W_1 \subset S \subset W_2$ et $S \cap V = \{0\}$. Puisque $W_1 \subset \mathcal{A}$, cet ensemble n'est pas vide. De plus $\mathcal{A}$ est partiellement ordonné pour l'inclusion.

(ii) **$\mathcal{A}$ est inductif** Nous prouvons maintenant que $\mathcal{A}$ est inductif [8]. Pour cela, soit une partie $\mathcal{A}'$ totalement ordonnée et $U = \bigcup_{A\in\mathcal{A}'} A$.

Alors, la partie U est un sous-espace vectoriel de E. En effet si $x, y \in U$, alors il existe $A_1, A_2 \in \mathcal{A}'$ tels que $x \in A_1$ et $y \in A_2$. Comme $\mathcal{A}'$ est totalement ordonné, l'un des ensembles parmi A_1 et A_2 est inclus dans l'autre. Sans perdre de généralité, disons $A_1 \subset A_2$. Alors les opérations s'effectuent dans A_2 : nous avons $x, y \in A_2$, et donc $\lambda x \in A_2 \subset U$ ainsi que $x + y \in A_2 \subset U$.

De plus, U contient W_1, et est contenu dans W_2. Ainsi, $U \in \mathcal{A}$ et majore $\mathcal{A}'$ pour l'inclusion. En bref, $\mathcal{A}$ est bien inductif.

(iii) **Utilisation de Zorn** Le lemme de Zorn 1.22 nous donne alors un élément maximal W de $\mathcal{A}$. Cet élément vérifie

(1) $W \cap V = \{0\}$,

(2) $W_1 \subset W \subset W_2$,

(3) pour tout $W' \in \mathcal{A}$, nous avons $W' \subset W$ par maximalité de W.

(iv) **Supplémentaire** Montrons que ce W est un supplémentaire de V. Soit $x \in E$. Le but est de trouver une décomposition de x en somme d'un élément de W et un de V. Comme $V + W_2 = E$, nous avons $v \in V$ et $w_2 \in W_2$ tels que

$$x = v + w_2. \tag{4.23}$$

7. Si vous ne savez pas ce que signifient les sigles « ZF » et « ZFC » vous ne devriez pas être en train de lire ceci, et encore moins penser à le resservir à un jury d'agrégation.

8. Définition 1.21.

Si $w_2 \in W$ alors c'est fini. Sinon ...

Soit $X = \mathrm{Span}\{W, w_2\}$. Vu que X contient strictement W et que W est maximum dans $\mathcal{A}$, la partie X n'est pas un élément de $\mathcal{A}$. Comme X est un sous-espace vectoriel de E tel que $W_1 \subset X \subset W_2$, la seule possibilité pour que X ne soit pas dans $\mathcal{A}$ est que $X \cap V \neq \{0\}$. Soit donc $y \neq 0$ dans $X \cap V$. Par définition de X,

$$y = w' + \lambda w_2 \tag{4.24}$$

pour $w' \in W$, $w_2 \in W_2$ et $\lambda \in \mathbb{K}$. Nous avons $\lambda \neq 0$, sinon nous aurions $y \in W \cap V$ et donc $y = 0$ puisque W est dans $\mathcal{A}$. La décomposition (4.24) permet alors d'écrire $w_2 = (y - w')/\lambda$ et finalement

$$x = v + \frac{1}{\lambda}(y - w') = \underbrace{v + \frac{1}{\lambda}y}_{\in V} - \underbrace{\frac{1}{\lambda}w'}_{\in W} . \tag{4.25}$$

La somme d'espaces vectoriels $E = V + W$ est donc établie.

□

Corolaire 4.21.
Tout sous-espace vectoriel d'un espace vectoriel possède un supplémentaire.

Démonstration. Soit un espace vectoriel E ainsi qu'un sous-espace vectoriel V. Si $V = E$ nous sommes O.K. Sinon nous considérons $v \in E \backslash V$ et nous posons $W_1 = \mathbb{K}v$ et $W_2 = E$.

Vu que V et W_1 sont des espaces vectoriels, nous avons $V \cap W_1 = \{0\}$, et puisque $W_2 = E$, nous avons $V + W_2 = E$. Le lemme 4.20 nous donne alors un supplémentaire de V. □

Proposition 4.22 (Base incomplète)**.**
Tout espace vectoriel (non réduit à $\{0\}$) possède une base.

Démonstration. Soit $\mathcal{A}$ l'ensemble des familles libres de E. Il n'est pas vide parce que $\{v\}$ en est une dès que v est non nul dans E. Rapidement :

— l'ensemble $\mathcal{A}$ est ordonné pour l'inclusion,

— si $\mathcal{A}'$ est une partie totalement ordonnée, l'union est un majorant,

— donc $\mathcal{A}$ est inductif,

— soit un maximum F de $\mathcal{A}$.

La partie F est libre parce qu'elle est dans $\mathcal{A}$. Elle est génératrice parce que si v n'est pas dans $\mathrm{Span}(F)$ alors la partie $F \cup \{v\}$ est encore libre, et majore strictement F pour l'inclusion, ce qui n'est pas possible.

Donc F est une base de E. □

Théorème 4.23 (Base incomplète, dimension quelconque)**.**
Soit une partie $\{e_i\}_{i \in I}$ génératrice de l'espace vectoriel E (ici, I est un ensemble quelconque[9]). Soit $I_0 \in I$ tel que $\{e_i\}_{i \in I_0}$ soit libre.

Alors il existe I_1 tel que $I_0 \subset I_1 \subset I$ tel que $\{e_i\}_{i \in I_1}$ soit une base de E.

Note : une telle partie I_0 existe en prenant un singleton. Mais l'existence n'est pas le sujet ici.

Démonstration. Soit $\mathcal{A}$ l'ensemble des parties J de I telles que $I_0 \subset J \subset I$ et telles que $\{e_i\}_{i \in J}$ soit libre.

Encore une fois, $\mathcal{A}$ est inductif pour l'ordre partiel donné par l'inclusion. Soit J un élément maximum par le lemme de Zorn 1.22. Puisque $J \in \mathcal{A}$, la partie $\{e_i\}_{i \in J}$ est libre. Mais elle est également génératrice parce que si e_k n'est pas dedans, J ne serait pas maximum, étant majorée par $J \cup \{k\}$.

Donc $\{e_i\}_{i \in J}$ engendre tous les e_i avec $i \in I$ et donc, tous les éléments de E. □

9. Un cas d'utilisation intéressant est de poser $I = E$ et $e_i = i$. Pensez-y.

4.24 ([11]).
Une preuve alternative du théorème de la base incomplète serait de prouver que l'ensemble des parties libres est inductif. De ce fait, la proposition 1.23 permet de dire que toute partie libre peut être complétée en une base.

4.1.2 Espace librement engendré

Définition 4.25 ([110]).
Soient un ensemble S et un corps $\mathbb{K}$. L'espace vectoriel ***librement engendré*** *sur S, noté $F_{\mathbb{K}}(S)$ est l'ensemble des applications $S \to \mathbb{K}$ qui sont non-nulles en un nombre fini de points de S.*

Autrement dit, $\sigma\colon S \to \mathbb{K}$ est dans $F_{\mathbb{K}}(S)$ si $\{x \in S$ tel que $\sigma(x) \neq 0\}$ est fini[10].

Le lemme suivant donne tout son sens à l'expression « librement » engendré. Il dit que $F_{\mathbb{K}}(S)$ possède une base indexée par S lui-même.

Lemme 4.26.
L'ensemble des applications δ_s données par

$$\begin{aligned} \delta_s\colon S &\to \mathbb{K} \\ t &\mapsto \begin{cases} 1 & \text{si } t = s \\ 0 & \text{sinon} \end{cases} \end{aligned} \tag{4.26}$$

avec $s \in S$ forment une base[11] *de $F_{\mathbb{K}}(S)$.*

Démonstration. Pour prouver que les δ_s sont générateurs, nous considérons $g\colon S \to \mathbb{K}$ non nul sur la partie finie $\{s_i\}_{i\in I}$ de S. Alors nous avons

$$g = \sum_{i\in I} g(s_i)\delta_{s_i}. \tag{4.27}$$

Pour prouver que les δ_s forment une partie libre, nous supposons avoir $\lambda_i \in \mathbb{K}$ tels que

$$g = \sum_{i\in I} \lambda_i \delta_{s_i} = 0 \tag{4.28}$$

Soit $j \in I$. Nous avons

$$0 = f(s_j) = \sum_{i\in I} \lambda_i \underbrace{\delta_{s_i}(s_j)}_{=\delta_{i,j}} = \lambda_j. \tag{4.29}$$

Donc les coefficients λ_i sont tous nuls, et nous avons prouvé que la partie est libre. □

Il est parfois pratique d'écrire les éléments de $F_{\mathbb{K}}(S)$ comme sommes « formelles » d'éléments de S. Cela va encore lorsque S est un ensemble n'ayant aucune somme bien définie.

Mais attention : si $S = \mathbb{R}$, l'élément $4+7$ de $F_{\mathbb{R}}(\mathbb{R})$ n'est pas 11. L'élément 11 de $F_{\mathbb{R}}(\mathbb{R})$ est un élément complètement différent. Bref, il n'est pas judicieux d'écrire les éléments de $F_{\mathbb{K}}(S)$ comme des combinaisons linéaires d'éléments de S. Pour $x \in S$ il vaut mieux écrire explicitement δ_x que x. La somme $\delta_x + \delta_y$ est parfaitement bien définie dans l'ensemble des applications de S vers $\mathbb{K}$.

4.2 Applications linéaires

Le lien entre matrice et applications linéaires est la définition 4.65 et toutes les propriétés qui s'en suivent.

10. Parce que nous l'aimons bien, nous ne résistons pas à faire un renvoi vers la définition 1.111.
11. Définition 4.5.

4.2.1 Définition

Définition 4.27.
*Soient des espaces vectoriels E et F sur le corps $\mathbb{K}$. Soit un sous-corps $\mathbb{L}$ de $\mathbb{K}$. Une application $T\colon E \to F$ est dite $\mathbb{L}$-**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.*

Nous noterons $\mathcal{L}_{\mathbb{L}}(E,F)$ l'espace des applications $\mathbb{L}$-linéaires de E vers F.

Si vous avez bien suivi, les égalités dans la définition 4.27 sont des égalités dans F.

4.28.
Le plus souvent, si E est un espace vectoriel sur $\mathbb{K}$, alors nous ne considérerons que les applications $\mathbb{K}$-linéaires. Autrement dit, nous écrirons le plus souvent simplement $\mathcal{L}(E,F)$ sans préciser le corps.

Il pourra pourtant arriver que, pour un espace vectoriel sur $\mathbb{C}$, nous considérions les applications seulement $\mathbb{R}$-linéaires. Ce sera le cas dans le lemme 27.140.

Lemme-Définition 4.29.
L'ensemble de toutes les applications linéaires de E vers F est noté $\mathcal{L}(E,F)$ et devient un espace vectoriel sur $\mathbb{K}$ avec les définitions suivantes :

(1) $(T_1 + T_2)(x) = T_1(x) + T_2(x)$,

(2) $(\lambda T)(x) = \lambda T(x)$.

Exemple 4.30.
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 $T_\lambda(x) = \lambda x$ est une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^m$ pour tout λ dans $\mathbb{R}$ et m dans $\mathbb{N}$. △

Exemple 4.31.
Soit $m \in \mathbb{N}$. 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 lorsque $v \neq 0$, parce que si a est un réel différent de 0 et 1, alors $av \neq v$, d'où

$$U_\lambda(ax) = \lambda(ax) + v \neq a(\lambda x + v) = aU_\lambda(x).$$

△

Exemple 4.32.
Soit A une matrice fixée de $\mathbb{M}(n \times n, \mathbb{R})$. La fonction $T_A\colon \mathbb{R}^m \to \mathbb{R}^n$ définie par $T_A(x) = Ax$ est une application linéaire. En effet,

— $T_A(x+y) = A(x+y) = Ax + Ay = T_A(x) + T_A(y)$,

— $T_A(ax) = A(ax) = a(Ax) = aT_A(x)$.

△

Lemme 4.33.
Si une application linéaire est inversible, alors son inverse est linéaire.

Démonstration. Soient une application linéaire inversible $f\colon E \to F$, ainsi que $x, y \in E$. Nous avons

$$f\big(f^{-1}(x) + f^{-1}(y)\big) = f\big(f^{-1}(x)\big) + f\big(f^{-1}(y)\big) = x + y = f\big(f^{-1}(x+y)\big). \tag{4.30}$$

En prenant f^{-1} des deux côtés, nous trouvons

$$f^{-1}(x) + f^{-1}(y) = f^{-1}(x+y). \tag{4.31}$$

De même :

$$f\big(\lambda f^{-1}(x)\big) = \lambda f\big(f^{-1}(x)\big) = \lambda x = f\big(f^{-1}(\lambda x)\big), \tag{4.32}$$

ce qui nous donne $\lambda f^{-1}(x) = f^{-1}(\lambda x)$. □

Définition 4.34 (Quelques ensembles d'applications linéaires)**.**
Soient E et F des espaces vectoriels.

- *L'ensemble des applications linéaires de E vers F est noté $\mathcal{L}(E, F)$, comme déjà dit en 4.29.*
- *Une application linéaire $E \to E$ est un **endomorphisme** de E. L'ensemble des endomorphismes de E est noté* $\mathrm{End}(E)$.
- *Un endomorphisme bijectif est un **automorphisme**. L'ensemble des automorphismes de E est noté* $\mathrm{Aut}(E)$.
- *Une application linéaire bijective $E \to F$ est un **isomorphisme** d'espace vectoriel. L'ensemble des isomorphismes de $E \to F$ est noté*[12] $\mathrm{GL}(E, F)$. *C'est un groupe par le lemme 4.33.*

Remarque 4.35.
Les ensembles définis en 4.34 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 11.179.

Définition 4.36.
*Si E est un espace vectoriel, si X est un espace vectoriel, et si $f\colon X \to E$ est une application, le **noyau** de f est le noyau de f lorsque E est vu comme un groupe pour l'addition*[13]*, c'est-à-dire la partie*

$$\ker(f) = \{x \in X \text{ tel que } f(x) = 0\}. \tag{4.33}$$

Proposition 4.37.
Le noyau d'une application linéaire est un sous-espace vectoriel.

Démonstration. Soit une application linéaire $f\colon E \to F$. Si $x, y \in \ker(f)$ et si $\lambda \in \mathbb{K}$ alors

$$f(x + y) = f(x) + f(y) = 0 + 0 = 0, \tag{4.34}$$

donc $x + y \in \ker(f)$, et

$$f(\lambda x) = \lambda f(x) = 0, \tag{4.35}$$

donc $\lambda x \in \ker(f)$. □

Proposition 4.38.
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 E et F, alors il existe une unique application linéaire $T\colon E \to F$ telle que $T(e_i) = f_i$ pour tout i.

Démonstration. En deux parties.

(i) **Existence** Soit $v \in E$. Vu que $\{e_i\}$ est une base, v se décompose de façon unique en $v = \sum_i v_i e_i$. Alors la définition

$$T(v) = \sum_i v_i f_i \tag{4.36}$$

est une bonne définition et satisfait aux exigences.

(ii) **Unicité** Soient T et U satisfaisant aux exigences. Alors pour tout i nous avons $T(e_i) = U(e_i)$. Si $v \in E$ s'écrit de la forme $v = \sum_i v_i e_i$ alors la linéarité impose $T(v) = \sum_i v_i T(e_i) = \sum_i v_i U(e_i) = U(v)$. Donc $T = U$.

12. Le fait d'utiliser une notation similaire à celle des matrices inversibles n'est pas anodine : le lecteur en est sans doute conscient.
13. Définition 2.5.

□

Lemme 4.39 ([1]).
Soient des espaces vectoriels V et W de dimension finie. Soient des bases $\{e_i\}$ de V et $\{f_\alpha\}$ de W. Nous posons

$$\begin{aligned}\varphi_{i\alpha}\colon V &\to W\\ v &\mapsto v_i f_\alpha\end{aligned} \tag{4.37}$$

où v_i est défini par la décomposition (unique) $v = \sum_i v_i e_i$.

Alors :

(1) La partie $\{\varphi_{i\alpha}\}$ est une base de $\mathcal{L}(V, W)$.

(2) Au niveau des dimensions : $\dim\big(\mathcal{L}(V,W)\big) = \dim(V)\dim(W)$.

Démonstration. Il faut prouver que $\{\varphi_{i\alpha}\}$ est libre et générateur.

(i) **Générateur** Soit une application linéaire $b\colon V \to W$. En décomposant $b(v)$ dans la base $\{f_\alpha\}$, nous définissons $b_\alpha\colon V \to \mathbb{K}$ par

$$b(v) = \sum_\alpha b_\alpha(v) f_\alpha. \tag{4.38}$$

Nous posons $b_{\alpha i} = b_\alpha(e_i)$. Ainsi,

$$b(v) = \sum_\alpha v_i b_{\alpha i} f_\alpha = \sum_{\alpha i} b_{\alpha i}\varphi_{i\alpha}(v). \tag{4.39}$$

Donc b peut être écrit comme combinaison linéaire des $\varphi_{i\alpha}$.

(ii) **Libre** Supposons que $\sum_{i\alpha} a_{i\alpha}\varphi_{i\alpha} = 0$ pour certains coefficients $a_{i\alpha} \in \mathbb{K}$. Nous avons, pour tout $v \in V$:

$$0 = \sum_{i\alpha} a_{i\alpha}\varphi_{i\alpha}(v) = \sum_{i\alpha} a_{i\alpha} v_i f_\alpha, \tag{4.40}$$

mais comme les f_α forment une base, chaque terme de la somme sur α est nul :

$$\sum_i a_{i\alpha} v_i = 0. \tag{4.41}$$

Et comme cela est valable pour tout v et donc, pour tout choix de v_i, nous avons $a_{i\alpha} = 0$ pour tout i et pour tout α.

La formule de dimension est simplement la cardinalité de la base trouvée ; c'est la définition 4.14.

□

4.2.2 Linéarité et bases

Proposition 4.40 ([111]).
Soient deux espaces vectoriels E et F. Une application linéaire[14] *$f\colon E \to F$ est injective si et seulement si* $\ker(f) = \{0\}$.

Démonstration. Nous supposons que f est injective. Si $x \in \ker(f)$, alors $f(x) = 0$. Or f est linéaire, donc $f(0) = 0$. Nous avons donc $f(x) = f(0)$ et donc $x = 0$ parce que f est injective.

Dans l'autre sens, soient x, y tels que $f(x) = f(y)$. Par linéarité de f nous avons $f(x - y) = 0$, et donc $x - y = 0$ parce que $\ker(f) = \{0\}$. Donc $x = y$ et f est injective. □

Proposition 4.41 ([111]).
Soit $f \in \mathcal{L}(E, F)$ où E et F sont deux espaces vectoriels.

(1) Si f est injective et si $\{v_i\}_{i\in I}$ est libre, alors $\{f(v_i)\}_{i\in I}$ est libre.

(2) Si f est surjective et si $\{v_i\}_{i\in I}$ est génératrice, alors $\{f(v_i)\}_{i\in I}$ est génératrice.

14. Définition 4.27.

(3) Si f est une bijection, alors l'image d'une base par f est une base.

Démonstration. En trois parties.

(i) **(1)** Nous devons montrer que $\{f(v_j)\}_{j\in J}$ est libre pour tout J fini dans I. Soit donc une partie finie $J \subset I$ et des scalaires[15] tels que $\sum_{j\in J} \lambda_j f(v_j) = 0$. La linéarité de f donne[16]

$$f\big(\sum_{j\in J} \lambda_j v_j\big) = 0. \tag{4.42}$$

Par injectivité de f nous avons alors $\sum_j \lambda_j v_j = 0$. Comme les v_j eux-même forment une partie libre, nous avons $\lambda_j = 0$ pour tout $j \in J$.

(ii) **(2)** Soit $y \in F$. Puisque f est surjective, il existe $x \in E$ tel que $f(x) = y$. Étant donné que $\{v_i\}_{i\in I}$ est générateur, il existe une partie finie $J \subset I$ et des scalaires $\lambda_j \in \mathbb{K}$ tels que

$$x = \sum_{j\in J} \lambda_j v_j. \tag{4.43}$$

En appliquant f aux deux côtés, et en tenant compte de la linéarité de f,

$$y = f(x) = \sum_{j\in J} \lambda_j f(v_j), \tag{4.44}$$

ce qui prouve que y est une combinaison linéaire des $f(v_j)$.

(iii) **(3)** Une base est à la fois libre et génératrice et une bijection est à la fois injective et surjective. Les deux premiers points permettent de conclure.

□

Corolaire 4.42 ([1])**.**
Si E et F sont des espaces vectoriels isomorphes de dimensions finies. Alors leurs dimensions sont égales.

Démonstration. Puisque E et F sont isomorphes, il existe une bijection $f\colon E \to F$. Par la proposition 4.41(3), l'image d'une base de E est une base de F. Donc les espaces E et F ont des bases contenant le même nombre d'éléments. □

4.2.3 Rang

La proposition 4.43 et le théorème 4.44 sont valables également en dimension infinie ; ce sera une des rares incursions en dimension infinie de ce chapitre.

Proposition-Définition 4.43.
L'image d'une application linéaire est un espace vectoriel. La dimension de cet espace est le ***rang*** *de ladite application linéaire.*

Démonstration. Soit une application linéaire $f\colon E \to F$. Nous considérons v, w dans l'image de f ainsi que λ dans le corps de base commun à E et F.

Soient $v_0 \in E$ et $w_0 \in E$ tels que $v = f(v_0)$ et $w = f(w_0)$. Alors $v + w = f(v_0 + w_0)$ et $\lambda v = f(\lambda v_0)$. Donc l'image est bien un espace vectoriel. □

Théorème 4.44 (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.

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* $\mathrm{Image}(f)$ *alors*

$$(u_s)_{s\in s} \cup (v_t)_{t\in T} \tag{4.45}$$

15. Des éléments du corps de base $\mathbb{K}$.
16. Voir les propriétés de la définition 4.27.

est une base de E.

En dimension finie, nous avons en plus la formule suivante :

$$\mathrm{rk}(f) + \dim(\ker f) = \dim E, \tag{4.46}$$

c'est-à-dire que le rang[17] *de f est égal à la codimension*[18] *du noyau.*

Démonstration. Nous devons montrer que

$$(u_s)_{s\in S} \cup (v_t)_{t\in T} \tag{4.47}$$

est libre et générateur.

Soit $x \in E$. Nous définissons les nombres x_t par la décomposition de $f(x)$ dans la base $\big(f(v_t)\big)$:

$$f(x) = \sum_{t\in T} x_t f(v_t). \tag{4.48}$$

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) :

$$x - \sum_t x_t v_t = \sum_{s\in S} x_s u_s. \tag{4.49}$$

Par conséquent

$$x = \sum_s x_s u_s + \sum_t x_t v_t. \tag{4.50}$$

En ce qui concerne la liberté nous écrivons

$$\sum_t x_t v_t + \sum_s x_s u_s = 0. \tag{4.51}$$

En appliquant f nous trouvons que

$$\sum_t x_t f(v_t) = 0 \tag{4.52}$$

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 25.18.

Proposition 4.45 ([112])**.**
Soit E, un espace vectoriel de dimension finie sur le corps $\mathbb{K}$. Soient V et W des sous-espaces vectoriels de E. Alors

$$\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W). \tag{4.53}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} \varphi\colon V \times W &\to E \\ (x, y) &\mapsto x + y. \end{aligned} \tag{4.54}$$

C'est une application linéaire dont l'image est $V + W$. Nous avons donc, pour commencer

$$\dim(V + W) = \dim\big(\mathrm{Image}(\varphi)\big). \tag{4.55}$$

17. Définition 4.43.
18. Définition 4.18.

Nous appliquons à présent le théorème du rang 4.44 à l'application φ :

$$\begin{aligned}\dim(V+W) &= \dim\big(\operatorname{Image}(\varphi)\big) && (4.56a)\\ &= \dim(V\times W) - \dim\big(\ker(\varphi)\big) && (4.56b)\\ &= \dim(V)+\dim(W) - \dim\big(\ker(\varphi)\big). && (4.56c)\end{aligned}$$

Nous devons maintenant étudier $\ker(\varphi)$. D'abord, $(v,w)\in V\times W$ appartient à $\ker(\varphi)$ si et seulement si $v+w=0$. Nous avons donc

$$\ker(\varphi) = \{(x,-x) \text{ tel que } x\in V\cap W\}. \tag{4.57}$$

Nous montrons à partir de cela que $\dim\big(\ker(\varphi)\big) = \dim(V\cap W)$ en montrant que l'application

$$\begin{aligned}\psi\colon V\cap W &\to \ker(\varphi)\\ x &\mapsto (x,-x)\end{aligned} \tag{4.58}$$

est un isomorphisme d'espaces vectoriels. D'abord ψ est injective parce que si $\psi(x)=\psi(y)$, alors $(x,-x)=(y,-y)$ et donc $x=y$. Ensuite, ψ est surjective parce qu'un élément générique de $\ker(\varphi)$ est $(x,-x)=\psi(x)$ avec $x\in V\cap W$. L'application ψ étant un isomorphisme d'espaces vectoriels, nous avons bien $\dim\big(\ker(\varphi)\big)=\dim(V\cap W)$. □

Corolaire 4.46.
Soient deux espaces vectoriels E et F de même dimension finie[19]*. Pour une application linéaire $f\colon E\to F$, les trois conditions suivantes sont équivalentes :*

(1) f est injective ;
(2) f est surjective ;
(3) f est bijective.

Démonstration. Si $f\colon E\to E$ est surjective, alors $\operatorname{rk}(f)=\dim(E)$, ce qui donne, par le théorème du rang 4.44, $\dim\big(\ker(f)\big)=0$, 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 $\operatorname{rk}(f)=\dim(E)$ ou encore que f est surjective. □

Exemple 4.47.
Le corolaire 4.46 n'est pas correct en dimension infinie. Par exemple en prenant $f(e_1)=f(e_2)=e_1$ et ensuite $f(e_k)=e_{k-1}$ pour tout $k\geqslant 2$. Cette application est surjective mais pas injective. △

Une conséquence du théorème du rang est que les endomorphismes ont un inverse à gauche et à droite égaux (lorsqu'ils existent). En résumé, ce que le corolaire 4.48 dit est que si $AB=\mathbb{1}$, alors $BA=\mathbb{1}$.

Corolaire 4.48.
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) les inverses à gauche et à droite sont égaux.

Tout cela tient également en remplaçant « gauche » par « droite ».

Démonstration. Soit g, un inverse à gauche de f : $gf=\mathrm{Id}$. Cela implique que f est injective et que g est surjective, et donc qu'elles sont toutes deux bijectives par le corolaire 4.46. Puisque f est bijective, elle admet également un inverse à droite, soit h. Nous avons : $gf=\mathrm{Id}$ et $fh=\mathrm{Id}$.

Alors $gfh=h$ parce que $gf=\mathrm{Id}$, mais également $gfh=g$ parce que $fh=\mathrm{Id}$. Donc $g=h$.[20] □

19. Les deux mots sont importants : les dimensions doivent être égales et finies.
20. C'est le même argument que celui employé pour la preuve du lemme 1.157 (2), à ceci près que nous devions montrer l'existence de l'inverse à droite.

C'est ce corolaire qui nous permet d'écrire f^{-1} sans plus de précisions dès que f est une bijection.

Exemple 4.49 (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

$$f(e_k) = e_{k+1}. \tag{4.59}$$

Il est injectif, mais pas surjectif. Si on pose

$$g(e_k) = \begin{cases} e_{k-1} & \text{si } k \geqslant 1 \\ 0 & \text{si } k = 0 \end{cases} \tag{4.60}$$

alors nous avons $gf = \mathrm{Id}$, mais pas $fg = \mathrm{Id}$ parce que ce $(fg)(e_0) = 0$. △

Lemme 4.50.
Si E et F sont des espaces vectoriels et si $f\colon E \to F$ est une application linéaire inversible, alors son inverse est également linéaire.

Démonstration. Nous avons $f^{-1}(x+y) = f^{-1}(x) + f^{-1}(y)$. En effet,

$$f\big(f^{-1}(x) + f^{-1}(y)\big) = f\big(f^{-1}(x)\big) + f\big(f^{-1}(y)\big) = x + y. \tag{4.61}$$

De la même façon,

$$f\big(\lambda f^{-1}(x)\big) = \lambda x, \tag{4.62}$$

donc $f^{-1}(\lambda x) = \lambda f^{-1}(x)$. □

Proposition 4.51.
Soient un espace vectoriel E de dimension finie, un endomorphisme $f\colon E \to E$ et une partie $\{v_i\}_{i\in I}$ tel que $\{f(v_i)\}_{i\in I}$ soit une base.

Alors $\{v_i\}_{i\in I}$ est une base.

Démonstration. Soit $x \in E$. Il existe une partie finie $J \subset I$ et des scalaires λ_j tels que

$$x = \sum_j \lambda_j f(v_j) = f\big(\sum_j \lambda_j v_j\big), \tag{4.63}$$

ce qui prouve que f est surjective. Le corolaire 4.46 nous dit alors que f est une bijection. L'application inverse est également linéaire par le lemme 4.50.

Une application linéaire bijective (comme f^{-1}) transforme une base en une base par la proposition 4.41. Donc

$$f^{-1}\big(\{f(v_i)\}\big) \tag{4.64}$$

est une base. □

Proposition 4.52.
Soit un espace vectoriel E de dimension finie et deux applications linéaires $f, g\colon E \to E$ telles que $g \circ f = \mathrm{Id}$. Alors f et g sont bijectives.

Démonstration. En plusieurs étapes

(i) **f est injective** Si $f(x) = f(y)$, alors en appliquant g nous avons

$$g\big(f(x)\big) = g\big(f(y)\big), \tag{4.65}$$

ce qui donne $x = y$.

(ii) **f est surjective** C'est maintenant le corolaire 4.46.

(iii) **g est surjective** Pour tout $x \in E$ nous avons $g\big(f(x)\big) = x$. Donc l'image de $f(E)$ par g est E.

(iv) g **est injective** C'est maintenant le corolaire 4.46.

□

Lemme 4.53 ([113]).
Soit une application linéaire $f\colon E \to F$.

(1) L'application f est injective si et seulement si il existe $g\colon F \to E$ telle que $g \circ f = \mathrm{Id}\,|_E$.

(2) L'application f est surjective si et seulement si 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 $g \circ f = \mathrm{Id}$.

Réciproquement, si 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

$$\begin{aligned} g\colon F &\to E \\ \sum_k a_k f(x_k) &\mapsto \sum_k a_k x_k. \end{aligned} \tag{4.66}$$

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

$$(f \circ g)(v) = \sum_k v_k (f \circ g)\big(f(x_k)\big) = f\left(\sum_k v_k x_k\right) = \sum_k v_k f(x_k) = v. \tag{4.67}$$

Réciproquement, si il existe $g\colon F \to E$ tel que $f \circ g = \mathrm{Id}$ alors f doit être surjective, parce que

$$F = \mathrm{Image}(f \circ g) = f\big(\mathrm{Image}(g)\big) \subset \mathrm{Image}(f). \tag{4.68}$$

□

4.3 Matrices

Les matrices et les applications linéaires sont deux choses différentes. Une application linéaire [21] est une application d'un espace vectoriel vers un autre, et une matrice est un simple tableau de nombres sur lesquels nous définissons des opérations, de telle sorte à fournir une structure d'espace vectoriel. Le lien entre ces opérations et les opérations correspondantes sur les applications linéaires sera fait plus tard. Voir la définition 4.65 et ce qui s'en suit.

4.3.1 Définitions

Les notions topologiques sur les espaces de matrices sont pour plus tard, à commencer par la définition 7.188.

Définition 4.54.
Soit un anneau $\mathbb{A}$ ainsi que des entiers m, n strictement positifs. L'ensemble $\mathbb{M}(n \times m, \mathbb{A})$ est l'ensemble des applications

$$\{1, \dots, n\} \times \{1, \dots, m\} \to \mathbb{A}, \tag{4.69}$$

*et est appelé ensemble des **matrices** $n \times m$ sur $\mathbb{A}$.*

Si A est une matrice, nous notons $A_{i,j}$ au lieu de $A(i,j)$ l'image de (i,j) par l'application A.

21. Définition 4.27.

Définition 4.55.
Quelques ensembles de matrices particuliers.

(1) Si $n = m$, alors :

— *nous disons que la matrice est* ***carrée****,*

— *nous notons $\mathbb{M}(n, \mathbb{A})$ pour $\mathbb{M}(n \times n, \mathbb{A})$,*

— *n est appelée* ***ordre*** *de la matrice.*

(2) Si $n = 1$, alors la matrice est appelée ***matrice-ligne****.*

(3) Si $m = 1$, alors la matrice est appelée ***matrice-colonne****.*

4.56.
On note les isomorphismes naturels $\mathbb{M}(1 \times m, \mathbb{A}) \simeq \mathbb{A}^m$ et $\mathbb{M}(n \times 1, \mathbb{A}) \simeq \mathbb{A}^n$.

Lemme-Définition 4.57.
Nous considérons les opérations suivantes sur $\mathbb{M}(n \times m, \mathbb{A})$:

Somme $(A + B)_{i,j} = A_{i,j} + B_{i,j}$,

Produit par un scalaire $(\lambda \cdot A)_{i,j} = \lambda A_{i,j}$ *pour tout $A, B \in \mathbb{M}(n \times m, \mathbb{A})$ et $\lambda \in \mathbb{A}$.*

Alors $\big(\mathbb{M}(n \times m, \mathbb{A}), +, \cdot\big)$ est un $\mathbb{A}$-module[22].

Lemme-Définition 4.58.
Avec la multiplication

$$\begin{aligned} \mathbb{M}(n \times p, \mathbb{A}) \times \mathbb{M}(p \times m, \mathbb{A}) &\to \mathbb{M}(n \times m, \mathbb{A}) \\ (A, B) &\mapsto (AB)_{i,j} = \sum_{k=1}^{p} A_{i,k} B_{k,j}, \end{aligned} \tag{4.70}$$

l'espace $\mathbb{M}(n, \mathbb{K})$ est une $\mathbb{K}$-algèbre[23].

Définition 4.59.
Pour un élément $A \in \mathbb{M}(n \times m, \mathbb{A})$ nous définissons encore

La transposée $A^t_{i,j} = A_{j,i}$,

La trace $\mathrm{Tr}(A) = \sum_i A_{i,i}$.

Remarque 4.60.
Quelques remarques directes sur les définitions.

(1) La motivation de cette définition pour le produit apparaîtra plus loin, mais le Frido n'étant pas un livre d'introduction, j'imagine que le lecteur a déjà une idée.

(2) Nous verrons plus loin en 9.10.3 que la définition de transposée d'une application linéaire n'est pas tout à fait évidente ; elle sera la définition 9.176.

Ici nous avons bien défini la transposée d'une matrice, pas d'une application linéaire.

Remarque 4.61.
Quelques remarques à propos de structures supplémentaires.

(1) Nous utiliserons (presque) tout le temps des matrices à coefficients dans un corps. Il est clair que, si $\mathbb{K}$ est un corps (commutatif), alors $\mathbb{M}(n \times m, \mathbb{K})$ a une structure d'espace vectoriel sur $\mathbb{K}$.

(2) Par ailleurs, sur les matrices carrées d'ordre n fixé, le produit de deux matrices est bien défini. Ainsi, $\mathbb{M}(n, \mathbb{A})$ se voit conférer une structure d'anneau, dont le neutre pour la multiplication

22. Définition 1.274
23. Définition 1.291.

est la matrice carrée $\mathbb{1}_n$ (notée aussi $\mathbb{1}$ lorsqu'il n'y a pas d'ambiguïté sur la taille), donnée par

$$\mathbb{1}_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases} \tag{4.71}$$

Il est vite vu que si A est une matrice carrée d'ordre n, alors $A\mathbb{1} = \mathbb{1}A = A$.

Lemme 4.62 ([1])**.**
Si A, B et C sont des matrices, nous avons

(1) $(AB)^t = B^t A^t$,

(2) $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB)$.

Démonstration. La première est un simple calcul :

$$(AB)^t_{i,j} = (AB)_{j,i} = \sum_k A_{j,k} B_{k,i} = \sum_k A^t_{k,j} B^t_{i,k} = (B^t A^t)_{i,j}. \tag{4.72}$$

Pour la seconde :

$$\mathrm{Tr}(ABC) = \sum_{ikl} A_{i,k} B_{k,l} C_{l,i} = \sum_{ikl} C_{l,i} A_{i,k} B_{k,l} = \sum_l (CAB)_{l,l} = \mathrm{Tr}(CAB). \tag{4.73}$$

□

4.63.
La seconde égalité est importante et est nommée **invariance cyclique** de la trace. Elle sert, entre autres nombreuses choses, à prouver que la trace d'une matrice d'une application linéaire ne dépend pas de la base choisie. Ce sera la proposition 9.192.

Lemme 4.64.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$. Si pour tout $x, y \in \mathbb{K}^n$ nous avons

$$\sum_{ij} A_{i,j} x_i y_j = \sum_{ij} B_{i,j} x_i y_j \tag{4.74}$$

alors $A = B$.

Démonstration. Il suffit de choisir $x_i = \delta_{i,k}$ et $y_j = \delta_{j,l}$ et d'effectuer les sommes ; par exemple

$$\sum_{ij} A_{i,j} \delta_{i,k} y_j = \sum_j A_{k,j} y_j. \tag{4.75}$$

Après avoir effectué toutes les sommes, nous nous retrouvons avec $A_{k,l} = B_{k,l}$, ce qui signifie $A = B$. □

4.3.2 Identifier matrices et applications linéaires

Voir dans l'index thématique, -2.1.1.

Soient deux espaces vectoriels de dimension finie E, F sur le corps $\mathbb{K}$. Nous considérons les bases[24] $\{e_i\}$ pour E et $\{f_\alpha\}$ pour F.

Définition 4.65.
Nous considérons l'application

$$\begin{aligned} \psi \colon \mathbb{M}(n \times m, \mathbb{K}) &\to \mathcal{L}(E, F) \\ A &\mapsto f_A \end{aligned} \tag{4.76}$$

24. C'est le théorème 4.12 qui nous permet de considérer des bases. Et ce théorème ne fonctionne que parce que nous avons supposé une dimension finie.

où f_A est définie par

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i} x_i f_\alpha \tag{4.77}$$

si x_i sont les coordonnées de $x \in E$ dans la base $\{e_i\}$.

4.66.
Nous allons prouver un certain nombre de résultats montrant que cette application a toutes les propriétés imaginables permettant d'identifier les matrices aux applications linéaires : elle est un isomorphisme pour toutes les structure que vous pouvez raisonnablement imaginer.

À cette application ψ il manque cependant une propriété importante : elle n'est pas canonique. Elle dépend des bases choisies. Autrement dit : nous avons à priori autant d'applications ψ différentes qu'il y a de choix de bases sur E et F [25].

Nous allons prouver maintenant quelques résultats montrant que les matrices et les applications linéaires, dans le cas des espaces vectoriels $\mathbb{K}^n$ sont deux présentations de la même chose.

Le fait que ψ est continue sera la proposition 12.113.

4.67.
Lorsque $A \in \mathbb{M}(n, \mathbb{K})$ est une matrice et $x \in \mathbb{K}^n$ un vecteur, nous notons Ax l'élément de $\mathbb{K}^n$ donné par

$$(Ax)_i = \sum_j A_{i,j} x_j. \tag{4.78}$$

Autrement dit, $Ax = f_A(x)$.

Cette convention et de nombreuses autres à propos de matrice sera rappelée dans -2.1.

Proposition-Définition 4.68.
Soient deux espaces vectoriels de dimension finie E, F sur le corps $\mathbb{K}$. Nous considérons les bases $\{e_i\}$ pour E et $\{f_\alpha\}$ pour F.

Nous considérons l'application

$$\begin{aligned} \psi \colon \mathbb{M}(n \times m, \mathbb{K}) &\to \mathcal{L}(E, F) \\ A &\mapsto f_A \end{aligned} \tag{4.79}$$

où f_A est définie par

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i} x_i f_\alpha \tag{4.80}$$

si x_i sont les coordonnées de $x \in E$ dans la base $\{e_i\}$.

Alors

(1) Nous avons

$$f_A(e_i)_\alpha = A_{\alpha,i}. \tag{4.81}$$

(2) Nous avons

$$f_A(e_i) = \sum_\alpha A_{\alpha,i} f_\alpha. \tag{4.82}$$

(3) Nous avons

$$\big(f_A(x)\big)_\alpha = \sum_i A_{\alpha,i} x_i. \tag{4.83}$$

(4) L'application ψ est une bijection.

Si f est une application linéaire, alors la matrice $\psi^{-1}(f)$ est la ***matrice associée*** *à f dans les bases choisies.*

Remarque : les bases ne sont supposées être canoniques en aucun sens du terme. Les dimensions de E et F ne sont pas non plus supposées identiques.

25. Bonne question. Est-ce qu'il y a moyen de construire deux choix de bases donnant la même application ψ ? Écrivez-moi si vous savez la réponse.

Démonstration. En nous rappelant que $(e_j)_i = \delta_{i,j}$ nous avons

$$f_A(e_j) = \sum_{i\alpha} A_{\alpha,i}(e_j)_i f_\alpha = \sum_\alpha A_{\alpha,j} f_\alpha, \tag{4.84}$$

donc $f_A(e_i)_\alpha = A_{\alpha,i}$. Cela prouve la formule du point (1).

Le point (2) est une simple somme sur α de (1).

La formule du point (3) est simplement la composante f_α de la définition 4.80.

Prouvons que ψ est injective. Si $f_A = f_B$, nous avons en particulier $f_A(e_i)_\alpha = f_B(e_i)_\alpha$ et donc $A_{\alpha,i} = B_{\alpha,i}$.

Prouvons que ψ est surjective. Pour cela nous considérons $f \in \mathcal{L}(E,F)$ et nous posons $A_{\alpha,i} = f(e_i)_\alpha$. Nous avons alors $f = f_A$ parce que

$$f_A(x) = \sum_{i\alpha} A_{\alpha,i} x_i f_\alpha = \sum_{i\alpha} f(e_i)_\alpha x_i f_\alpha = \sum_\alpha f(\sum_i x_i e_i)_\alpha f_\alpha = \sum_\alpha f(x)_\alpha f_\alpha = f(x). \tag{4.85}$$

□

La proposition suivante montre que le produit matriciel correspond à la composition d'applications linéaires, pourvu que l'on travaille avec les bases canoniques sur $\mathbb{K}^n$.

Proposition 4.69 ([1])**.**
Soit un corps commutatif $\mathbb{K}$. Nous considérons des espaces vectoriels E et F munis de bases $\{e_i\}_{i=1,\dots,n}$ et $\{f_\alpha\}_{\alpha=1,\dots,m}$.

L'application déjà définie[26]

$$\psi\colon \mathbb{M}(m \times n, \mathbb{K}) \to \mathcal{L}(E,F) \tag{4.86}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. Le fait que ψ soit une bijection est la proposition 4.68. Nous devons montrer que c'est une application linéaire.

Pour $\lambda \in \mathbb{K}$ nous avons le calcul

$$\psi(\lambda A)(e_k) = f_{\lambda A}(e_k) = \sum_{\alpha i} (\lambda A)_{\alpha,i} \underbrace{(e_k)_i}_{=\delta_{k,i}} f_\alpha = \lambda \sum_\alpha A_{\alpha,k} f_\alpha = \lambda f_A(e_j). \tag{4.87}$$

Donc $\psi(\lambda A) = \lambda\psi(A)$.

Si $A, B \in \mathbb{M}(n,\mathbb{K})$ nous avons de la même façon, $f_{A+B} = f_A + f_B$. □

Proposition 4.70.
Soient des espaces vectoriels E, F et G de dimensions n, m et p munis de bases[27] *$\{e_i\}$, $\{f_i\}$ et $\{g_i\}$. Nous considérons les applications*

$$\psi\colon \mathbb{M}(m \times n, \mathbb{K}) \to \mathcal{L}(E,F) \tag{4.88a}$$
$$\psi\colon \mathbb{M}(p \times m, \mathbb{K}) \to \mathcal{L}(F,G) \tag{4.88b}$$
$$\psi\colon \mathbb{M}(p \times n, \mathbb{K}) \to \mathcal{L}(E,G). \tag{4.88c}$$

Nous avons

$$\psi(A) \circ \psi(B) = \psi(AB) \tag{4.89}$$

pour toutes matrices $A \in \mathbb{M}(p \times m, \mathbb{K})$ et $B \in \mathbb{M}(m \times n, \mathbb{K})$.

26. Notez la position du n et du m. Sachez noter les bornes des sommes écrites dans la démonstration.
27. Avec trois ensembles, nous renonçons à utiliser des alphabets différents pour numéroter les éléments des bases.

Démonstration. Nous considérons les applications linéaires associées à A et B : $f_A\colon F \to G$ et $f_B\colon E \to F$ et la composée $f_A \circ f_B\colon E \to G$. Et puis c'est le calcul :

$$(f_A \circ f_B)(e_k) = f_A\big(\sum_{ij} B_{i,j}(e_k)_j f_i\big) \tag{4.90a}$$
$$= \sum_i B_{i,k} f_A(f_i) \tag{4.90b}$$
$$= \sum_i B_{i,k} \sum_{rs} A_{r,s}(f_i)_s g_r \tag{4.90c}$$
$$= \sum_{ir} B_{i,k} A_{r,i} g_r \tag{4.90d}$$
$$= \sum_r (AB)_{r,k} g_r \tag{4.90e}$$
$$= f_{AB}(e_k). \tag{4.90f}$$

Donc $f_A \circ f_B = f_{AB}$ comme il se doit. □

Nous pouvons particulariser au cas où $E = F = G$.

Proposition 4.71.
Si E est un espace vectoriel muni d'une base $\{e_i\}$, alors l'application

$$\psi\colon \mathbb{M}(n, \mathbb{K}) \to \mathrm{End}(E) \tag{4.91}$$

est un isomorphisme d'algèbre[28] *et d'anneaux*[29].

Démonstration. Le fait que ψ soit un isomorphisme d'algèbre est juste la combinaison entre les propositions 4.69 et 4.70.

En ce qui concerne l'isomorphisme d'anneaux, il faut en plus identifier les neutres. Le neutre pour la composition d'applications linéaires est l'application identité et le neutre pour la multiplication de matrices est la matrice identité. Nous devons donc montrer que $\psi(\delta) = f_\delta = \mathrm{Id}$. Juste un calcul :

$$f_\delta(x) = \sum_{ij} \delta_{i,j} x_j e_i = \sum_i x_i e_i = x. \tag{4.92}$$

Donc oui, f_δ est l'identité. □

Le fait que ψ est continue sera la proposition 12.113.

Voilà. Soyez bien conscient que l'application ψ dont nous avons beaucoup parlé est surtout intéressante dans le cas des espaces de la forme $\mathbb{K}^n$. Dans ce cas, nous avons une identification canonique entre $\mathbb{M}(n, \mathbb{K})$ et $\mathrm{End}(\mathbb{K}^n)$ qui est un isomorphisme d'anneaux et d'algèbres.

Nous verrons que ce ψ respecte encore les inverses[30] et les déterminants[31].

4.72.
Il convient de ne pas confondre matrice et application linéaire (bien que nous le ferons sans vergogne). Une matrice est un bête tableau de nombres, tandis qu'une application linéaire est une application entre deux espaces vectoriels vérifiant certaines propriétés.

Cependant si les espaces vectoriels E et F sont munis de bases, alors il y a une application

$$\psi\colon \mathbb{M}(m \times n, \mathbb{K}) \to \mathcal{L}(E, F) \tag{4.93}$$

qui a toutes les propriétés imaginables[32].

28. Définition 1.291.
29. Définition 1.39
30. Proposition 4.89.
31. Proposition 9.10.
32. Et elle en aura encore plus lorsque nous aurons vu les déterminants.

Cette application dépend des bases choisies. Il n'y a donc pas de trucs comme « la matrice de telle application linéaire » ou comme « voici une matrice, nous considérons l'application linéaire associée ».

Cependant, sur des espaces comme $\mathbb{R}^n$ ou plus généralement sur $\mathbb{K}^n$, nous avons une base canonique et toute personne raisonnable utilise toujours la base canonique (sauf mention du contraire). Dans ces cas il est sans danger de dire « la matrice associée à telle application linéaire » sans préciser les bases.

Mais si un jour vous utilisez une base autre que la base canonique sur $\mathbb{R}^n$, précisez-le et plutôt deux fois qu'une[33].

4.73.
Tant que nous sommes à parler de matrice et d'applications linéaires, les plus acharnés anti-abus de language[34] remarqueront qu'il n'est pas vrai que « étant donné une base, une application linéaire a une matrice ».

En effet, une base est une partie libre et génératrice (définition 4.5). Or une partie d'un ensemble n'est pas muni d'un ordre. Toutes les permutations de colonnes de la matrice sont encore possible d'après l'ordre que l'on met sur les vecteurs de la base.

Encore une fois, la base canonique n'a pas de problème parce que les $\{e_i\}$ de $\mathbb{R}^n$ viennent avec un ordre indiscutable. Plus généralement, très souvent, lorsqu'on construit une base, la construction suggère un ordre.

4.3.3 Déterminant

Définition 4.74.
*Si $A \in \mathbb{M}(n, \mathbb{K})$ nous définissons le **déterminant** de A par la formule*

$$\det(A) = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} A_{i,\sigma(i)} \tag{4.94}$$

où la somme est effectuée sur tous les éléments du groupe symétrique[35] *S_n et où $(-1)^\sigma$ représente la parité de la permutation σ.*

En se souvenant que $|S_n| = n!$, nous sommes frappés de stupeur devant le fait que le nombre de termes dans la somme croît de façon factorielle (c'est plus qu'exponentiel, pour info) en la taille de la matrice. Cette formule est donc sans espoir pour une matrice plus grande que 3×3 ou à la rigueur 4×4 à la main. À l'ordinateur, il est possible de monter plus haut, mais pas tellement.

4.3.4 Déterminant en petite dimension

En dimension deux, le déterminant de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est le nombre

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb. \tag{4.95}$$

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

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{4.96}$$

33. Au passage, non, les coordonnées polaires ne sont pas une base de $\mathbb{R}^2$. C'est un système de coordonnées, et ce n'est pas la même chose.

34. Dont l'auteur de ces lignes fait partie.

35. Pour le groupe symétrique, c'est la définition 1.179, le fait que ce soit un groupe fini est le lemme 1.182, et pour la somme sur un groupe fini c'est la définition 1.254.

4.3.5 Manipulations de lignes et de colonnes

Nous voudrions savoir ce qu'il se passe avec le déterminant d'une matrice lorsque nous substituons à une ligne ou une colonne, une combinaison des autres lignes et colonnes. Lorsqu'une matrice est donnée, nous notons C_j sa j^{e} colonne.

Lemme 4.75 ([1]).
Si A et B sont des matrices, alors

$$(AB)^t = B^t A^t. \tag{4.97}$$

Démonstration. Il suffit de calculer les éléments de matrice :

$$(AB)^t_{i,j} = (AB)_{j,i} = \sum_k A_{j,k} B_{k,i} = \sum_k B^t_{i,k} A^t_{k,j} = (B^t A^t)_{i,j}. \tag{4.98}$$

□

Lemme 4.76 ([1, 114]).
Si A est une matrice, alors $\det(A) = \det(A^t)$.

Démonstration. Nous commençons par écrire la définition du déterminant :

$$\det(A^t) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^n (A^t)_{i,\sigma(i)} = \sum_\sigma \epsilon(\sigma) \prod_i A_{\sigma(i),i}. \tag{4.99}$$

Pour chaque σ séparément, en utilisant la proposition 1.261 pour ré-indexer le produit :

$$\prod_i A_{\sigma(i),i} = \prod_i A_{i,\sigma^{-1}(i)}. \tag{4.100}$$

Nous profitons du fait que l'application $\varphi\colon S_n \to S_n$ donnée par $\varphi(\sigma) = \sigma^{-1}$ soit une permutation de S_n pour appliquer la définition 1.254 et faire la somme sur σ^{-1} :

$$\det(A^t) = \sum_\sigma \epsilon(\sigma) \prod_i A_{i,\sigma^{-1}(i)} = \sum_\sigma \epsilon(\sigma^{-1}) \prod_i A_{i,\sigma(i)} = \det(A) \tag{4.101}$$

où nous avons utilisé le fait que $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$ (corolaire 1.206). □

Le fait que $\det(A) = \det(A^t)$ permet, dans toutes les propositions du type « ce qui arrive au déterminant si on change telle ligne ou colonne » de ne donner qu'une preuve pour la partie « ligne » et déduire automatiquement le cas « colonne ». Le lemme suivant donne un exemple d'utilisation.

Lemme 4.77 ([1]).
Soit une matrice A. Nous considérons la matrice B obtenue à partir de A par la permutation de lignes $L_k \leftrightarrow L_l$ ainsi que la matrice C obtenue à partir de A^t par la permutation de colonnes $C_k \leftrightarrow C_l$.

Alors $C^t = B$.

Démonstration. Calculons les éléments de matrice de C :

$$C_{i,j} = \begin{cases} (A^t)_{i,j} & \text{si } j \neq k, j \neq l \\ (A^t)_{i,k} & \text{si } j = l \\ (A^t)_{i,l} & \text{si } j = k \end{cases} = \begin{cases} A_{j,i} & \text{si } j \neq k, j \neq l \\ A_{k,i} & \text{si } j = l \\ A_{l,i} & \text{si } j = k. \end{cases} \tag{4.102}$$

Ensuite nous prouvons que $C^t = B$ en écrivant les éléments de C^t :

$$(C^t)_{i,j} = C_{j,i} = \begin{cases} A_{i,j} & \text{si } i \neq k, i \neq l \\ A_{k,j} & \text{si } i = l \\ A_{l,j} & \text{si } i = k. \end{cases} \tag{4.103}$$

Cette dernière expression est la matrice A après permutation des lignes $L_k \leftrightarrow L_l$, c'est-à-dire la matrice B. □

Pour la suite nous écrivons δ la matrice « identité », c'est-à-dire celle dont les entrées sont précisément les $\delta_{i,k}$. Nous écrivons également $E_{i,j}$ la matrice contenant des zéros partout sauf en (i,j) où elle a un 1, c'est-à-dire

$$(E_{i,j})_{k,l} = \delta_{i,k}\delta_{j,l}. \tag{4.104}$$

Proposition 4.78 (Permuter des lignes ou des colonnes $L_k \leftrightarrow L_l$[115, 1])**.**
Soient une matrice $A \in \mathbb{M}(n,\mathbb{K})$, deux entiers $k \neq l$ inférieurs ou égaux à n.

(1) Si B est la matrice obtenue à partir de A en permutant deux lignes ou deux colonnes, alors

$$\det(A) = -\det(B). \tag{4.105}$$

(2) Si B est la matrice obtenue à partir de A par la permutation de lignes $L_k \leftrightarrow L_l$. Alors

$$B = SA \tag{4.106}$$

avec $S = \delta + E_{k,l} + E_{l,k} - E_{k,k} - E_{l,l}$.
Autrement dit : la matrice S est une matrice de permutations de lignes.

(3) La matrice S vérifie $\det(S) = -1$

(4) Nous avons

$$\det(SA) = \det(S)\det(A). \tag{4.107}$$

Démonstration. Point par point

(i) **(1) pour les colonnes** Soient k et l fixés, et considérons la permutation des colonnes C_k et C_l. Nous notons α la permutation (k,l) dans S_n (groupe symétrique, définition 1.179). Nous avons

$$B_{i,j} = A_{i,\alpha(j)}, \tag{4.108}$$

ou encore : $A_{i,j} = B_{i,\alpha(j)}$. Par définition,

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma)\prod_{i=1}^{n} A_{i,\sigma(i)} \tag{4.109}$$

C'est le moment d'utiliser la proposition 1.257 à propos de somme sur des groupes avec $G = S_n$, $h = \alpha$ et

$$f(\sigma) = \epsilon(\sigma)\prod_i A_{i,\sigma(i)}. \tag{4.110}$$

Nous savons que $\epsilon(\alpha) = -1$ et que ϵ est un morphisme par la proposition 1.205(1), donc

$$f(\alpha\sigma) = \epsilon(\alpha\sigma)\prod_i A_{i,(\alpha\sigma)(i)} = -\epsilon(\sigma)\prod_i B_{i,\sigma(i)}. \tag{4.111}$$

Avec ça, nous concluons :

$$\det(A) = \sum_{\sigma\in S_n} f(\sigma) = \sum_{\sigma} f(\alpha\sigma) = -\sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n} B_{i,\sigma(i)} = -\det(B). \tag{4.112}$$

(ii) **(1) pour les lignes** Que se passe-t-il si nous permutons les lignes L_k et L_l ? Si nous notons B' la matrice obtenue à partir de A par la permutation de lignes $L_k \leftrightarrow L_l$, et C celle obtenue de A^t après permutation de colonnes $C_k \leftrightarrow C_l$ alors nous avons $C^t = B'$. Le lemme 4.77 nous dit que $C^t = B'$. En utilisant le lemme 4.76 sur le déterminant de la transposée,

$$\det(B') = \det(C^t) = \det(C) = -\det(A^t) = -\det(A). \tag{4.113}$$

Voilà qui prouve le résultat pour les permutation de lignes.

(iii) **(2)** Si $k = l$, il n'y a pas de permutation, et il est vite vu que la matrice S est l'identité parce qu'il y a quatre fois le terme $E_{k,k}$. Nous supposons donc que $k \neq l$; en particulier $\delta_{k,l} = 0$.

Il s'agit surtout d'un beau calcul :

$$(SA)_{i,j} = \sum_m S_{i,m}A_{m,j} = A_{i,j} + \sum_m (\delta_{k,i}\delta_{l,m} + \delta_{l,i}\delta_{l,m} - \delta_{k,i}\delta_{k,m} - \delta_{l,i}\delta_{l,m})A_{m,j} \tag{4.114a}$$

$$= A_{i,j} + \delta_{k,i}A_{l,j} + \delta_{l,i}A_{k,j} - \delta_{k,i}A_{k,j} - \delta_{l,i}A_{l,j}. \tag{4.114b}$$

Si $i \neq j$ et $i \neq l$, alors $(SA)_{i,j} = A_{i,j}$. Si $i = k$, alors

$$(SA)_{k,j} = A_{k,j} + A_{l,j} - A_{k,j} = A_{l,j}, \tag{4.115}$$

c'est-à-dire que la k^{e} ligne de SA est la l^{e} ligne de A.

Avec $i = l$ nous obtenons la k^{e} ligne de A.

Tout cela montre que SA est la matrice A dans laquelle les lignes k et l ont été échangées, c'est-à-dire $SA = B$.

(iv) **(3)** En utilisant la définition du déterminant,

$$\det(S) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} S_{i,\sigma(i)} \tag{4.116a}$$

$$= \sum_{\sigma} \epsilon(\sigma) \prod_i \big(\delta_{i,\sigma(i)} + \delta_{k,i}\delta_{l,\sigma(i)} + \delta_{l,i}\delta_{k,\sigma(i)} - \delta_{k,i}\delta_{k,\sigma(i)} - \delta_{l,i}\delta_{l,\sigma(i)}\big). \tag{4.116b}$$

Nous utilisons l'associativité et la commutativité du produit pour séparer les facteurs $i = k$ et $i = l$ des autres :

$$\det(S) = \sum_{\sigma} \epsilon(\sigma) \prod_{\substack{i \neq k \\ i \neq l}} \delta_{i,\sigma(i)}(\delta_{k,\sigma(k)} + \delta_{l,\sigma(k)} - \delta_{k,\sigma(k)})(\delta_{l,\sigma(l)} + \delta_{k,\sigma(l)} - \delta_{l,\sigma(l)}). \tag{4.117}$$

À cause des facteurs $i \neq k$ et $i \neq l$, les σ pour lesquels le tout n'est pas nul doivent vérifier $\delta_{i,\sigma(i)} = 1$ pour tout i différent de k et l. Les deux seuls sont donc $\sigma = \mathrm{Id}$ et la permutation $\sigma = (k, l)$. Pour $\sigma = \mathrm{Id}$, nous avons

$$\prod_{\substack{i \neq k \\ i \neq l}} \delta_{i,i}(\delta_{k,k} + \delta_{l,k} - \delta_{k,k})(\delta_{l,l} + \delta_{k,l} - \delta_{l,l}) = 0. \tag{4.118}$$

Dernier espoir : $\sigma = (k, l)$. Pour ce terme nous avons $\epsilon(\sigma) = -1$ et

$$\prod_{\substack{i \neq k \\ i \neq l}} \delta_{i,i}(\delta_{k,l} + \delta_{l,l} - \delta_{k,l})(\delta_{l,k} + \delta_{k,k} - \delta_{l,k}) = 1. \tag{4.119}$$

Au final dans $\det(S)$, il n'y a que le terme $\sigma = (k, l)$ qui est non nul, et il vaut -1. Donc

$$\det(S) = -1. \tag{4.120}$$

(v) **(4)** Il s'agit de mettre bout à bout les points déjà prouvés :

$$\det(SA) = -\det(A) = \det(S)\det(A). \tag{4.121}$$

□

Corolaire 4.79 ([115]).
Soit une matrice $A \in \mathbb{M}(n, \mathbb{K})$. Si deux lignes ou deux colonnes de A sont égales, alors $\det(A) = 0$.

Démonstration. Si deux colonnes sont égales, la matrice ne change pas lorsqu'on les permute, alors que le déterminant change de signe. La seule possibilité est que $\det(A) = -\det(A)$, ce qui signifie que $\det(A) = 0$. □

Notons que si pour $k \neq l$ nous avons $C_k = \lambda C_l$, alors nous avons aussi $\det(A) = 0$.

La réciproque n'est pas vraie : il existe des matrices dont le déterminant est nul et dont aucune entrée n'est nulle. Par exemple

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}. \tag{4.122}$$

Proposition 4.80 ([115])**.**
Soient $A \in \mathbb{M}(n, \mathbb{K})$, et $v \in \mathbb{K}^n$. Si B est la matrice A avec la substitution $L_j \to L_j + v$ et C est la matrice A avec la substitution $L_j \to v$, alors

$$\det(B) = \det(A) + \det(C). \tag{4.123}$$

Démonstration. En utilisant l'associativité de la multiplication,

$$\begin{aligned}
\det(B) &= \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} B_{i,\sigma(i)} && (4.124a)\\
&= \sum_{\sigma} \epsilon(\sigma) \big(\prod_{i \neq j} B_{i,\sigma(i)}\big) B_{j,\sigma(j)} && (4.124b)\\
&= \sum_{\sigma} \epsilon(\sigma) \big(\prod_{i \neq j} A_{i,\sigma(i)}\big) (A_{j,\sigma(j)} + v_{\sigma(j)}) && (4.124c)\\
&= \sum_{\sigma} \epsilon(\sigma) \prod_{i} A_{i,\sigma(i)} + \sum_{\sigma} \epsilon(\sigma) \prod_{i \neq j} C_{i,\sigma(i)} v_{\sigma(j)} && (4.124d)\\
&= \det(A) + \sum_{\sigma} \epsilon(\sigma) \prod_{i \neq j} C_{i,\sigma(i)} C_{j,\sigma(j)} && (4.124e)\\
&= \det(A) + \det(C). && (4.124f)
\end{aligned}$$

Justifications :

— 4.124d parce que pour $i \neq j$ nous avons $A_{i,\sigma(i)} = C_{i,\sigma(i)}$

— 4.124e parce que $v_{\sigma(j)} = C_{j,\sigma(j)}$.

□

Proposition 4.81 (Combinaison de lignes ou colonnes $L_k \to L_k + \lambda L_l$[115])**.**
Soient une matrice $A \in \mathbb{M}(n, \mathbb{K})$, deux entiers $k \neq l$ inférieurs ou égaux à n.

(1) Si B est la matrice obtenue à partir de A par la substitution $L_k \to L_k + \lambda L_l$ ou $C_k \to C_k + \lambda C_l$, alors

$$\det(A) = \det(B). \tag{4.125}$$

(2) Si B est la matrice A dans laquelle nous avons opéré la substitution $L_k \to L_k + \lambda L_l$, alors

$$B = UA \tag{4.126}$$

avec $U = \delta + \lambda E_{k,l}$, c'est-à-dire que U est une matrice de combinaison de lignes.

(3) La matrice U vérifie $\det(U) = 1$.

(4) Nous avons

$$\det(UA) = \det(U)\det(A). \tag{4.127}$$

Démonstration. Point par point.

(i) **(1)** Soit la matrice C obtenue à partir de A par $L_k \to \lambda L_l$. En considérant le vecteur $v = \lambda L_l$, nous sommes dans la situation de la proposition 4.80. Donc

$$\det(B) = \det(A) + \det(C). \tag{4.128}$$

Mais dans la matrice C, nous avons $L_k = \lambda L_l$, ce qui implique $\det(C) = 0$ par le corolaire 4.79. Donc $\det(A) = \det(B)$ comme il se devait.

(ii) **(2)** Encore un calcul :

$$(UA)_{i,j} = \sum_m \big(\delta_{i,m} + \lambda(E_{k,l})_{i,m}\big)A_{m,j} = A_{i,j} + \lambda \sum_m \delta_{k,i}\delta_{l,m}A_{m,j} = A_{i,j} + \lambda\delta_{l,i}A_{k,j}. \tag{4.129}$$

Cela donne, pour $i = k$ la ligne

$$(UA)_{k,j} = A_{k,j} + \lambda A_{l,j}, \tag{4.130}$$

ce qui correspond bien à $L_k \to L_k + \lambda L_l$.

(iii) **(3)** Nous calculons le déterminant de $U = \delta + \lambda E_{k,l}$ avec $k \neq l$. Nous avons dans un premier temps :

$$\det(U) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} (\delta_{i,\sigma(i)} + \lambda\delta_{k,i}\delta_{l,\sigma(i)}). \tag{4.131}$$

Puisque nous avons toujours $\delta_{k,i}\delta_{l,i} = 0$, le terme $\sigma = \mathrm{Id}$ donne 1.

Pour les $\sigma \neq \mathrm{Id}$, le facteur $\lambda\delta_{k,i}\delta_{l,\sigma(i)}$ ne s'annule pas, uniquement si $i = k$ et $\sigma(i) = l$. Donc le seul terme non nul autre que $\sigma = \mathrm{Id}$ peut provenir de $\sigma = (k, l)$. Pour ce terme, nous isolons les termes $i = l$ et $i = k$:

$$(\delta_{k,\sigma(k)} + \lambda\delta_{k,k}\delta_{k,\sigma(k)})(\delta_{l,\sigma(l)} + \lambda\delta_{k,l}\delta_{k,\sigma(l)}). \tag{4.132}$$

Le dernier facteur est nul.

(iv) **(4)** En mettant bout à bout les résultats prouvés,

$$\det(UA) = \det(A) = \det(U)\det(A). \tag{4.133}$$

□

Proposition 4.82 (Multiplication par un scalaire d'une ligne ou colonne $L_k \to \lambda L_k$[115])**.**
Soient une matrice $A \in \mathbb{M}(n, \mathbb{K})$, un entier $k \neq l$ inférieur ou égal à n. Soit la matrice B obtenue à partir de A en multipliant la ligne L_k par $\lambda \in \mathbb{K}$.

(1) $\det(B) = \lambda \det(A)$

(2) En considérant la matrice $T = \delta + (\lambda - 1)E_{k,k}$, nous avons

$$B = TA, \tag{4.134}$$

c'est-à-dire que la matrice T est une matrice de multiplication de ligne par un scalaire.

(3) Nous avons $\det(T) = \lambda$.

(4) Et aussi : $\det(TA) = \det(T)\det(A)$

Démonstration. Point par point.

(i) **(1)** La matrice B est donnée par les éléments

$$B_{i,j} = \begin{cases} A_{i,j} & \text{si } i \neq k \\ \lambda A_{i,j} & \text{si } i = k \end{cases} \tag{4.135}$$

c'est-à-dire $B_{i,j} = \big(1 + (\lambda - 1)\delta_{i,k}\big)A_{i,j}$. Nous mettons cela dans la définition du déterminant de B :

$$\det(B) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} B_{i,\sigma(i)} = \sum_{\sigma} \prod_{i} \big(1 + (\lambda - 1)\delta_{\sigma(i),k} A_{i,\sigma(i)}\big). \tag{4.136}$$

L'associativité du produit dans $\mathbb{K}$ nous permet de séparer le produit de la façon suivante :

$$\prod_{i=1}^{n} \big(1 + (\lambda - 1)\delta_{\sigma(i),k}\big)A_{i,\sigma(i)} = \prod_{i} \big(1 + (\lambda - 1)\delta_{\sigma(i),k}\big) \prod_{i} A_{i,\sigma(i)} = \lambda \prod_{i} A_{i,\sigma(i)}. \tag{4.137}$$

En remettant dans (4.136), nous trouvons $\det(B) = \det(A)$.

(ii) **(2)** C'est un cas particulier de la proposition 4.81(2) en prenant $k = l$ et en adaptant le λ.

(iii) **(3)** Nous calculons le déterminant de la matrice $T = \delta + (\lambda - 1)E_{k,k}$. La formule du déterminant donne

$$\det(T) = \sum_{\sigma} \epsilon(\sigma) \prod_{i=1}^{n} \big(\delta_{i,\sigma(i)} + (\lambda - 1)\delta_{k,i}\delta_{k,\sigma(i)}\big). \tag{4.138}$$

Si $i \neq \sigma(i)$, alors non seulement $\delta_{i,\sigma(i)} = 0$, mais en plus $\delta_{k,i}\delta_{k,\sigma(i)} = 0$. Donc seul $\sigma = \mathrm{Id}$ reste dans la somme sur $\sigma \in S_n$. Il reste donc

$$\det(T) = \prod_{i=1}^{n} \big(1 + (\lambda - 1)\delta_{k,i}\big) = \left(\prod_{i \neq k} 1\right)(1 + (\lambda - 1)) = \lambda \tag{4.139}$$

où nous avons utilisé encore l'associativité pour isoler le facteur $i = k$.

(iv) **(4)** Il faut mettre bout à bout les résultats déjà établis :

$$\det(TA) = \lambda \det(A) = \det(T)\det(A). \tag{4.140}$$

□

4.3.6 Réduction de Gauss

Nous avons vu les matrices d'opérations élémentaire sur les lignes et colonnes :

— Permutation de lignes $L_k \leftrightarrow L_l$: $S(n; k, l) = \delta + E_{k,l} + E_{l,k} - E_{k,k} - E_{l,l}$, proposition 4.78.

— Combinaisons de lignes $L_k \to L_k + \lambda L_l$: $U(n; k, l, \lambda) = \delta + \lambda E_{k,l}$, proposition 4.81.

— Multiplication d'une ligne par un scalaire $L_k \to \lambda L_k$: $T(n; k, \lambda) = \delta + (\lambda - 1)E_{k,k}$, proposition 4.82.

Ces matrices seront dans la suite, notées G. Et elles vérifient l'importante propriété

$$\det(GA) = \det(G)\det(A) \tag{4.141}$$

pour toute matrice A.

Proposition 4.83 (Réduction de Gauss[1])**.**
Soit une matrice $A \in \mathbb{M}(n, \mathbb{K})$ de déterminant non nul : $\det(A) \neq 0$. *Alors il existe des matrices $G_1, \dots, G_N$ toutes de type S, U ou T telles que*

$$G_1 \dots G_N A = \delta. \tag{4.142}$$

Démonstration. Nous procédons par récurrence sur n. D'abord pour $n = 1$, la matrice A contient un seul élément $A_{1,1}$ qui est non nul par hypothèse. Nous pouvons multiplier sa ligne par $1/A_{1,1}$ pour obtenir le résultat. Plus précisément, nous avons l'égalité

$$T(1; 1, \frac{1}{A_{1,1}})A = \delta \tag{4.143}$$

dans $\mathbb{M}(1,\mathbb{K})$. Notons que $\mathbb{K}$ est un corps (donc $A_{1,1}$ est inversible) commutatif, ce qui permet d'écrire $1/A_{1,1}$ sans ambiguïté.

Supposons le résultat prouvé pour n, et voyons ce qu'il se passe pour $n+1$. Puisque $\det(A) \neq 0$, aucune de ses colonnes n'est nulle (corolaire 4.79). Il existe donc un k tel que $A_{k,1} \neq 0$.

Par la proposition 4.78, la matrice

$$B^{(1)} = S(n+1;k,1)A \tag{4.144}$$

est une matrice telle que $B^{(1)}_{1,1} = A_{k,1} \neq 0$. Ensuite, par la proposition 4.82 la matrice

$$B^{(2)} = T(n+1;1,\frac{1}{A_{k,1}})B^{(1)} \tag{4.145}$$

vérifie $B^{(2)}_{1,1} = 1$.

Puisque la multiplication par la matrice $U(n+1;k;l;\lambda)$ réalise, par la proposition 4.81, la substitution $L_k \to L_k + \lambda L_l$, la matrice

$$B^{(3)} = \prod_{k=2}^{n+1} U(n+1;k,1,-B^{(1)}_{k,1})B^{(1)} \tag{4.146}$$

a toute sa première colonne nulle à l'exception de $B^{(3)}_{1,1} = 1$.

Nous n'avons pas donné de nom ni démontré de théorèmes à propos de la substitution $C_k \to C_k + \lambda C_l$. En passant éventuellement par les transposées et en utilisant les lemmes 4.75 et 4.76 nous obtenons une matrice $U'(n+1;k,l,\lambda)$ ayant la propriété que la matrice

$$B^{(4)} = \prod_{k=2}^{n+1} U'(n+1;k,1,-B^{(3)}_{1,k})B^{(3)} \tag{4.147}$$

vérifie $B^{(4)}_{1,j} = B^{(4)}_{j,1} = 0$ pour tout j sauf $j=1$. En d'autres termes, la matrice $B^{(4)}$ est de la forme

$$B^{(4)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix} \tag{4.148}$$

où A' est une matrice de taille n.

Voyons quelques propriétés de A'. Nous savons que

$$B^{(4)} = \prod_i G_i A \tag{4.149}$$

où les G_i sont de type S, T ou U. Puisque $\det(SA) = \det(S)\det(A)$ (et idem pour T et U), nous avons

$$\det(B^{(4)}) = \prod_i \det(G_i)\det(A), \tag{4.150}$$

et comme aucun des $\det(G_i)$ n'est nul, nous avons encore $\det(B^{(4)}) \neq 0$, ce qui implique $\det(A') \neq 0$.

La récurrence peut avoir lieu. Il existe des matrices G'_i telles que

$$G'_1 \dots G'_M A' = \delta \tag{4.151}$$

où les G'_i sont de taille n, ainsi que le δ. En remarquant que

$$S(n+1;k,l) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & S(n;k-1,l-1) & \\ 0 & & & \end{pmatrix}, \tag{4.152}$$

et pareillement pour les matrices T et U, nous voyons qu'en prenant

$$G_i = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & G'_i & \\ 0 & & & \end{pmatrix}, \tag{4.153}$$

nous avons

$$\prod_{i=1}^{M} G_i B^{(3)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \prod_{i=1}^{M} G'_i A' & \\ 0 & & & \end{pmatrix} = \delta_{n+1} \tag{4.154}$$

où nous avons mis un indice sur le dernier δ pour être plus explicite. □

4.3.7 Matrices inversibles

Proposition-Définition 4.84.
Soit une matrice $A \in \mathbb{M}(n, \mathbb{K})$. Si les matrices B_1 et B_2 de $\mathbb{M}(n, \mathbb{K})$ vérifient

$$AB_1 = B_1 A = \delta \tag{4.155}$$

et

$$AB_2 = B_2 A = \delta, \tag{4.156}$$

alors $B_1 = B_2$. Dans ce cas, nous disons que A est inversible et nous notons A^{-1} l'unique matrice telle que $AA^{-1} = A^{-1}A = \delta$.

Démonstration. La preuve est réalisée dans le cas général par le lemme 1.157. Mais si vous en voulez une preuve avec les notations d'ici, en voici une.

Nous avons $AB_1 = AB_2$. En multipliant à gauche par B_1, nous trouvons $B_1AB_1 = B_1AB_2$. En remplaçant B_1A par δ des deux côtés, il reste $B_1 = B_2$. □

Lemme 4.85 ([115])**.**
Si $A \in \mathbb{M}(n, \mathbb{K})$, alors il existe au plus une matrice $B \in \mathbb{M}(n, \mathbb{K})$ telle que $AB = \delta$.

Démonstration. Soient des matrices $B, C \in \mathbb{M}(n, \mathbb{K})$ telles que $AB = AC = \delta$. Nous allons montrer que $B = C$.

Pour cela nous considérons les applications linéaires $f_A, f_B, f_C \in \mathrm{End}(\mathbb{K}^n)$ associées par la proposition 4.68. Puisque $AB = \delta$, par la proposition 4.70, nous avons $f_A \circ f_B = f_{AB} = \mathrm{Id}$. La proposition 4.52 nous dit alors que f_A et f_B sont bijectives.

En particulier, comme $\{e_i\}$ est une base, son image par f_B est une base par la proposition 4.41. La proposition 4.51 dit alors que $\{f_B(e_i)\}$ est une base. Nous décomposons $f_B(e_k) - f_C(e_k)$ dans cette base :

$$f_B(e_k) - f_C(e_k) = \sum_j \alpha_j f_B(e_j) \tag{4.157}$$

où les α_j dépendent à priori de k. Puisque $f_A \circ (f_B - f_C) = 0$, nous avons

$$0 = f_A\big(f_B(e_k) - f_C(e_k)\big) = \sum_j (f_A \circ f_B)(e_j) = \sum_j \alpha_j e_j. \tag{4.158}$$

Donc les α_j sont tous nuls.

Nous en déduisons que $f_B(e_k) = f_C(e_k)$, et donc $f_B = f_C$. Cela implique que $B = C$ par la proposition 4.68(4). □

Proposition 4.86 ([115])**.**
Si $A, B \in \mathbb{M}(n, \mathbb{K})$ vérifient $AB = \delta$, alors $BA = \delta$.

Démonstration. L'astuce est de poser $C = BA - \delta + B$ et de montrer que $C = B$. Pour cela, un rapide calcul commence par montrer que

$$AC = ABA - A + AB = AB = \delta. \tag{4.159}$$

Donc C est également un inverse à droite de A. Le lemme 4.85 donne alors $C = B$. □

Corolaire 4.87.
Soit $A \in \mathbb{M}(n, \mathbb{K})$. Si il existe $B \in \mathbb{M}(n, \mathbb{K})$ tel que $AB = \delta$, alors A est inversible et son inverse est B.

Démonstration. Il s'agit d'une paraphrase de la proposition 4.86 et de la définition 4.84. □

Lemme 4.88.
Si une matrice A n'est pas inversible, alors le produit AB n'est inversible pour aucune matrice B.

Démonstration. Soient une matrice non inversible A, ainsi qu'une matrice quelconque B. Supposons que AB soit inversible. Alors

$$AB(AB^{-1}) = \delta. \tag{4.160}$$

Donc la matrice $B(AB^{-1})$ est un inverse de A. Contradiction. □

Proposition 4.89.
Une matrice est inversible si et seulement si son application linéaire associée est inversible. Dans ce cas, nous avons

$$f_A^{-1} = f_{A^{-1}}. \tag{4.161}$$

Démonstration. Dans le sens direct, si A est inversible nous avons $AA^{-1} = \delta$. Donc

$$f_A \circ f_{A^{-1}} = f_{AA^{-1}} = f_\delta = \mathrm{Id} \tag{4.162}$$

où nous avons utilisé la proposition 4.70 pour la composition et la proposition 4.71 pour l'identité. L'égalité (4.162) indique que f_A est inversible et que son inverse est $f_{A^{-1}}$.

Dans l'autre sens, l'application f_A^{-1} existe. Soit $B \in \mathbb{M}(n, \mathbb{K})$ sa matrice. Alors nous avons

$$f_\delta = \mathrm{Id} = f_A \circ f_B = f_{AB}. \tag{4.163}$$

Le fait que l'application $\psi\colon A \to f_A$ soit une bijection[36] implique que $AB = \delta$, c'est-à-dire que A est inversible et que $B = A^{-1}$. □

Lemme 4.90 ([1])**.**
Soient une matrice inversible $A \in \mathbb{M}(n, \mathbb{R})$ et $r < n$. Il existe une permutation $\sigma \in S_n$ telle que la matrice $a \in \mathbb{M}(r, \mathbb{R})$ donnée par

$$a_{i,j} = A_{i,\sigma(j)} \tag{4.164}$$

soit inversible.

4.3.8 Inversibilité et déterminant

Proposition 4.91.
Une matrice au déterminant non nul est inversible.

Démonstration. Si A est une matrice telle que $\det(A) \neq 0$, alors la proposition 4.83 nous donne des matrices $G_1, \dots, G_N$ telles que

$$G_1 \dots G_N A = \delta. \tag{4.165}$$

Donc la matrice $G_1 \dots G_N$ est un inverse de A par le corolaire 4.87. □

36. Proposition 4.68(4).

Proposition 4.92.
Si une matrice A a une ligne ou une colonne de zéros, alors

(1) $\det(A) = 0$,

(2) *A n'est pas inversible.*

Démonstration. Par définition, nous avons

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}. \tag{4.166}$$

Si la k^{e} ligne est nulle, alors $A_{k,\sigma(k)} = 0$ pour tout σ. Donc tous les produits contiennent un facteur nul. Donc $\det(A) = 0$.

Pour toute matrice B nous avons

$$(AB)_{k,k} = \sum_{l} A_{k,l} B_{l,k}. \tag{4.167}$$

Si la k^{e} ligne de A est nulle nous avons $(AB)_{k,k} = 0$ et donc pas $AB = \delta$. Donc A n'est pas inversible. □

4.3.9 Quelques ensembles de matrices particuliers

Certains ensembles de matrices ont une importance particulière, que nous développerons plus tard.

Définition 4.93 (Groupe linéaire de matrices)**.**
On note $\mathrm{GL}(n, \mathbb{A})$ *l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{A}$, qui sont inversibles. En d'autres termes,* $\mathrm{GL}(n, \mathbb{A}) = U(\mathbb{M}(n, \mathbb{A}))$.

Définition 4.94 (Groupe orthogonal de matrices)**.**
On dit qu'une matrice A est ***orthogonale*** *si son inverse est sa transposée, c'est-à-dire si $A^{-1} = A^t$. On note* $\mathrm{O}(n, \mathbb{A})$ *l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{A}$, qui sont orthogonales.*

4.3.10 Déterminant et combinaisons de lignes et colonnes

Proposition 4.95.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$ telles que $\det(A) \neq 0$ *et* $\det(B) \neq 0$. *Alors*

$$\det(AB) = \det(A)\det(B). \tag{4.168}$$

Démonstration. La proposition 4.83 nous donne des matrices de permutations de lignes et de colonnes $G_1, \dots, G_N$ et $G'_1, \dots, G'_N$ telles que[37]

$$G_1 \dots G_N A = \delta \tag{4.169a}$$
$$G'_1 \dots G'_N B = \delta. \tag{4.169b}$$

Nous avons

$$(G'_1 \dots G'_N) \underbrace{(G_1 \dots G_N) A}_{=\delta} B = \delta. \tag{4.170}$$

En prenant le déterminant des deux côtés et en tenant compte de (4.141),

$$1 = \det(\delta) = \det\big(G'_1 \dots G'_N G_1 \dots G_N AB\big) = \det(G'_1 \dots G'_N)\det(G_1 \dots G_N)\det(AB). \tag{4.171}$$

37. Les plus acharnés préciseront que pour avoir le même N des deux côtés, il a fallu compléter avec des matrices δ là où il y en avait le moins.

Mais en même temps, les équations 4.169 donnent

$$\det(G_1 \dots G_N) = \det(A)^{-1} \tag{4.172a}$$
$$\det(G'_1 \dots G'_N) = \det(B)^{-1}. \tag{4.172b}$$

Cela pour dire que

$$1 = \det(A)^{-1} \det(B)^{-1} \det(AB), \tag{4.173}$$

et donc ce qu'il nous fallait. □

Proposition 4.96.
Soient des matrices $A, B \in \mathbb{M}(n, \mathbb{K})$ telles que $\det(A) = 0$ et $\det(B) \neq 0$. Alors

$$\det(AB) = \det(BA) = \det(A)\det(B) = 0. \tag{4.174}$$

Démonstration. Il existe des matrices de manipulations de lignes et de colonnes $G_1, \dots, G_N$ telles que $G_1 \dots G_N B = \delta$. Donc

$$0 = \det(A) = \det(G_1 \dots G_N BA) = \det(G_1 \dots G_N)\det(BA). \tag{4.175}$$

Donc $\det(BA) = 0$. □

4.3.11 Transvections

Nous nommons $E_{i,j}$ la matrice remplie de zéros sauf à la case i, j qui vaut 1. Autrement dit

$$(E_{i,j})_{k,l} = \delta_{i,k}\delta_{j,l}. \tag{4.176}$$

Définition 4.97.
*Une **matrice de transvection** est une matrice de la forme*

$$T_{i,j}(\lambda) = \mathrm{Id} + \lambda E_{i,j} \tag{4.177}$$

avec $i \neq j$.

*Une **matrice de dilatation** est une matrice de la forme*

$$D_i(\lambda) = \mathrm{Id} + (\lambda - 1)E_{i,i}. \tag{4.178}$$

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 permutation** associée est la matrice d'entrées*

$$(P_\sigma)_{i,j} = \delta_{i,\sigma(j)}. \tag{4.179}$$

Lemme 4.98.
La matrice $T_{i,j}(\lambda)A = (\mathbb{1} + \lambda E_{i,j})A$ est la matrice A à qui on a effectué la substitution

$$L_i \to L_i + \lambda L_j. \tag{4.180}$$

La matrice $AT_{i,j}(\lambda)$ est la substitution

$$C_j \to C_j + \lambda C_i. \tag{4.181}$$

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 k, l de la matrice $E_{i,j}A$:

$$(E_{i,j}A)_{k,l} = \sum_m (E_{i,j})_{k,m} A_{m,l} \tag{4.182a}$$

$$= \sum_m \delta_{i,k}\delta_{j,m} A_{m,l} \tag{4.182b}$$

$$= \delta_{i,k} A_{j,l}. \tag{4.182c}$$

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

$$A + \lambda E_{i,j} A \tag{4.183}$$

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

$$(AE_{i,j}) = A_{k,i}\delta_{j,l}. \tag{4.184}$$

Pour les matrices de permutation, nous avons

$$(AP_\sigma)_{k,l} = A_{k,\sigma(l)} \tag{4.185}$$

et

$$(P_\sigma A)_{k,l} = \sum_m \delta_{k,\sigma(m)} A_{m,l} = \sum_m \delta_{\sigma^{-1}(k),m} A_{m,l} = A_{\sigma^{-1}(k),l}. \tag{4.186}$$

□

4.3.12 Mineur, rang

Pour la définition du rang d'une matrice, nous en donnons une qui est clairement inspirée de l'application linéaire associée.

Définition 4.99 ([115]).
*Le **rang** d'une matrice de $\mathbb{M}(n, \mathbb{K})$ est la dimension de la partie de $\mathbb{K}^n$ engendrée par ses colonnes.*

Il est possible d'exprimer le rang d'une matrice de façon plus « intrinsèque » via le concept de mineur.

Définition 4.100 ([116]).
Les mineurs d'une matrice sont les déterminants de ses sous-matrices carrées.

Dans la suite nous désignerons souvent par le mot « mineur » la sous-matrice carrée elle-même au lieu de son déterminant.

Lorsque A est une matrice, nous notons f_A l'application linéaire associée à la matrice A par l'application (4.76).

Lemme 4.101.
Soit $\mathbb{K}$ un corps commutatif[38]*. Si A est une matrice carrée d'ordre n et de rang r à coefficients dans $\mathbb{K}$, alors il existe des vecteurs $(x_i)_{i=1,\dots,n}$ formant une base de $\mathbb{K}^n$ tels que*

$$f_A(x_i) \neq 0 \tag{4.187}$$

pour $i \leqslant r$ et

$$f_A(x_i) = 0 \tag{4.188}$$

pour $i > r$.

38. Comme toujours.

Démonstration. Soit V le sous-espace de $\mathbb{K}^n$ engendré par les colonnes de A. Nous considérons la base canonique $\{e_i\}$ de $\mathbb{K}^n$, ainsi que v_i le vecteur créé par la i^{e} colonne de A. Nous avons

$$v_i = f_A(e_i). \tag{4.189}$$

Les vecteurs v_i engendrent V, donc nous pouvons en extraire une base par le théorème 4.16(1). Soit donc $\{v_j\}_{i\in J}$ une base de V avec $J \subset \{1, \dots, n\}$.

La base de $\mathbb{K}^n$ que nous cherchons commence par les vecteurs $\{e_j\}_{j\in J}$. Ces vecteurs vérifient $f_A(e_j) = v_j \neq 0$ parce que des vecteurs d'une base ne sont jamais nuls.

Pour la suite de la base, nous pourrions penser au théorème de la base incomplète[39], mais les vecteurs ainsi complétant la base ne sont pas garantis de s'annuler par f_A. Voir l'exemple 4.102.

L'idée est d'utiliser le noyau de f_A qui est un sous-espace vectoriel par la proposition 4.37. Soit une base[40] $\{z_k\}$ de $\ker(f)$. Les vecteurs $\{e_j\}_{j\in J}$ forment une base de $\text{Image}(f_A)$. Puisque les z_i forment une base de $\ker(f_A)$, le théorème du rang 4.44 dit alors que $\{e_j\}_{j\in J} \cup \{z_k\}$ est une base de $\mathbb{K}^n$.

Il y a r éléments dans J parce que l'espace engendré par les colonnes de A est de dimension r par hypothèse. Donc il y a $n - r$ éléments dans les $\{z_k\}$ pour que le tout ait le bon nombre d'éléments. □

Exemple 4.102.
Soit la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}. \tag{4.190}$$

Elle est de rang 1. En suivant l'idée de la démonstration, nous commençons la base de $\mathbb{R}^2$ par le vecteur e_1 qui vérifie

$$f_A(e_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{4.191}$$

L'utilisation du théorème de la base incomplète ne permet pas de trouver un second vecteur de base v tel que $f_A(v) = 0$. En effet ce théorème donne juste l'existence d'une completion de la base, mais pas de propriétés particulières de la base obtenue. Elle pourrait donner $v = e_2$ comme second vecteur de base. Mais alors

$$f_A(v) = f_A(e_2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \neq 0. \tag{4.192}$$

Au contraire, le noyau de f_A est donné par le sous-espace engendré par $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Une base convenable est donc $\{e_1, e_1 - e_2\}$. △

Proposition 4.103.
Le rang d'une application linéaire[41] *est égal au rang de sa matrice*[42] *dans n'importe quelle base.*

4.3.13 Matrices équivalentes et semblables

Définition 4.104.
Deux relations d'équivalence entre les matrices.

(1) *Deux matrices A et B sont* ***équivalentes*** *dans $\mathbb{M}(n, \mathbb{K})$ si il existe $P, Q \in \text{GL}(n, \mathbb{K})$ telles que $A = PBQ^{-1}$.*

(2) *Deux matrices sont* ***semblables*** *si il existe une matrice $P \in \text{GL}(n, \mathbb{K})$ telle que $A = PBP^{-1}$.*

39. Théorème 4.12(2).
40. Cette base contient $n - r$ éléments, mais ce n'est pas très important pour la suite.
41. Définition 4.43.
42. Définition 4.99.

Lemme 4.105.
Une matrice de rang[43] *r dans $\mathbb{M}(n,\mathbb{K})$ est équivalente à la matrice par blocs*

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.193}$$

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 $i > r$, qui existe par le lemme 4.101.

Nous considérons la matrice inversible P telle que $Pe_i = f_i$; ses colonnes sont donc précisément les f_i, si bien que

$$APe_i = Af_i = \begin{cases} 0 & \text{si } i > r \\ \neq 0 & \text{sinon.} \end{cases} \tag{4.194}$$

La matrice AP se présente donc sous la forme

$$AP = \begin{pmatrix} M & 0 \\ * & 0 \end{pmatrix} \tag{4.195}$$

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

$$QAPe_i = \begin{cases} e_i & \text{si } i < r \\ 0 & \text{sinon.} \end{cases}. \tag{4.196}$$

Cela signifie que QAP est la matrice J_r. □

Corolaire 4.106 (Équivalence et rang)**.**
Deux matrices sont équivalentes[44] *si et seulement si elles sont de même rang.*

Démonstration. D'abord il y a des implicites dans l'énoncé. Puisque nous voulons, soit par 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 carrées d'ordre n.

Par le lemme 4.105, 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[45] : $A = PBQ^{-1}$ avec P et Q inversibles. Alors

$$\begin{aligned} \mathrm{Image}(PBQ^{-1}) &= \{PBQ^{-1}v \text{ tel que } v \in \mathbb{K}^n\} && (4.197a) \\ &= PB\underbrace{\{Q^{-1}v \text{ tel que } v \in \mathbb{K}^n\}}_{=\mathbb{K}^n} && (4.197b) \\ &= P\big(B(\mathbb{K}^n)\big). && (4.197c) \end{aligned}$$

L'ensemble $B(\mathbb{K}^n)$ est un sous-espace vectoriel de $\mathbb{K}^n$. Comme le rang de P est maximum, la dimension de $P\big(B(\mathbb{K}^n)\big)$ est la même que celle de $B(\mathbb{K}^n)$. Par conséquent

$$\dim\Big(\mathrm{Image}(PBQ^{-1})\Big) = \dim\big(B(\mathbb{K}^n)\big) = \mathrm{rk}(B). \tag{4.198}$$

Le membre de gauche de cela n'est autre que $\mathrm{rk}(A) = \dim\big(\mathrm{Image}(PBQ^{-1})\big)$. □

43. Définition 4.99.
44. Définition 4.104(1).
45. Définition 4.104.

4.3.14 Algorithme des facteurs invariants

Proposition 4.107 (Algorithme des facteurs invariants[89]).
Soit $(\mathbb{A}, \delta)$ *un anneau euclidien muni de son stathme et* $U \in \mathbb{M}(n \times m, \mathbb{A})$. *Alors il existe* $d_1, \dots, d_s \in \mathbb{A}^*$ *et des matrices* $P \in \mathrm{GL}(m, \mathbb{A})$, $Q \in \mathrm{GL}(n, \mathbb{A})$ *tels que nous ayons*

$$U = P \begin{pmatrix} d_1 & & & \\ & \ddots & & 0 \\ & & d_s & \\ & 0 & & 0 \end{pmatrix} Q \tag{4.199}$$

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

$$\begin{aligned} C_1 &\leftrightarrow C_{j_0} \\ L_1 &\leftrightarrow L_{i_0} \end{aligned} \tag{4.200}$$

Ensuite nous traitons la première colonne jusqu'à amener des zéros partout en dessous de $u_{1,1}$ de la façon suivante : pour chaque ligne successivement nous calculons la division euclidienne

$$u_{i,1} = q u_{1,1} + r_i, \tag{4.201}$$

et nous faisons

$$L_i \to L_i - q L_1, \tag{4.202}$$

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 contents [46]. Dans ce cas nous permutons $L_1 \leftrightarrow L_i$, ce qui aura pour effet de strictement diminuer le stathme de $u_{1,1}$ parce qu'on va mettre en $u_{1,1}$ le nombre r_i dont le stathme est strictement plus petit que celui de $u_{1,1}$.

En faisant ce jeu de division euclidienne puis échange, on diminue toujours le stathme de $u_{1,1}$, donc ça finit par s'arrêter, c'est-à-dire qu'à un certain moment, la division euclidienne de $u_{i,1}$ par $u_{1,1}$ va donner un reste nul et nous serons contents.

Une fois la première colonne ramenée à la forme

$$C_1 = \begin{pmatrix} u_{1,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.203}$$

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 [47]

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \tag{4.204}$$

46. Si il est zéro, nous passons à la ligne suivante

47. Nous nommons toujours par la même lettre U la matrice originale et la matrice modifiée, comme il est d'usage en informatique.

Si l'élément $u_{1,1}$ ne divise pas un des éléments de A, disons $a_{i,j}$, alors nous opérons

$$C_1 \to C_1 - C_j. \tag{4.205}$$

Cela nous détruit un peu la première colonne, mais ne change pas $u_{1,1}$. Nous avons maintenant

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ * & & & \\ u_{i,j} & & A & \\ * & & & \\ 0 & & & \end{pmatrix} \tag{4.206}$$

Et nous refaisons tout le jeu depuis le début. Cependant, lorsque nous allons nous attaquer à la ligne i, $u_{1,1}$ ne divisera pas $u_{i,j}$, 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 $u_{1,1}$ et réciproquement, diminuera encore strictement le stathme. Encore une fois, nous allons travailler jusqu'à avoir la matrice sous la forme

$$U = \begin{pmatrix} u_{1,1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}, \tag{4.207}$$

sauf que cette fois, le stathme de $u_{1,1}$ est strictement plus petit que la fois précédente. Si $u_{1,1}$ 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 (4.207) avec $u_{1,1}$ 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

$$U = \begin{pmatrix} u_{1,1} & & 0 \\ & u_{2,2} & \\ 0 & & B \end{pmatrix}. \tag{4.208}$$

Sous cette forme nous avons $u_{1,1} \mid u_{2,2}$ et $u_{1,1}$ divise tous les éléments de B. En effet $u_{1,1}$ 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 sur 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). □

4.4 Changement de base

Soit un espace vectoriel E muni de deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$. Les deux bases sont liées entre elles par

$$f_\alpha = \sum_i Q_{i,\alpha} e_i. \tag{4.209}$$

Ici Q n'est pas une application linéaire $E \to E$: Q est seulement un tableau de nombres, donnant les coordonnées des vecteurs f_α dans la base de e_i. Éventuellement Q peut être vu comme une application linéaire $\mathbb{K}^n \to \mathbb{K}^n$.

Dans la suite nous nommerons Q^{-1} la matrice inverse de Q. Inverse au sens des bêtes tableaux de nombres, sans interprétation en tant qu'application linéaire. De même pour Q^t qui est la transposée de Q.

4.4.1 Changement de base : vecteurs de base

Lemme 4.108.
Soit un espace vectoriel E sur $\mathbb{K}$ ainsi que deux bases $(e_i)_{i=1,\dots,n}$, $(f_\alpha)_{\alpha=1,\dots,n}$ de E liées par $f_\alpha = \sum_i Q_{i,\alpha} e_i$. Alors

$$e_i = \sum_\alpha Q^{-1}_{\alpha,i} f_\alpha. \tag{4.210}$$

Démonstration. Nous multiplions l'égalité $f_\alpha = \sum_i Q_{i,\alpha} e_i$ par le nombre[48] $Q^{-1}_{\alpha,j} \in \mathbb{K}$ et nous sommons sur α :

$$\sum_\alpha Q^{-1}_{\alpha,j} f_\alpha = \sum_{i\alpha} (A_{i,\alpha} Q^{-1}_{\alpha,j}) e_i = e_j. \tag{4.211}$$

□

4.4.2 Changement de base : coordonnées

Proposition 4.109.
Soit un espace vectoriel E sur $\mathbb{K}$. Soient deux bases $(e_i)_{i=1,\dots,n}$ et $(f_\alpha)_{\alpha=1,\dots,n}$ liées par $f_\alpha = \sum_i Q_{i,\alpha} e_i$. Nous considérons un même vecteur dans les deux bases : $\sum_i x_i e_i = \sum_\alpha y_\alpha f_\alpha$. Alors

(1) $y_\alpha = \sum_i Q^{-1}_{\alpha,i} x_i$

(2) $x_i = \sum_\alpha Q_{i,\alpha} y_\alpha$.

Démonstration. Soit un vecteur $x \in E$. Il peut être écrit dans les deux bases :

$$x = \sum_i x_i e_i = \sum_\alpha y_\alpha f_\alpha. \tag{4.212}$$

En remplaçant e_i par sa valeur (4.210) nous avons l'égalité

$$\sum_{i\alpha} x_i Q^{-1}_{\alpha,i} f_\alpha = \sum_\alpha y_\alpha f_\alpha. \tag{4.213}$$

Puisque les f_α sont linéairement indépendants, l'égalité des sommes donne l'égalité de chacun des termes :

$$y_\alpha = \sum_i x_i Q^{-1}_{\alpha,i}. \tag{4.214}$$

En identifiant $x \in E$ au vecteur dans $\mathbb{K}^n$ de ses coordonnées dans la base $\{e_i\}$ nous pouvons écrire

$$y_\alpha = (Q^{-1} x)_\alpha, \tag{4.215}$$

Le point (1) est prouvé.

En ce qui concerne le point (2), nous repartons encore de (4.212), mais nous y substituons la définition des f_α :

$$\sum_i x_i e_i = \sum_{\alpha i} y_\alpha Q_{i,\alpha} e_i. \tag{4.216}$$

Vous voulez des détails ? Allez, une étape de plus que le strict nécessaire : nous écrivons

$$\sum_i \big(x_i - \sum_\alpha y_\alpha Q_{i,\alpha}\big) e_i = 0. \tag{4.217}$$

Par linéaire indépendance des e_i, nous avons annulation de tous les coefficients, c'est-à-dire

$$x_i = \sum_\alpha Q_{i,\alpha} y_\alpha, \tag{4.218}$$

comme annoncé. □

48. Attention à la bonne interprétation de ce nombre : on fait bien référence à l'élément situé en (α, j) de la matrice Q^{-1}, et pas autre chose.

4.110.
Attention à l'ordre des indices dans la dernière égalité : la matrice Q vient avec les indices dans l'ordre i, α, tandis que la matrice Q^{-1} vient avec les indices dans l'ordre opposé : α, i. C'est pour cela qu'il est intéressant de noter avec des lettres latines les indices se rapportant à la première base, et avec des lettres grecques ceux se rapportant à la seconde base.

4.111.
Les formules de changement de coordonnées de la proposition 4.109 s'écrivent souvent de la façon suivante :

(1) $y_\alpha = (Q^{-1}x)_\alpha$
(2) $y = Q^{-1}x$.
(3) $x_i = (Qy)_i$
(4) $x = Qy$

Ces égalités reposent sur un petit paquet d'abus de notations qu'il convient de bien comprendre. Ici, x et y sont les éléments de $\mathbb{K}^n$ donnés par les composantes de x dans les bases $\{e_i\}$ et $\{f_\alpha\}$, et Q est vu comme une matrice, un opérateur linéaire sur $\mathbb{K}^n$. Autrement dit, le choix des bases permet d'identifier E avec $\mathbb{K}^n$ et la matrice Q avec l'application linéaire f_Q de la proposition 4.68.

4.4.3 Changement de base : matrice d'une application linéaire

Proposition 4.112.
Soit une application linéaire $t\colon E \to E$ de matrices A et B dans les bases $\{e_i\}$ et $\{f_\alpha\}$. Si les bases sont liées par

$$f_\alpha = \sum_i Q_{i,\alpha} e_i, \tag{4.219}$$

alors les matrices A et B sont liées par

$$B = Q^{-1}AQ. \tag{4.220}$$

Démonstration. L'hypothèse sur le fait que A et B sont les matrices de t signifie que pour tout $x \in E$,

$$t(x) = \sum_{ij} A_{j,i} x_i e_j = \sum_{\alpha\beta} B_{\alpha,\beta} y_\beta f_\alpha. \tag{4.221}$$

En remplaçant e_j par son expression (4.210) en termes des f_α et x_i par son expression $x_i = (Qy)_i$ (proposition 4.109), nous avons

$$(By)_\alpha = \sum_{ij\alpha} A_{j,i} (Qy)_i Q^{-1}_{\alpha,j} f_\alpha \tag{4.222a}$$
$$= \sum_{i\alpha} (Q^{-1}A)_{\alpha,i} (Qy)_i f_\alpha \tag{4.222b}$$
$$= \sum_\alpha (Q^{-1}AQy)_\alpha f_\alpha. \tag{4.222c}$$

Puisque les f_α forment une base, nous en déduisons $Q^{-1}AQy = By$. Et comme y est un élément quelconque de $\mathbb{K}^n$, nous en déduisons l'égalité de matrices

$$B = Q^{-1}AQ. \tag{4.223}$$

□

Il s'agit bien d'une égalité de matrices ou, à la limite, d'applications linéaires sur $\mathbb{K}^n$, et non d'une égalité d'application linéaire sur E.

4.5 Espaces de polynômes

Attention : les polynômes en-soi, font l'objet de la définition 1.303.

Pour chaque $k > 0$ donné, nous définissons

$$\mathcal{P}_{\mathbb{R}}^k = \{p : \mathbb{R} \to \mathbb{R} \,|\, p : x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k,\ a_i \in \mathbb{R},\ \forall i = 0, \ldots, k\}. \tag{4.224}$$

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 $\mathcal{P}_{\mathbb{R}}^k$ est donc un espace vectoriel muni des opérations héritées de $\mathcal{P}_{\mathbb{R}}$.

La base canonique de l'espace $\mathcal{P}_{\mathbb{R}}^k$ est donnée par les monômes $\mathcal{B} = \{x \mapsto x^j \,|\, j = 0, \ldots, k\}$. 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 applications linéaires de $\mathcal{P}_{\mathbb{R}}^k$ vers d'autres espaces vectoriels.

L'isomorphisme canonique $\phi : \mathcal{P}_{\mathbb{R}}^k \to \mathbb{R}^{k+1}$ Nous définissons ϕ par les relations suivantes

$$\phi(x^j) = e_{j+1}, \qquad \forall j \in \{0, \ldots, k\}.$$

Cela veut dire que pour tout p dans $\mathcal{P}_{\mathbb{R}}^k$, 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 4.113.
Soit $k = 5$ on a

$$\phi(-8 - 7x - 4x^2 + 4x^3 + 2x^5) = \begin{pmatrix} -8 \\ -7 \\ -4 \\ 4 \\ 0 \\ 2 \end{pmatrix}. \tag{4.225}$$

△

Cette application est clairement bijective et respecte les opérations d'espace vectoriel, donc c'est un isomorphisme d'espaces vectoriels. L'existence d'un isomorphisme entre $\mathcal{P}_{\mathbb{R}}^k$ 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 $d : \mathcal{P}_{\mathbb{R}}^k \to \mathcal{P}_{\mathbb{R}}^{k-1}$ L'application de dérivation d fait exactement ce qu'on attend d'elle

$$d(x^0) = d(1) = 0, \qquad d(x^j) = j x^{j-1}, \quad \forall j \in \{1, \ldots, k\}.$$

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 4.114.
Soit $k = 3$ on a

$$d(-8 - 12x + 4x^3) = -12(1) + 4(3x^2) = -12 + 12x^2. \tag{4.226}$$

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 $I : \mathcal{P}_{\mathbb{R}}^{k} \to \mathcal{P}_{\mathbb{R}}^{k+1}$ Nous pouvons définir une application qui est « à une constante près » l'application réciproque de la dérivation. Cette application est définie sur les éléments de base par

$$I(x^j) = \frac{x^{j+1}}{j+1}. \tag{4.227}$$

Bien entendu, la raison d'être et la motivation de cette définition apparaîtront lorsque nous développerons une théorie générale de l'intégration.

Exemple 4.115.
Soit $k = 4$ on a

$$I(6 + 2x + x^2 + x^4) = 6x + x^2 + \frac{x^3}{3} + \frac{x^5}{5}. \tag{4.228}$$

△

Remarque : étant donné que dans la définition de I nous avons décidé d'intégrer entre zéro et x, tous les polynômes dans $\mathcal{P}_{\mathbb{R}}^{k+1}$ qui sont l'image par I d'un polynôme de $\mathcal{P}_{\mathbb{R}}^{k}$ 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 $\mathcal{P}_{\mathbb{R}}^{k+1}$, 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 applications 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 11.3.2, nous verrons que toutes les normes ne sont pas équivalentes sur l'espace des polynômes.

4.6 Projection et orthogonalité

Proposition 4.116.
Si nous écrivons $\mathtt{proj}_Y$ *l'opération de projection sur la droite qui sous-tend* Y*, alors nous avons*

$$\| \mathtt{proj}_Y X \| = \frac{X \cdot Y}{\|Y\|}. \tag{4.229}$$

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

$$X = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} l \\ 0 \\ 0 \end{pmatrix} \tag{4.230}$$

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.

Un simple calcul montre que $X \cdot Y = xl + y \cdot 0 + 0 \cdot 0 = xl$. Par ailleurs, nous avons $\| \mathtt{proj}_Y X \| = x$. Par conséquent,

$$\| \mathtt{proj}_Y X \| = \frac{X \cdot Y}{l} = \frac{X \cdot Y}{\|Y\|}. \tag{4.231}$$

□

Corolaire 4.117.
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 4.116. □

4.7 Dualité

Proposition 4.118.
Si A est la matrice d'une application linéaire, alors le rang de cette application linéaire est égal au rang de A, c'est-à-dire à la taille de la plus grande matrice carrée de déterminant non nul contenue dans A.

Définition 4.119.
Soit E un espace vectoriel sur $\mathbb{K}$.

*Une **forme linéaire** sur E est une application linéaire de E sur son corps de base $\mathbb{K}$.*

*Le **dual algébrique** de E, noté E^*, est l'ensemble des formes linéaires sur E. Ainsi, $E^* = \mathrm{GL}(E, \mathbb{K})$.*

Nous verrons plus tard qu'en dimension infinie, les applications linéaires ne sont pas toujours continues. Nous définirons donc aussi un concept de dual topologique. Voir la proposition 11.61, la remarque 11.64 et la définition 25.1.

Lemme-Définition 4.120.
Si E est un espace vectoriel et si $\{e_i\}$ est une base de E, alors nous définissons

$$\begin{aligned} e^* \colon E &\to \mathbb{K} \\ e_j &\mapsto \delta_{i,j}, \end{aligned} \tag{4.232}$$

et sa prolongation par linéarité.

*Ces élements du dual E^* forment une base appelée **base duale**.*

Notons que si $v \in E$ est un vecteur, ça n'a aucun sens à priori de parler de v^*. Il s'agit bien de définir *toute* la base $\{e_i^*\}$ à partir de toute la base $\{e_i\}$.

Lemme 4.121 ([117])**.**
Soit un espace vectoriel E de dimension finie.

(1) Si α est une forme linéaire non nulle sur E, alors il existe $x \in E$ tel que $\alpha(x) = 1$.

(2) Si $x \neq 0$ dans E, alors il existe une forme linéaire α sur E telle que $\alpha(x) = 1$.

Démonstration. En deux parties.

(i) **Pour (1)** Puisque α est non nulle, nous pouvons considérer $v \in E$ tel que $\alpha(v) \neq 0$. Alors en posant $x = \alpha(v)^{-1}v$, nous avons le résultat.

(ii) **Pour (2)** Soit un vecteur non nul que nous écrivons sous la forme $x = \sum_{i=1}^n x_i e_i$ (pour une certaine base $\{e_i\}_{i=1,\dots,n}$ de E). Supposons que $x_k \neq 0$. Alors la forme

$$\begin{aligned} \alpha \colon E &\to \mathbb{K} \\ y &\mapsto y_k / x_k \end{aligned} \tag{4.233}$$

fait l'affaire.

□

Lemme 4.122.
Soit un espace vectoriel de dimension finie E sur le corps $\mathbb{K}$. Si $\{\alpha_1, \dots, \alpha_n\}$ est une base de E^*, alors l'application

$$\begin{aligned} \Phi \colon E &\to \mathbb{K}^n \\ x &\mapsto \big(\alpha_1(x), \dots, \alpha_n(x)\big) \end{aligned} \tag{4.234}$$

est un isomorphisme d'espaces vectoriels.

Démonstration. En deux parties.

(i) **Φ est injective** Soit $z \in \ker(\Phi)$. Nous avons $\alpha_i(z) = 0$ pour tout i. Si $z \neq 0$, alors le lemme 4.121 dit qu'il existe $\beta \in E^*$ tel que $\beta(z) \neq 0$.

Décomposons un tel β dans la base de $\{\alpha_i\}$:

$$\beta = \sum_{i=1}^{n} \beta_i \alpha_i. \tag{4.235}$$

Alors nous avons

$$0 \neq \beta(z) = \sum_{i=1}^{n} \beta_i \underbrace{\alpha_i(z)}_{=0} = 0. \tag{4.236}$$

Contradiction. Donc $\ker(\Phi) = \{0\}$ et Φ est injective.

(ii) **Φ est surjective** Les espaces vectoriels E, E^* et $\mathbb{K}^n$ ont tout trois, une dimension n. Donc Φ est une application linéaire injective entre deux espaces de même dimension. Elle est donc surjective par le corolaire 4.46.

□

Proposition-Définition 4.123 ([117]).
*Soit un espace vectoriel E de dimension finie sur le corps $\mathbb{K}$. Toute base du dual E^* est duale d'une unique base de E. Cette base est dite **préduale**.*

Démonstration. Nous considérons une base $\mathcal{F} = \{\alpha_1, \dots, \alpha_n\}$ de E^*. Nous devons prouver qu'il existe une unique base de E dont la base duale est $\mathcal{F}$.

(i) **Existence** Le lemme 4.122 nous indique que l'application

$$\begin{aligned} \Phi \colon E &\to \mathbb{K}^n \\ x &\mapsto \big(\alpha_1(x), \dots, \alpha_n(x)\big) \end{aligned} \tag{4.237}$$

est un isomorphisme d'espaces vectoriels.

Soit la base canonique de $\mathbb{K}^n$: $(\epsilon_1, \dots, \epsilon_n)$. Puisque Φ est un isomorphisme, $\big(\Phi^{-1}(\epsilon_i)\big)_{i=1,\dots,n}$ est une base de E. Nous allons montrer qu'elle est préduale de (α_i). Nous posons $e_i = \Phi^{-1}(\epsilon_i)$ et nous calculons :

$$\begin{aligned} \alpha_i(e_j) &= \alpha_i\big(\Phi^{-1}(\epsilon_j)\big) && (4.238a) \\ &= \Phi\big(\Phi^{-1}(\epsilon_j)\big)_i && (4.238b) \\ &= (\epsilon_i)_j && (4.238c) \\ &= \delta_{i,j} && (4.238d) \end{aligned}$$

Justifications :

— Pour 4.238b, nous remarquons que $\alpha_i(x) = \Phi(x)_i$.

— Pour 4.238d, nous utilisons le fait que les ϵ_j forment la base canonique de $\mathbb{K}^n$.

(ii) **Unicité** Soit une base préduale (e_i) de (α_i). Nous avons, par définition, que $\alpha_i(e_j) = \delta_{i,j}$. Donc

$$\big(\alpha_1(e_j), \dots, \alpha_n(e_j)\big) = \epsilon_j. \tag{4.239}$$

Nous appliquons Φ^{-1} à cette dernière équation pour obtenir $e_j = \Phi^{-1}(\epsilon_j)$. Donc les e_j sont déterminés de façon unique à partir des α_i.

□

4.7.1 Orthogonal

Définition 4.124.
*Soit E, un espace vectoriel, et F un sous-espace de E. L'**orthogonal** de F est la partie $F^\perp \subset E^*$ donnée par*

$$F^\perp = \{\alpha \in E^* \text{ tel que } \forall x \in F, \alpha(x) = 0\}. \tag{4.240}$$

Cette définition d'orthogonal via le dual n'est pas du pur snobisme. En effet, la définition « usuelle » qui ne parle pas de dual,

$$F^\perp = \{y \in E \text{ tel que } \forall x \in F, y \cdot x = 0\}, \tag{4.241}$$

demande la donnée d'un produit scalaire. Évidemment dans le cas de $\mathbb{R}^n$ muni 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.

La définition 4.124, au contraire, est intrinsèque : elle ne dépend que de la structure d'espace vectoriel.

Si $B \subset E^*$, on note B^o son orthogonal :

$$B^o = \{x \in E \text{ tel que } \omega(x) = 0, \forall \omega \in B\}. \tag{4.242}$$

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.

Proposition 4.125.
Soient un espace vectoriel E et F, un sous-espace vectoriel de E. Nous avons

$$\dim F + \dim F^\perp = \dim E. \tag{4.243}$$

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 4.12. Soit $\{e_1^*, \dots, e_n^*\}$ la base duale. Alors nous prouvons que $\{e_{p+1}^*, \dots, e_n^*\}$ est une base de $F^\perp$.

D'abord, 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, $e_{p+k}^* \in F^\perp$.

Ensuite, en tant que partie d'une base de F^*, c'est une partie libre. Il reste à montrer que c'est générateur.

Enfin $F^\perp \subset \operatorname{Span}\{e_k^*, k \in \{p+1, \dots, n\}\}$ 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 $\omega = \sum_{k=p+1}^n \omega_k e_k^*$. □

La proposition 9.172 donnera une version plus terre à terre de la proposition 4.125 en disant que si nous avons un produit scalaire, alors $E = F \oplus F^\perp$ où $F^\perp$ est cette fois défini comme l'orthogonal pour le produit scalaire.

4.8 Représentation de groupe

Définition 4.126 (Représentation)**.**
*Soit un groupe G. Une **représentation** de G est un couple (E, ρ) où E est un espace vectoriel et ρ est une application $\rho\colon G \to \mathrm{GL}(E)$ vérifiant*

$$\rho(g) \circ \rho(h) = \rho(gh). \tag{4.244}$$

pour tout $g, h \in G$.

Définition 4.127.
*Une représentation[49] $\rho\colon G \to \mathrm{GL}(E)$ est **fidèle** si elle est injective*

*La dimension de E est le **degré** de la représentation (E, ρ).*

49. Définition 4.126.

Le fait que la représentation $\rho\colon G \to \mathrm{GL}(E)$ soit fidèle ne dit pas que chacun des $\rho(g)$ est injectif.

Proposition 4.128.
Soit un corps $\mathbb{K}$. Si G est un groupe dans $\mathbb{M}(n,\mathbb{K})$ (c'est-à-dire un groupe de matrices $n \times n$ à coefficients dans $\mathbb{K}$), alors l'application

$$\begin{aligned} \rho\colon G &\to \mathrm{GL}(\mathbb{K}^n) \\ A &\mapsto f_A \end{aligned} \tag{4.245}$$

où f_A est l'application linéaire associée à A, est une représentation de G.

Démonstration. La représentation dont nous parons n'est autre que l'application ψ de la définition 4.65, dont nous connaissons beaucoup de propriétés. La proposition 4.71 dit, entre autres, que $\psi(AB) = \psi(A)\psi(B)$, c'est à dire que $\rho(AB) = \rho(A)\rho(B)$, et donc que ρ est une représentation. □

4.9 Somme directe d'espaces vectoriels

Si V et W sont des espaces vectoriels, ce que nous notons $V \oplus W$ n'est rien d'autre que l'espace vectoriel de l'ensemble $V \times W$.

Proposition-Définition 4.129 ([118, 119])**.**
Si V et W sont des espaces vectoriels sur le même corps $\mathbb{K}$, alors les définitions

(1) $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$

(2) $\lambda(v, w) = (\lambda v, \lambda w)$

donnent une structure d'espace vectoriel sur $V \times W$.

Cet espace sera noté $V \oplus W$ et est appelé ***somme directe*** *de V et W.*

Définition 4.130 (Sous-espaces en somme directe[120])**.**
Soient un espace vectoriel E sur $\mathbb{K}$ ainsi que des sous-espaces vectoriels $\{F_i\}_{i\in I}$ (I est un ensemble fini ou infini). Nous disons que les F_i sont ***en somme directe*** *si pour tout élément $u \in \sum_{i\in I} F_i$, il existe un unique ensemble $\{u_i\}_{i\in I}$ de vecteurs tel que*

(1) $u = \sum_{i\in I} u_i$

(2) $u_i \in F_i$ *pout tout i,*

(3) $\{j \in I$ *tel que* $u_j \neq 0\}$ *est fini.*

Si l'espace vectoriel E est un espace vectoriel topologique, nous avons la définition 7.151 qui donne des conditions de compatibilité entre les topologies.

Lemme 4.131.
Soit une espace vectoriel E et deux sous-espaces F_1, F_2 en somme directe : $E = F_1 \oplus F_2$. Alors l'application

$$\begin{aligned} \psi\colon F_1 \times F_2 &\to E \\ (x, y) &\mapsto x + y \end{aligned} \tag{4.246}$$

est une bijection.

Démonstration. L'application est injective parce que si $\psi(x_1, x_2) = \psi(y_1, y_2)$, alors $x_1 + x_2 = y_1 + y_2$. Nommons u ce vecteur. Par unicité de l'ensemble $\{u_1, u_2\}$ tel que $u = u_1 + u_2$, nous avons automatiquement $\{x_1, x_2\} = \{y_1, y_2\}$ et donc $x_1 = y_1$ et $x_2 = y_2$.

En ce qui concerne la surjectivité, si $u \in E$, il existe un ensemble $\{u_1, u_2\}$ avec $u_i \in F_i$ tel que $u = u_1 + u_2 = \psi(u_1, u_2)$. □

Lemme 4.132 ([1, 120])**.**
Soient un espace vectoriel E ainsi que des sous-espaces vectoriels F_i. Nous avons équivalence entre les assertions suivantes.

(1) Les F_i sont en somme directe[50].

(2) Si $\sum_{i\in I} u_i = 0$ avec $u_i \in F_i$ et si $\{j \in I$ tel que $u_j \neq 0\}$ est fini, alors tous les u_i sont nuls.

(3) Chaque espace F_k est en somme directe avec la somme des précédents, c'est à dire que pour tout k,

$$\left(\sum_{i=1}^{k-1} F_i\right) \cap F_k = \{0\}. \tag{4.247}$$

(4) Pour tout k,

$$F_k \cap \left(\sum_{i\neq k} F_i\right) = \{0\}. \tag{4.248}$$

Proposition 4.133 ([121]).
Soient E un espace vectoriel de dimension finie, et deux sous-espaces F_1 et F_2 satisfaisant

(1) $F_1 \cap F_2 = \{0\}$,

(2) $\dim(F_1) + \dim(F_2) \geqslant \dim(E)$.

Alors $E = F_1 \oplus F_2$.

Démonstration. Soient une base $\{e_i\}_{i\in I}$ de F_1 et $\{f_\alpha\}$ de F_2. Nous commençons par prouver que la partie $B = \{e_i\} \cup \{f_\alpha\}$ est libre.

Supposons en effet, avoir des coefficients a_i et b_α tels que

$$\sum_i a_i e_i + \sum_\alpha b_\alpha f_\alpha = 0. \tag{4.249}$$

Cela implique que $\sum_i a_i e_i = -\sum_\alpha b_\alpha f_\alpha$. Or $\sum_i a_i e_i \in F_1$ et $-\sum_\alpha b_\alpha f_\alpha \in F_2$. Donc les éléments $\sum_i a_i e_i$ et $\sum_\alpha b_\alpha f_\alpha$ sont dans $F_1 \cap F_2 = \{0\}$. Nous avons alors les égalités

$$\sum_i a_i e_i = 0 \tag{4.250}$$

et

$$\sum_\alpha b_\alpha f_\alpha = 0. \tag{4.251}$$

La première implique $a_i = 0$ pour tout i et la seconde implique $b_\alpha = 0$ pour tout α.

Donc B est une partie libre de E contenant $\dim(F_1) + \dim(F_2) \geqslant \dim(E)$ éléments. La proposition 4.17(1) nous indique alors qu'en réalité $\dim(F_1) + \dim(F_2) = \dim(E)$. Comme B est une partie libre contenant $\dim(E)$ éléments, c'est une base par la proposition 4.17(2). □

4.9.1 Structure réelle

Proposition-Définition 4.134 ([122, 119, 1]).
Soit un espace vectoriel E sur $\mathbb{C}$. Il existe une application $\sigma\colon E \to E$ telle que

(1) $\sigma^2 = \sigma$ (involution)

(2) Pour tout $\alpha, \beta \in \mathbb{R}$, $f(\alpha x + \beta y) = \bar{\alpha}\sigma(x) + \bar{\beta}\sigma(y)$.

*Une telle application est une **structure réelle** sur E.*

Démonstration. La proposition 4.22 nous permet de considérer une base $\{e_i\}_{i\in I}$ de E. Alors, nous définissons

$$\sigma\big(\sum_{i\in I} \lambda_i e_i\big) = \sum_{i\in I} \bar{\lambda}_i e_i. \tag{4.252}$$

Notez que la somme est toujours finie. □

50. Définition 4.130.

Proposition 4.135.
Soit un espace vectoriel E sur $\mathbb{C}$ et une structure réelle[51] *σ sur E. Nous posons*

$$E_{\mathbb{R}} = \{v \in E \text{ tel que } \sigma(v) = v\}. \tag{4.253}$$

Alors

(1) La partie $E_{\mathbb{R}}$ est un espace vectoriel réel.

(2) Nous avons la décomposition en somme directe[52]

$$E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}. \tag{4.254}$$

Démonstration. En plusieurs parties.

(i) **Espace vectoriel réel** Si $v, w \in E_{\mathbb{R}}$, alors

$$\sigma(v + w) = \sigma(v) + \sigma(w) = v + w, \tag{4.255}$$

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

$$\sigma(\lambda x) = \lambda\sigma(x) = \lambda x. \tag{4.256}$$

Donc $E_{\mathbb{R}}$ est un espace vectoriel réel.

(ii) **Première somme directe** Nous définissons

$$E^+ = \{v \in E \text{ tel que } \sigma(v) = v\}, \tag{4.257a}$$
$$E^- = \{v \in E \text{ tel que } \sigma(v) = -v\} \tag{4.257b}$$

Nous prouvons que

$$\begin{aligned} \psi\colon E^+ \times E^- &\to E \\ (a, b) &\mapsto a + b \end{aligned} \tag{4.258}$$

est un isomorphisme d'espaces vectoriels.

Puisque ψ est linéaire, il suffit de prouver qu'elle est bijective.

(i) **Surjectif** Si $v \in E$, alors en posant $v_+ = \frac{1}{2}\big(v + \sigma(v)\big)$, $v_- = \frac{1}{2}\big(v - \sigma(v)\big)$, nous avons

$$v = v_+ + v_-, \quad v_+ \in E^+, \quad v_- \in E^-, \tag{4.259}$$

et donc $v = \psi(v_+, v_-)$.

(ii) **Injectif** Supposons $\psi(a, b) = \psi(\alpha, \beta)$. Alors $a + b = \alpha + \beta$ et donc $a - \alpha = \beta - b$. Comme $a - \alpha \in E^+$ et $\beta - b \in E^-$, nous savons que $a - \alpha = \beta - b \in E^+ \cap E^-$. Étant donné que $E^+ \cap E^- = \{0\}$, nous avons $a - \alpha = \beta - b = 0$.

Nous avons donc la somme directe $E = E^+ \oplus E^-$.

(iii) **Conclusion** Par définition, $E^+ = E_{\mathbb{R}}$. Il nous reste à voir que $E^- = iE^+$. Nous prouvons les inclusions dans les deux sens.

(i) $E^- \subset iE^+$ Soit $v \in E^-$. Nous avons $iv \in E^+$; en effet

$$\sigma(iv) = \bar{i}\sigma(v) = -i\sigma(v) = iv. \tag{4.260}$$

Donc $iv \in E^+$ pour $v \in E^-$.

(ii) $iE^+ \subset E^-$ Soit $v \in E^+$, et voyons que $iv \in E^-$. En effet,

$$\sigma(iv) = -i\sigma(v) = -iv. \tag{4.261}$$

□

51. Définition 4.134.
52. Définition 4.129.

4.136.
Lorsque nous avons une structure réelle σ sur un espace vectoriel complexe E, nous écrivons $E = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ sans préciser dans la notation « $E_{\mathbb{R}}$ » que cet ensemble dépend du choix de σ. En particulier si F est un sous-espace vectoriel de E, nous utiliserons la notation $F_{\mathbb{R}}$ relativement à la même involution que celle utilisée pour E.

Lemme 4.137.
Soit un espace vectoriel complexe E muni d'une structure réelle σ. Si F est un sous-espace de E alors $F_{\mathbb{R}} = E_{\mathbb{R}} \cap F$.

Démonstration. Par définition,

$$F_{\mathbb{R}} = \{v \in F \text{ tel que } \sigma(v) = v\}. \tag{4.262}$$

(i) $\underline{F_{\mathbb{R}} \subset F}$ C'est dans la définition de $F_{\mathbb{R}}$ (sous-ensemble de F).

(ii) $\underline{F_{\mathbb{R}} \subset E_{\mathbb{R}}}$ Si $v \in F_{\mathbb{R}}$, alors $\sigma(v) = v$. Mais cette égalité est précisément celle qui permet d'être dans $E_{\mathbb{R}}$.

□

Vous remarquerez que ce lemme ne fonctionne que parce que nous avons choisi la même structure réelle sur F que sur E.

Chapitre 5

Classification de certains groupes

5.1 Théorèmes de Sylow

Lemme 5.1.
Soient H et K des sous-groupes finis de G. Alors

$$\mathrm{Card}(HK) = \frac{|H| \cdot |K|}{|H \cap K|}. \tag{5.1}$$

Attention : dans ce lemme, l'ensemble HK n'est pas spécialement un groupe. Ce serait le cas si H normalisait K, c'est-à-dire si nous avions $hkh^{-1} \in K$, $\forall (h,k) \in H \times K$.

Théorème 5.2 (Théorème de Cauchy[123])**.**
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 5.3 (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

$$\begin{aligned} \varphi\colon G &\to S_n \\ \varphi(x)g &\mapsto xg \end{aligned} \tag{5.2}$$

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 5.4.
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 5.5.
En mettant bout à bout les lemmes 5.3 et 5.4, 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 5.6.
*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 5.7.
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 5.8.
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

$$\begin{aligned}
\mathrm{Card}\big(\mathrm{GL}(n, \mathbb{F}_p)\big) &= (p^n - 1)(p^n - p)\dots(p^n - p^{n-1}) && (5.3a)\\
&= p \cdot p^2 \cdots p^{n-1}(p^n - 1)(p^{n-1} - 1)\dots(p - 1) && (5.3b)\\
&= p^{\frac{n(n-1)}{2}} m && (5.3c)
\end{aligned}$$

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

$$|T| = p^{\frac{n(n-1)}{2}}, \tag{5.4}$$

et T est un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$. □

Proposition 5.9.
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 (5.2), 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 2.13 de Lagrange). Le nombre r est alors une puissance de p. □

Lemme 5.10.
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

$$gSg^{-1} \cap H \tag{5.5}$$

soit un p-Sylow de H.

1. Définition 4.43.

Démonstration. Nous considérons l'ensemble G/S sur lequel H agit. Si $a \in G$, le stabilisateur de $[a]$ dans G/S est

$$\begin{aligned} \mathrm{Fix}\big([a]\big) &= \{h \in H \text{ tel que } [ha] = [a]\} && (5.6a)\\ &= \{h \in H \text{ tel que } a^{-1}ha \in S\} && (5.6b)\\ &= aSa^{-1} \cap H. && (5.6c) \end{aligned}$$

Nous cherchons $a \in G$ tel que l'entier

$$\frac{\mathrm{Card}(H)}{\mathrm{Card}\left(aSa^{-1} \cap H\right)} \tag{5.7}$$

soit premier avec p. En effet, dans ce cas le groupe $\mathrm{Fix}([a])$ est un p-Sylow de H parce que $|H : aSa^{-1} \cap H|$ ne divise pas p. La formule des orbites (équation (2.67)) nous dit que

$$\frac{|H|}{|aSa^{-1} \cap H|} = \mathrm{Card}\left(\mathcal{O}_{[a]}\right). \tag{5.8}$$

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

$$p \mid \mathrm{Card}(G/S) = \frac{|G|}{|S|} \tag{5.9}$$

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 (5.7) soit vérifiée. □

Théorème 5.11 (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. En plusieurs points.

(1) Nous savons de la remarque 5.5 que G est un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_p)$ et que ce dernier a un p-Sylow par la proposition 5.8. Par conséquent G possède un p-Sylow par le lemme 5.10.

(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 5.10 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,

$$H = aSa^{-1} \cap H \tag{5.10}$$

et $H \subset aSa^{-1}$, ce qui signifie que H est inclus dans 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 vient du fait 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

$$E_S = \{T \in E \text{ tel que } s \cdot T = T, \forall s \in S\}. \tag{5.11}$$

2. Définition 5.6.

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

$$|E| \equiv |E_S| \mod p. \tag{5.12}$$

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

$$sTs^{-1} = T \tag{5.13}$$

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

$$g = s_1 t_1 \cdots s_r t_r \tag{5.14}$$

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

$$gtg^{-1} = s_1 t_1 \dots s_r t_r t t_r^{-1} s_r^{-1} \dots t_1^{-1} s_1^{-1} \in T. \tag{5.15}$$

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,

$$S = aTa^{-1} = T. \tag{5.16}$$

Ceci achève la démonstration des théorèmes de Sylow.

□

Proposition 5.12.
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

$$(gsg^{-1})^q = gs^q g^{-1} = e. \tag{5.17}$$

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 5.13 ([124])**.**
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 5.14 ([124])**.**
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 parcourt les sous-groupes d'ordre p dans G. Chacun de ces ensembles possède $p-1$ éléments et le lemme 5.13 nous assure qu'ils sont disjoints. Par conséquent nous avons $n(p-1)$ éléments d'ordre p dans G. □

Corolaire 5.15.
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 5.14 nous dit alors que le nombre d'éléments d'ordre p dans G est $p-1$. Donc tout élément est générateur. □

5.2 Groupe monogène

Groupe monogène : définition 1.269 ; groupe cyclique : définition 1.270.

Le théorème suivant donne quelques informations à propos des groupes monogènes. Il impliquera dans le corolaire 19.16 qu'un groupe monogène d'ordre n possède $\varphi(n)$ générateurs où φ est la fonction indicatrice d'Euler définie en 19.11.

Théorème 5.16.
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.

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

$$\begin{aligned} f\colon \mathbb{Z} &\to G \\ p &\mapsto a^p. \end{aligned} \tag{5.18}$$

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 2.6) nous indique que

$$\mathbb{Z}/\ker f = \mathbb{Z}/n\mathbb{Z} = \operatorname{Image} f = G. \tag{5.19}$$

□

Le lemme suivant donne une démonstration alternative, avec une construction plus explicite de l'isomorphisme.

Lemme 5.17 ([1])**.**
À propos de groupes monogènes[3] *et cycliques.*

(1) Soit un groupe cyclique G de cardinal n dont g est un générateur. Alors il existe un isomorphisme

$$\phi\colon G \to (\mathbb{Z}/n\mathbb{Z}, +) \tag{5.20}$$

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

$$\phi\colon G \to (\mathbb{Z}, +) \tag{5.21}$$

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.

3. Définition 1.269.

Démonstration. Commençons par enfoncer une porte ouverte : comme le groupe est monogène, l'ordre du groupe est égal à l'ordre de son générateur. Nous séparons les cas selon que l'ordre soit fini ou non.

(i) **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 :

$$\phi(g^m) = [m]_n. \tag{5.22}$$

Cela est une bonne définition parce qu'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 un isomorphisme entre G et $\mathbb{Z}/n\mathbb{Z}$.

(i) **Morphisme** Pour l'identité, si $x = e$ alors $m = 0$ et $\phi(e) = [0]_n$. Et si $x = g^k$, $y = g^l$ alors $\phi(xy) = \phi(g^{k+l}) = [k + l]_n = [k]_n + [l]_n = \phi(x) + \phi(y)$.

(ii) **Injectif** Supposons $\phi(g^k) = \phi(g^l)$ avec $k \geqslant l$. Nous avons $h^k = h^l$, donc $h^{k-l} = e$, ce qui donne $k - l \in [0]_n$ ou encore $[k]_n = [l]_n$. En particulier $g^k = g^l$.

(iii) **Surjectif** La classe $[k]_n$ est l'image de g^k.

(ii) **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.

□

5.3 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 5.18 ([125]).
Pour chaque $x \in (\mathbb{Z}/n\mathbb{Z})^*$ nous considérons l'application

$$\begin{aligned} \sigma_x\colon \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ y &\mapsto xy. \end{aligned} \tag{5.23}$$

L'application

$$\sigma\colon \big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \big) \to \mathrm{Aut}\,\big(\mathbb{Z}/n\mathbb{Z}, +\big) \tag{5.24}$$

ainsi définie est un isomorphisme de groupes.

L'énoncé de ce théorème s'écrit souvent rapidement par

$$\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z}) = (\mathbb{Z}/n\mathbb{Z})^*, \tag{5.25}$$

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

$$f([r]) = f(r[1]) = rf([1]) = [r]f([1]). \tag{5.26}$$

En particulier, puisque 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]

$$\begin{aligned} \sigma\colon \operatorname{Aut}((\mathbb{Z}/n\mathbb{Z},+)) &\to \big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \big) \\ f &\mapsto f([1]) \end{aligned} \tag{5.27}$$

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 5.17. 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

$$f\big(g([1])\big) = f\big(g([1])[1]\big) = g([1])f([1]) = \sigma(f)\sigma(g). \tag{5.28a}$$

□

Ce dernier résultat s'étend aux groupes cycliques.

Proposition 5.19.
Si G est un groupe cyclique[5] *d'ordre n, alors*

$$\operatorname{Aut}(G) = \big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \big). \tag{5.29}$$

Démonstration. Vu que G est cyclique, le lemme 5.17 nous dit que G est isomorphe à $(\mathbb{Z}/n\mathbb{Z}),+$. Maintenant le théorème 5.18 nous indique que

Les égalités suivantes sont en réalité des isomorphismes de groupes :

$$\operatorname{Aut}(G) = \operatorname{Aut}\Big(\mathbb{Z}/n\mathbb{Z},+\Big) \tag{5.30a}$$

$$= \Big((\mathbb{Z}/n\mathbb{Z})^*, \ \cdot\ \Big) \tag{5.30b}$$

Jusitifications.

— Pour (5.30a). Le lemme 5.17 nous dit que G est isomorphe à $(\mathbb{Z}/n\mathbb{Z},+)$, et le lemme 1.37 dit que des groupes isomorphes ont des groupes d'isomorphismes isomorphes.

— Pour (5.30b). C'est le théorème 5.18.

□

Corolaire 5.20.
Si p divise $q-1$ alors $\operatorname{Aut}(\mathbb{F}_q)$ possède un unique sous-groupe d'ordre p.

Démonstration. Si a est un générateur de $\mathbb{F}_q^*$ alors le groupe

$$\operatorname{gr}\Big(a^{\frac{q-1}{p}}\Big) \tag{5.31}$$

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 3.30 nous enseigne que le groupe donné en (5.31) 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.

5. Définition 1.270.

5.4 Groupes abéliens finis

Source : [74].

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. C'est le théorème suivant.

Théorème 5.21 (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,

$$m = p^\alpha m' \tag{5.32a}$$

$$q = p^\beta q' \tag{5.32b}$$

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 5.22.
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 :

$$\begin{aligned} \tilde{\varphi}\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(x) \\ (\bar{k}, h) &\mapsto x^{kl}\varphi(h) \end{aligned} \tag{5.33}$$

où l est encore à déterminer, et

$$\begin{aligned} p\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(y, H) \\ (\bar{k}, h) &\mapsto y^k h. \end{aligned} \tag{5.34}$$

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

$$\begin{tikzcd} \ker(p) \arrow[r, hook] & \mathbb{Z}/b\mathbb{Z} \times H \arrow[r, "p"] \arrow[d, "\tilde{\varphi}"'] & \mathrm{gr}(y, H) \arrow[dl, "\hat{\varphi}"] \\ & \mathrm{gr}(x) & \end{tikzcd} \tag{5.35}$$

que l'on voudra être commutatif. Puisque p est surjective, les théorèmes d'isomorphismes nous disent que

$$\mathrm{gr}(y, H) \simeq \frac{\mathbb{Z}/b\mathbb{Z} \times H}{\ker p}. \tag{5.36}$$

Si $[\bar{k}, h]$ est la classe de $(\bar{k}, h)$ modulo $\ker(p)$ alors nous voudrions définir $\hat{\varphi}$ par

$$\hat{\varphi}\big([\bar{k}, h]\big) = \tilde{\varphi}(\bar{k}, h). \tag{5.37}$$

Pour que cela soit bien défini, il faut que si $(\bar{r}, z) \in \ker p$, alors,

$$\hat{\varphi}\big([\bar{k}\bar{r}, hz]\big) = \hat{\varphi}\big([\bar{k}, h]\big), \tag{5.38}$$

c'est-à-dire que $\tilde{\varphi}(\bar{r}, z) = e$. Du coup la définition (5.37) n'est bonne que si et seulement si

$$\ker(p) \subset \ker(\tilde{\varphi}). \tag{5.39}$$

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

$$\ker(p) = \{(-m\beta, y^{m\beta}) \text{ tel que } m \in \mathbb{Z}\}. \tag{5.40}$$

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

$$\tilde{\varphi}(\beta, y^{-\beta}) = x^{\beta l}\varphi(y^{-\beta}) = x^{\beta l + \alpha}. \tag{5.41}$$

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,

$$e = \varphi(y^b) = \varphi(y^{-\beta b/\beta}) = \varphi(y^{-\beta})^{b/\beta} = x^{b\beta/\alpha}. \tag{5.42}$$

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

$$\begin{aligned} \psi\colon G &\to \operatorname{gr}(x) \oplus \ker(\varphi) \\ g &\mapsto \big(\varphi(g), g\varphi(g)^{-1}\big) \end{aligned} \tag{5.43}$$

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,

$$\varphi\big(g\varphi(g)^{-1}\big) = \varphi(g)\varphi(g)^{-1} = e. \tag{5.44}$$

L'application ψ est un morphisme parce que, en utilisant le fait que G est abélien,

$$\begin{aligned} \psi(g_1g_2) &= \big(\varphi(g_1g_2), g_1g_2\varphi(g_1g_2)^{-1}\big) && (5.45a) \\ &= \big(\varphi(g_1)\varphi(g_2), g_1\varphi(g_1)^{-1}g_2\varphi(g_2)^{-1}\big) && (5.45b) \\ &= \psi(g_1)\psi(g_2). && (5.45c) \end{aligned}$$

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 2.6) appliqué à φ nous avons

$$|G| = |\operatorname{gr}(x)| \cdot |\ker(\varphi)|. \tag{5.46}$$

□

Théorème 5.23.
Tout groupe abélien fini (non trivial) se décompose en

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/d_r\mathbb{Z} \tag{5.47}$$

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

$$H_1 = \mathrm{gr}(x_1) = \mathbb{F}_{n_1}. \tag{5.48}$$

D'après la proposition 5.22(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

$$G = \mathbb{F}_{d_1} \oplus \ldots \oplus \mathbb{F}_{d_r} = \mathbb{F}_{s_1} \oplus \ldots \oplus \mathbb{F}_{s_q}. \tag{5.49}$$

L'exposant de G est d_r et s_q. Donc $d_r = s_q$. Les complémentaires étant égaux nous avons

$$\mathbb{F}_{d_1} \oplus \ldots \oplus \mathbb{F}_{d_{r-1}} = \mathbb{F}_{s_1} \oplus \ldots \oplus \mathbb{F}_{s_{q-1}}. \tag{5.50}$$

En continuant nous trouvons $r = q$ et $d_i = s_i$. □

5.5 Groupes d'ordre pq

Lemme 5.24.
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[6].

Démonstration. Soit n_q le nombre de q-Sylow ; par le théorème de Sylow 5.11(1) le groupe G possède des q-Sylow et par 5.11(4),

$$n_q \in [1]_q. \tag{5.51}$$

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 5.11(3) parce que le conjugué de H est encore un q-Sylow alors que H est l'unique q-Sylow.

Vu que

$$1 < p = |H| < pq = |G|, \tag{5.52}$$

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 de lire le théorème suivant, n'oubliez pas de lire la définition d'un produit semi-direct 2.47.

Théorème 5.25 ([126]).
Soient deux nombres premiers distincts[7] *p et q avec $q > p$.*

6. Pas de sous-groupes normaux non triviaux, 1.169.
7. Le cas $p = q$ sera traité par la proposition 5.28.

(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*

$$G = \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} \tag{5.53}$$

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[8].

Démonstration. Division de la preuve en plusieurs parties.

(i) **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 5.11). 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 2.13) $|H \cap K|$ 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

$$\begin{aligned} \psi\colon H \times K &\to HK \\ (h,k) &\mapsto hk \end{aligned} \tag{5.54}$$

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

$$h^{-1}h' = (k'k^{-1})^{-1} \in H \cap K = \{e\}. \tag{5.55}$$

Par conséquent $|pq| = |H \times K| = |HK|$, et $HK = G$. Le corolaire 2.49 nous indique que

$$G = H \times_\varphi K \tag{5.56}$$

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.

(ii) **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 (5.56) est en réalité un produit direct (φ est triviale) et nous avons

$$G = \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/pq\mathbb{Z}. \tag{5.57}$$

(iii) **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 2.6, nous avons

$$|\varphi(\mathbb{Z}/p\mathbb{Z})| = \frac{|\mathbb{Z}/p\mathbb{Z}|}{|\ker\varphi|}, \tag{5.58}$$

8. 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 corolaire 5.20 nous indique que $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ possède un unique sous-groupe d'ordre p que nous notons Γ ; c'est-à-dire que $\Gamma = \mathrm{Image}(\varphi)$. Vu que $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ est un morphisme, Γ est généré par $\varphi(\bar{1})$ qui est alors un élément d'ordre p, comme annoncé.

(iv) **Unicité** Nous nous attaquons maintenant à l'unicité. Soient φ et φ' deux morphismes non triviaux $\mathbb{Z}/p\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$. Étant donné que $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ ne possède qu'un seul sous-groupe d'ordre p, nous savons que $\mathrm{Image}(\varphi) = \mathrm{Image}(\varphi') = \Gamma$. Nous pouvons donc parler de φ'^{-1} en tant qu'application de $\mathbb{Z}/p\mathbb{Z}$ dans Γ. Nous montrons que

$$\begin{aligned} f\colon \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} &\to \mathbb{Z}/q\mathbb{Z} \times_{\varphi'} \mathbb{Z}/p\mathbb{Z} \\ (h,k) &\mapsto (h, \alpha(k)) \end{aligned} \tag{5.59}$$

où $\alpha = \varphi'^{-1} \circ \varphi$ est un isomorphisme de groupes. Le calcul est immédiat :

$$\begin{aligned} f(h_1,k_1)f(h_2mk_2) &= \big(h_1, \alpha(k_1)\big)(h_2, \alpha(k_2)) && (5.60a)\\ &= \big(h_1\varphi'(\alpha(k_1))h_2m\alpha(k_1k_2)\big) && (5.60b)\\ &= f\big(h_1\varphi(k_1)h_2, k_1k_2\big) && (5.60c)\\ &= f\big((h_1,k_1),(h_2,k_2)\big). && (5.60d) \end{aligned}$$

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 5.26 ([35]).
Soit G un groupe fini d'ordre pq où p et q sont deux nombres premiers distincts vérifiant

$$\begin{cases} p \neq 1 \mod q & (5.61a)\\ q \neq 1 \mod p. & (5.61b) \end{cases}$$

Alors G est cyclique, abélien et

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{5.62}$$

Démonstration. Soient n_p et n_q les nombres de p-Sylow et q-Sylow. Par le théorème de Sylow 5.11, 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 exposées 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

$$\begin{aligned} (xyx^{-1})y^{-1} &\in T && (5.63a)\\ x(yx^{-1})y^{-1}) &\in S, && (5.63b) \end{aligned}$$

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 $S \cap T$ 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

$$|ST| = |S||T| = pq = |G|. \tag{5.64}$$

Par conséquent $G = ST$. En nous rappelant que $S \cap T = \{e\}$ et que S et T sont normaux, le lemme 1.175 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

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{5.65}$$

□

Théorème 5.27 (Théorème de Burnside[74])**.**
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 :

$$\mathcal{Z}_G = \{z \in G \text{ tel que } \sigma_x(z) = z, \forall x \in G\} = \text{Fix}_G(G). \tag{5.66}$$

Nous utilisons l'équation aux classes (2.38) 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 5.28.
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 5.2 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 5.27, 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. □

5.6 Groupe symétrique, groupe alterné

La définition des permutations et du groupe symétrique sont 1.179. Voir aussi le thème 7.

5.6.1 Le groupe alterné

Définition 5.29.
Le groupe A_n des permutations paires[9] *dans S_n est le* ***groupe alterné****.*

Proposition 5.30.
À propos du groupe alterné dans le groupe symétrique.

(1) Le groupe alterné A_n est un sous-groupe caractéristique[10] *de S_n*

(2) Le sous-groupe A_n est d'indice 2 dans S_n.

9. Définition 1.201.
10. Définition 1.167.

(3) Le sous-groupe A_n est l'unique sous-groupe d'indice[11] *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 2.6, il existe un isomorphisme

$$f\colon S_n/\ker(\epsilon) \to \mathrm{Image}(\epsilon). \tag{5.67}$$

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 3.29, 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 $\varphi\colon S_n \to S_n/H$:

$$S_n \xrightarrow{\varphi} S_n/H \xrightarrow{\theta} \{-1, 1\}. \tag{5.68}$$

La composition $\theta \circ \varphi$ est alors un homomorphisme surjectif de S_n sur $\{-1, 1\}$ et nous avons $\theta \circ \varphi = \epsilon$ par la proposition 1.205. L'enchainement (5.68) nous montre que $H = \ker(\theta \circ \varphi) = \ker(\epsilon) = A_n$. □

Proposition 5.31 ([127])**.**
Le groupe symétrique S_n peut être écrit comme un produit semi-direct[12] *du groupe alterné :*

$$S_n = A_n \times_\varphi \mathbb{Z}/2\mathbb{Z} \tag{5.69}$$

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

$$1 \xrightarrow{i} A_n \xrightarrow{i} S_n \xrightarrow{\epsilon} \{\pm 1\} \longrightarrow 1 \tag{5.70}$$

où les i représentent des inclusions et ϵ est la signature définie en 1.202. 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 2.48 nous dit que

$$S_n \simeq A_n \times_\varphi \{\mathrm{Id}, \sigma\} \tag{5.71}$$

où φ est l'action adjointe de $\{\mathrm{Id}, \sigma\}$ sur A_n. □

Proposition 5.32.
Si $\beta \in S_n$ est une transposition, nous avons les égalités suivantes d'ensembles :

$$S_n = A_n \cup A_n\beta = A_n \cup \beta A_n. \tag{5.72}$$

Démonstration. Les parties A_n et βA_n ont le même nombre d'éléments. En effet, l'application

$$\begin{aligned} \varphi\colon A_n &\to A_n\beta \\ \sigma &\mapsto \sigma\beta \end{aligned} \tag{5.73}$$

est une bijection.

De plus ces deux ensembles sont disjoints à cause de la proposition 1.205. En effet si $\sigma \in A_n$, alors $\epsilon(\sigma) = 1$. Mais un élément de $A_n\beta$ est de la forme $\sigma\beta$ avec $\sigma \in A_n$. Or ϵ est un homomorphisme, donc $\epsilon(\sigma\beta) = \epsilon(\sigma)\epsilon(\beta) = -1$.

Enfin, la proposition 5.30(2) dit que A_n est d'indice deux dans S_n. Donc la partie

$$A_n \cup A_n\beta \tag{5.74}$$

contient $|S_n|/2 + |S_n|/2 = |S_n|$ éléments. C'est donc S_n. □

11. Définition 2.12.
12. Définition 2.47.

Lemme 5.33.
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 5.34 ([128]).
Soit $n \geqslant 3$. Les 3-cycles $c_i = (1, 2, i)$ avec $i = 3, \dots, n$ engendrent le groupe alterné A_n.

Démonstration. Soit H, le groupe engendré par les c_i. D'abord nous avons

$$c_i = (1, 2, i) = (1, 2)(2, i), \tag{5.75}$$

de telle sorte que $\epsilon(c_i) = 1$. Par conséquent nous avons $H \subset A_n$. Nous 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 $\{1, \dots, n-1\}$. Par l'hypothèse de récurrence, cette restriction, appartenant à A_{n-1}, se décompose en produit des $c_3, \dots, c_{n-1}$ 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 $(i = 1, \dots, n-1)$ en vertu du point précédent. □

Proposition 5.35.
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 $\{1, \dots, n\}$ 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. Comme $n \geqslant 5$, nous pouvons prendre s et t, des éléments distincts dans $\{1, \dots, n\} \backslash \{j_1, j_2, j_3\}$ et poser $\tau = (st)$. Puisque la signature est un homomorphisme et que τ et α sont impairs, l'élément $\tau\alpha$ est pair (lemme et proposition 1.204 et 1.200) et est donc dans A_n. Les supports de τ et φ étant disjoints, ces derniers commutent et nous avons

$$(\tau\alpha)\sigma(\tau\alpha)^{-1} = \tau(\alpha\sigma\alpha^{-1})\tau^{-1} = \tau\varphi\tau^{-1} = \varphi. \tag{5.76}$$

Donc σ et φ sont conjugués par $\tau\alpha$ qui est dans A_n. □

Théorème 5.36 ([35]).
Le groupe alterné A_n est simple[13] *pour $n \geqslant 5$.*

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 5.34) et que tous les 3-cycles sont conjugués dans A_n (proposition 5.35), 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

$$\theta = (\alpha^{-1}\sigma\alpha)\sigma^{-1} \tag{5.77}$$

est dans N. De plus en utilisant le lemme 1.195 et le fait que $\alpha^{-1} = (ikj)$ nous avons

$$\theta = (ikj)(j\sigma(j)m). \tag{5.78}$$

13. Pas de sous-groupes normaux non triviaux, définition 1.169.

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

$$\theta = (ik)(j\sigma(j)). \tag{5.79}$$

C'est à priori un 2×2-cycle. Mais si de plus $\sigma(j) = i$, alors

$$\theta = (ijk) \tag{5.80}$$

qui est un 3-cycle ; et si $\sigma(j) = k$, alors

$$\theta = (ikj) \tag{5.81}$$

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

$$\theta = (i, j, \sigma(j), m, k) \tag{5.82}$$

est un 5-cycle. Si $\sigma(j) = i$, alors

$$\theta = (ikj)(jim) = (imk) \tag{5.83}$$

qui est un 3-cycle. Si $\sigma(j) = k$, alors

$$\theta = (ikj)(jkm) = (ikm) \tag{5.84}$$

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

$$\underbrace{(abe)^{-1}\theta(abe)}_{\in N}\theta^{-1} = (aeb)(ab)(cd)(abe)(an)(cd) = (abe) \in N. \tag{5.85}$$

Si θ est le 5-cycle $(abcde)$, alors l'élément suivant est dans N :

$$(abc)^{-1}\theta(abc)\theta^{-1} = (acb)(abcde)(abc)(aedcb) = (acd). \tag{5.86}$$

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.

Lemme 5.37.
Le groupe alterné[14] A_6 *n'accepte pas de sous-groupes normaux d'ordre* 60.

14. Définition 5.29.

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 $\mathrm{gr}(a)$ et $\mathrm{gr}(b)$ sont deux 5-Sylow dans A_6. En effet, 5 est un nombre premier, et est la plus grande puissance de 5 dans la décomposition de 60 ; donc $\mathrm{gr}(a)$ est un 5-Sylow dans G. D'autre part, l'ordre de A_6 (qui est $\frac{1}{2} \cdot 6!$) ne possède également que 5 à la puissance 1 dans sa décomposition.

En vertu du théorème de Sylow 5.11(3), les 5-Sylow $\mathrm{gr}(a)$ et $\mathrm{gr}(b)$ 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. □

Le théorème suivant montre que tout groupe peut être vu, en agissant sur lui-même, comme une partie du groupe symétrique.

Théorème 5.38.
Un groupe G est isomorphe à un sous-groupe de son groupe symétrique $S(G)$.

Démonstration. Nous considérons φ, la translation à gauche :

$$\begin{aligned} \varphi \colon G &\to S(G) \\ g &\mapsto t_g \end{aligned} \tag{5.87}$$

où $f_g(h) = gh$. Étant donné que

$$\varphi(gh) = ghx = g(t_h x) = t_g \circ t_h(x), \tag{5.88}$$

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 $\mathrm{Image}(\varphi)$ est donc un sous-groupe de $S(G)$, et φ est un isomorphisme vers ce groupe. □

Lemme 5.39.
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

$$a' = s(a) = (\tau s \tau^{-1})(a) = (\tau s)(b) = \tau(b') \tag{5.89a}$$

$$b' = s(b) = (\tau s \tau^{-1})(b) = (\tau s)(a) = \tau(a'). \tag{5.89b}$$

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 puisque a, b et c sont distincts,

$$\{b\} = \{b, c\} \cap \{a, b\} = \{s(b)\}. \tag{5.90}$$

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 5.40 ([129, 124]).
Tout groupe simple[15] *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 5.11(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 $n_5 = 6$.

Par le point (3) du théorème de Sylow, le groupe G agit transitivement sur l'ensemble des 5-Sylow par l'action adjointe :

$$g \cdot S = gSg^{-1}. \tag{5.91}$$

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

$$\begin{aligned}
(gkg^{-1}) \cdot S &= gkg^{-1}Ggk^{-1}g^{-1} && (5.92a)\\
&= gkTk^{-1}g^{-1} && (5.92b)\\
&= gTg^{-1} && (5.92c)\\
&= S && (5.92d)
\end{aligned}$$

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 il est 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

$$G = D(G) \subset D(S_6) = A_6 \tag{5.93}$$

parce que le groupe dérivé du groupe symétrique est le groupe alterné (lemme 5.33).

L'ensemble $\theta^{-1}(A_6)$ est distingué dans G. En effet si $\sigma \in A_6$ et si $g \in G$ nous avons

$$\theta\big(g\theta^{-1}(\sigma)g^{-1}\big) = \theta(g)\sigma\theta(g)^{-1} \in A_6. \tag{5.94}$$

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

$$[\sigma] = \{h\sigma \text{ tel que } h \in H\}. \tag{5.95}$$

Évidemment A_6 agit sur X de façon naturelle. Au niveau de la cardinalité,

$$\mathrm{Card}(X) = \frac{|A_6|}{|G|} = \frac{360}{60} = 6. \tag{5.96}$$

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 $\varphi(A_6) = A_6$.

15. Définition 1.169.

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 le 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. □

5.6.2 Sous-groupes normaux

5.41 ([130]).
Soit le groupe V_4 engendré par les doubles transpositions de S_4. Nous savons de l'exemple 1.199(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 5.42.
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

$$(s\sigma s^{-1})(b) = (s\sigma)(a) = s(a) = b. \tag{5.97}$$

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 5.43 (Sous-groupes normaux de S_n [130]).
Les sous-groupes normaux de S_n *ne sont pas légions.*

(1) Pour $n = 4$*, les 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 5.39, s n'est pas dans le centre de S_n et il existe $u \in S_n$ tel que $us \neq su$. Comme u est un produit de transpositions (proposition 1.200), il existe une transposition t telle que $st \neq ts$. Le sous-groupe H est normal et puisque $s \in H$ 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 1.199). Nous en concluons que $sts^{-1}t^{-1}$ est un produit de deux transpositions appartenant à 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 élément communs.

(i) **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. Puisque H est normal et que les cycles de longueur trois sont une classe de conjugaison (exemple 1.199) et que A_n est engendré par ceux-ci (proposition 5.34), $A_n \subset H$. 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.

(ii) **Second cas** Le groupe H contient un élément de la forme $(ab)(cd)$ avec a, b, c, d distincts dans $\{1, \dots, n\}$.

(i) **Si $n = 3$** Impossible parce que avec $n = 3$ nous n'avons pas quatre éléments distincts.

(ii) **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

$$4 \leqslant |H| \leqslant 24. \tag{5.98}$$

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 5.30(3)). Quid de $|H| = 8$?

D'après le corolaire 2.14 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.

(i) **x d'ordre 2** Prenons la décomposition de x en cycles disjoints. Puisqu'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.

(ii) **x d'ordre 4** L'élément x est alors un cycle de longueur 4, et H contient tous les cycles de longueur 4 ; par exemple, le produit $(abcd)(bacd) = (adc)$. Le sous-groupe H contient alors A_4 (parce qu'il contient tous les 3-cycles).

(iii) **Si $n \geqslant 5$** Soit un élément e[16] distinct de a, b, c et d. Par notre liste préférée des classes de conjugaison (exemple 1.199(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 :

$$(ce)(ab)(ab)(cd) = (ce)(cd) = (ecd). \tag{5.99}$$

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.

(iv) **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 $n = 4$.

□

5.6.3 Indice

Théorème 5.44.
Tout sous-groupe d'indice n dans S_n est isomorphe à S_n.

Démonstration. Pour $n = 1$, il n'y a pas de sous-groupe. 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.

(i) **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 implique que σ est une transposition. Donc

$$H = \{\mathrm{Id}, (12)\} \tag{5.100}$$

ou l'identité avec (23), ou avec (13). Dans tous les cas c'est isomorphe à S_2.

(ii) **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 5.25 nous dit que H est soit cyclique[17] (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

$$H = \mathbb{Z}_3 \times_\varphi \mathbb{Z}_2 \tag{5.101}$$

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 (5.101). D'abord un point de notations : ici nous considérons les groupes $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ munis de l'addition. Donc

16. e n'est pas l'élément neutre ici
17. Définition 1.270.

1 n'est pas le neutre. Ensuite nous savons du théorème 5.18 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 5.31, 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 5.31 que $S_n = A_n \times_\varphi \mathbb{Z}_2$; en comparant avec (5.101) 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 $(\mathbb{Z}_3, +)$.

(iii) **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 à priori pas un groupe) donnée par $g \cdot [s] = [gs]$.

(i) **Morphisme $\varphi\colon S_n \to S_E$** Le φ défini par l'action est un morphisme parce que

$$\varphi(g_1g_2)[s] = [g_1g_2s] = \varphi(g_1)[g_2s] = \varphi(g_1)\varphi(g_2)[s]. \tag{5.102}$$

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 $s = g^{-1}t$ convient.

(ii) **$\ker(\varphi)$ 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 $\varphi(\sigma z \sigma^{-1}) = \varphi(\sigma)\varphi(z)\varphi(\sigma^{-1}) = \mathrm{Id}_E$.

(iii) **$\ker(\varphi) = \bigcap_{g\in S_n} gHg^{-1}$** 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

$$\varphi(z)[s] = [zs] = [zhs^{-1}s] = [s], \tag{5.103}$$

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.

(iv) **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

$$|\ker(\varphi)| \leqslant |H| = (n-1)! < |A_n|. \tag{5.104}$$

Mais les seuls sous-groupes normaux de S_n sont A_n, S_n et $\{\mathrm{Id}\}$ (proposition 5.43). 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

$$\varphi\colon S_n \to S_E. \tag{5.105}$$

Soit une fonction de numérotation $\psi\colon E \to \{1,\dots,n\}$. Avec cela nous définissons un isomorphisme de groupes

$$\begin{aligned} \tilde{\psi}\colon S_E &\to S_n \\ \sigma &\mapsto \psi\sigma\psi^{-1}. \end{aligned} \tag{5.106}$$

(v) **Fixateur** 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 $\psi[\mathrm{Id}]$:

$$(\tilde{\psi}\circ\varphi)(h)\psi[\mathrm{Id}] = \psi\circ\varphi(h)[\mathrm{Id}] = \psi[h] = \psi[\mathrm{Id}]. \tag{5.107}$$

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 $\sigma\psi[\mathrm{Id}] = \psi[\mathrm{Id}]$. Puisque $\sigma \in S_n$ nous avons $s \in S_E$ tel que $\sigma = \tilde{\psi}(s)$. Pour ce s nous avons donc

$$\big(\tilde{\psi}(s) \circ \psi\big)[\mathrm{Id}] = \psi[\mathrm{Id}], \tag{5.108}$$

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 $\sigma = \tilde{\psi}\big(\varphi(h)\big)$.

(vi) **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,

$$\tilde{\psi} \circ \varphi \colon H \to \mathrm{Fix}\big(\psi[\mathrm{Id}]\big) \tag{5.109}$$

est un isomorphisme de groupe. Mais le stabilisateur d'un point dans S_n est S_{n-1}.

□

5.7 Isométries du cube

Les isométries du cube proviennent de [89].

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-groupe de G constitué des éléments de déterminant positif. Nous notons

$$\mathcal{D} = \{D_1, \dots, D_4\} \tag{5.110}$$

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

$$\rho \colon G \to S_4. \tag{5.111}$$

Nous montrons que ce 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,

$$\overrightarrow{ON} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \overrightarrow{OF} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}. \tag{5.112}$$

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 :

$$\ker(\rho) = \{\mathrm{Id}, s_0\}. \tag{5.113}$$

Le premier théorème d'isomorphisme 2.6 nous permet d'écrire le quotient de groupes :

$$\frac{G}{\{\mathrm{Id}, s_0\}} \simeq S_4. \tag{5.114}$$

Une classe d'équivalence modulo $\ker(\rho)$ dans G est donc toujours de la forme $\{f, f \circ s_0\}$. Et comme 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 qu'en tant que matrice, $s_0 = -\mathbb{1}$, donc les problèmes de commutativité ne se posent pas. L'application

$$\begin{aligned} \varphi\colon \frac{G}{\{\mathrm{Id}, s_0\}} &\to G^+ \\ [g] &\mapsto \begin{cases} g & \text{si } \det(g) > 0 \\ g \circ s_0 & \text{sinon} \end{cases} \end{aligned} \tag{5.115}$$

est un isomorphisme de groupes. Et enfin nous pouvons écrire

$$G^+ \simeq S_4. \tag{5.116}$$

Nous allons maintenant utiliser le corolaire 2.49 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 $\{f, f \circ s_0\}$.

Puisque $G^+ \simeq S_4$ et $\ker(\rho) \simeq \mathbb{Z}/2\mathbb{Z}$ nous pouvons écrire de façon plus brillante que

$$G \simeq S_4 \times_\sigma \mathbb{Z}/2\mathbb{Z}. \tag{5.117}$$

Mais étant donné que la conjugaison par s_0 est triviale, le produit semi-direct est un produit direct :

$$G \simeq S_4 \times \mathbb{Z}/2\mathbb{Z}. \tag{5.118}$$

Il est maintenant du meilleur gout de pouvoir identifier géométriquement ces éléments. Les éléments de $\mathbb{Z}/2\mathbb{Z} = \{\mathrm{Id}, s_0\}$ ne font pas de mystère. Dans S_4 nous avons les classes de conjugaison des éléments Id, (12), (123), (1234) et (12)(34) déterminées durant l'exemple 1.199.

(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 :

$$B \to D \to E \to B. \tag{5.119}$$

C'est une rotation d'angle $\frac{2\pi}{3}$. En tout, nous avons donc 8 rotations d'ordre 3.
Notons à ce propos que la différence entre (234) et (243) est que la première réalise une rotation d'angle $2\pi/3$ tandis que la seconde, 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 [18], c'est-à-dire les rotations d'angle π autour des mêmes axes. Cela fait 3 éléments d'ordre 2.

18. En fait c'est (13)(24), le carré de la précédente, mais c'est la même classe de conjugaison.

Chapitre 6

Corps

6.1 Généralités

6.1.
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 19.34 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 [89] 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.

6.1.1 Corps ordonnés

Nous avons vu la définition de corps totalement ordonné en 1.318.

Définition 6.2 ([60])**.**
Un corps est ***formellement réel*** *si* -1 *n'est pas une somme de carrés.*

Proposition 6.3.
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 1.318(1b). Si $a \leqslant 0$ alors $-a \geqslant 0$ et

$$a^2 = (-a)^2 \geqslant 0. \tag{6.1}$$

Comme $-1 < 0$, il ne peut donc pas être écrit comme un carré. A fortiori comme somme de carrés. □

6.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$

Proposition 6.4 ([131, 1])**.**
L'identité est l'unique automorphisme du corps $\mathbb{R}$*.*

Démonstration. Soit un automorphisme $\sigma\colon \mathbb{R} \to \mathbb{R}$. Comme pour tout automorphisme,

$$\sigma(a) = \sigma(1a) = \sigma(1)\sigma(a). \tag{6.2}$$

Donc $\sigma(1) = 1$.

(i) **Identité sur les rationnels** De plus

$$\sigma(n) = \sigma(1 + \ldots + 1) = \sigma(1) + \ldots + \sigma(1) = n, \tag{6.3}$$

et

$$\sigma\left(\frac{1}{n}\right) + \ldots + \sigma\left(\frac{1}{n}\right) = \sigma\left(\frac{1}{n} + \ldots + \frac{1}{n}\right) = \sigma(1) = 1. \tag{6.4}$$

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.

(ii) **Positive sur les positifs** Si $x > 0$ alors $\sigma(x) = \sigma(\sqrt{x})^2 > 0$.

(iii) **Croissance** Si $x > y$ alors $x - y > 0$ et $\sigma(x - y) > 0$. Cela donne $\sigma(x) > \sigma(y)$.

(iv) **Identité sur les réels** Soit un irrationnel $x \in \mathbb{R}$ et une suite (q_i) dans $\mathbb{Q}$ 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 $\mathbb{Q}$,

$$q_i + \epsilon > \sigma(x). \tag{6.5}$$

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,

$$q_i < \sigma(x) < q_i + \epsilon \tag{6.6}$$

pour tout $\epsilon > 0$ dans $\mathbb{Q}$ et $i > N$. Ce ϵ étant fixé nous pouvons prendre la limite des inégalités (6.6) :

$$x \leqslant \sigma(x) \leqslant x + \epsilon. \tag{6.7}$$

Cela étant valable pour tout $\epsilon > 0$ dans $\mathbb{Q}$, nous avons bien $x = \sigma(x)$.

□

Remarque 6.5.
Certains[131] 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

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = x + y \tag{6.8}$$

parce que ϕ est un morphisme. D'autre part,

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = \phi(a + b). \tag{6.9}$$

Donc

$$\phi^{-1}(x + y) = \phi^{-1}\big(\phi(a) + \phi(b)\big) = \phi^{-1}\big(\phi(a + b)\big) = a + b = \phi^{-1}(x) + \phi^{-1}(y). \tag{6.10}$$

Proposition 6.6.
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 6.4.

En ce qui concerne les imaginaires purs,

$$-1 = \sigma(-1) = \sigma(ii) = \sigma(i)^2. \tag{6.11}$$

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 $x \in \mathbb{R}$

6.1.3 Corps premier

Définition 6.7.
Un corps est ***premier*** *si il est son seul sous-corps. Le* ***sous-corps premier*** *d'un corps est l'intersection de tous ses sous-corps.*

Lemme 6.8.
Le sous corps premier d'un corps est un corps.

Démonstration. Soit un corps $\mathbb{K}$ et son sous-corps premier $\mathbb{P}$. Si $x, y \in \mathbb{P}$, alors pour tout sous-corps $\mathbb{L}$ de $\mathbb{K}$ nous avons $x, y \in \mathbb{L}$ et donc $xy \in \mathbb{L} \subset \mathbb{P}$. Même type de vérification pour les autres propriétés d'un corps. □

Lemme 6.9.
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 6.10.
Soit p un nombre premier. Nous notons $\mathbb{F}_p = \mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$.

Nous verrons plus loin (section 19.4) comment nous pouvons définir $\mathbb{F}_{p^n}$ lorsque p est premier, ainsi que l'unicité d'un tel corps.

Nous avons par exemple

$$\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\} \tag{6.12}$$

avec la loi $2 = 0$.

Notons que $\mathbb{F}_p$ est un corps[2] possédant p éléments. L'ensemble $\mathbb{F}_p^*$ est un groupe d'ordre $p-1$.

Lemme 6.11.
Les corps $\mathbb{Q}$ et $\mathbb{Z}/p\mathbb{Z}$ (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 3.66 que $\mathbb{F}_p$ est un corps lorsque p est premier. □

Lemme 6.12.
Si $k \in \mathbb{Z}$, nous notons $k_{\mathbb{K}}$ l'élément $k \cdot 1_{\mathbb{K}} = \sum_{i=1}^{k} 1_{\mathbb{K}}$. L'application

$$\begin{aligned} s\colon \mathbb{Z} &\to \mathbb{K} \\ k &\mapsto k_{\mathbb{K}} \end{aligned} \tag{6.13}$$

est un morphisme d'anneaux.

Si la caractéristique de $\mathbb{K}$ est 0, alors s est injective.

Proposition 6.13 (Sous-corps premier[132])**.**
Soit $\mathbb{K}$, un corps[3] de caractéristique p.

(1) Si $p > 0$ alors le sous-corps premier de $\mathbb{K}$ est isomorphe à $\mathbb{Z}$.

(2) Si $p = 0$ alors le sous-corps premier de $\mathbb{K}$ est isomorphe à $\mathbb{Q}$.

Démonstration. Pour les besoins de notations, dans la suite si $k \in \mathbb{Z}$, nous notons $k_{\mathbb{K}}$ l'élément $k \cdot 1_{\mathbb{K}} = \sum_{i=1}^{k} 1_{\mathbb{K}}$. Notons aussi $\mathbb{P}$ le sous-corps premier de $\mathbb{K}$.

2. Quand p est premier, $\mathbb{Z}/p\mathbb{Z}$ est un corps, proposition 3.66.
3. Je rappelle que tous les corps du Frido sont commutatifs, sauf mention explicite du contraire.

Étant donné que $1_{\mathbb{K}}$ est dans tout sous-corps, nous devons avoir $\mathbb{Z}1_{\mathbb{K}} \subseteq \mathbb{P}$. Nous avons donc toujours

$$\mathbb{Z}1_{\mathbb{K}} \subset \mathbb{P} \subset \mathbb{K}. \tag{6.14}$$

Maintenant, le lemme 1.296 nous dit la nature de $\mathbb{Z}1_{\mathbb{K}}$ en tant qu'anneaux, et nous pouvons séparer les cas en fonction de la caractéristique.

(i) **Si $p > 0$** Nous avons $\mathbb{Z}1_{\mathbb{K}} \simeq \mathbb{Z}/p\mathbb{Z}$ par le lemme 1.296. L'ensemble $\mathbb{Z}/p\mathbb{Z}$ étant un corps, $\mathbb{Z}1_{\mathbb{K}}$ est un corps et c'est donc le corps premier de $\mathbb{K}$.

(ii) **Si $p = 0$** Si $p = 0$, alors $\mathbb{Z}1 \simeq \mathbb{Z}$ en tant qu'anneaux. Dommage que $\mathbb{Z}$ ne soit pas un corps. Soit $n \in \mathbb{Z}1_{\mathbb{K}}$ et $m \in (\mathbb{Z}1_{\mathbb{K}})^*$. L'élément $nm^{-1} \in \mathbb{K}$ est dans $\mathbb{P}$ parce que $\mathbb{P}$ est un corps. Donc l'application

$$\begin{aligned} \psi\colon \mathbb{Q} &\to \mathbb{P} \\ ab^{-1} &\mapsto a_{\mathbb{K}} b_{\mathbb{K}}^{-1} \end{aligned} \tag{6.15}$$

prend bien ses valeurs dans $\mathbb{P}$ et est un morphisme de corps.

La surjectivité de ψ est ce dont nous venons de parler. En ce qui concerne l'injectivité, si $\psi(ab^{-1}) = \psi(xy^{-1})$, alors $a_{\mathbb{K}} b_{\mathbb{K}}^{-1} = x_{\mathbb{K}} y_{\mathbb{K}}^{-1}$ et la commutativité de $\mathbb{K}$ nous permet d'écrire

$$a_{\mathbb{K}} y_{\mathbb{K}} = x_{\mathbb{K}} b_{\mathbb{K}}. \tag{6.16}$$

Étant donné le lemme 6.12, nous déduisons que $(ay - xb)_{\mathbb{K}} = 0$. Et comme $z \mapsto z_{\mathbb{K}}$ est injective, nous en déduisons que $ay - xb = 0$ dans $\mathbb{Z}$, et donc que $ab^{-1} = xy^{-1}$ dans $\mathbb{Q}$ par le lemme 1.331.

□

Proposition 6.14.
Soit $\mathbb{K}$ un corps et $\mathbb{P}$ son sous-corps premier[4]*. Si $\sigma \in \mathrm{Aut}(\mathbb{K})$ alors $\sigma|_{\mathbb{P}} = \mathrm{Id}$, c'est-à-dire que $\sigma(x) = x$ pour tout $x \in \mathbb{P}$.*

Démonstration. Soit $\mathbb{P}$ le sous-corps premier de $\mathbb{K}$. La proposition 6.13 nous dit que $\mathbb{P}$ est soit $\mathbb{Z}$ soit $\mathbb{Q}$ en fonction de la caractéristique de $\mathbb{K}$. Nous allons faire le cas de caractéristique nulle, c'est à dire $\mathbb{P} \simeq \mathbb{Q}$.

Tout élément de $\mathbb{P}$ peut être écrit sous la forme $k_{\mathbb{K}} l_{\mathbb{K}}^{-1}$ pour des $k, l \in \mathbb{Z}$. Nous calculons la valeur de σ sur ces éléments. D'abord σ satisfait $\sigma(1_{\mathbb{K}})$ et préserve la somme, dont $\sigma(k_{\mathbb{K}}) = k_{\mathbb{K}}$ pour tout $k \in \mathbb{Z}$[5]. Étant donné que σ préserve également l'inverse et le produit,

$$\sigma(k_{\mathbb{K}} l_{\mathbb{K}}^{-1}) = \sigma(k_{\mathbb{K}})\sigma(l_{\mathbb{K}})^{-1} = k_{\mathbb{K}} l_{\mathbb{K}}^{-1}. \tag{6.17}$$

Et voila. □

6.1.4 Petit théorème de Fermat

Théorème 6.15 (Petit théorème de Fermat)**.**
Soit p un nombre premier.

(1) Si $x \in \mathbb{Z}/p\mathbb{Z}$ alors $x^p = x$.

(2) Si $x \in (\mathbb{Z}/p\mathbb{Z})^*$, alors $x^{p-1} = 1$.

(3) Si $x \in (\mathbb{Z}/p\mathbb{Z})^*$ alors $x^{-1} = x^{p-2}$.

(4) Si x est premier avec p, alors $x^{p-1} \in [1]_p$.

4. Définition 6.7.
5. Dans le cas de caractéristique non nulle, $\mathbb{P} = \mathbb{Z}$ et vous pouvez vous arrêter ici.

Démonstration. Étant donné que $\mathbb{Z}/p\mathbb{Z}$ est un corps commutatif et que p est premier, la proposition 3.43 nous indique que $\sigma(x) = x^p$ est un automorphisme. La proposition 6.14 nous indique alors que

$$a^p = a. \tag{6.18}$$

Si a est inversible alors $a^{p-1} = a^p a^{-1} = 1$.

Pour (4). Le nombre x n'est pas premier avec p uniquement lorsque x est multiple de p. Dans ce cas c'est $[a]_p = 0$ dans $\mathbb{Z}/p\mathbb{Z}$ et donc $a^{p-1} = 0$. □

Exemple 6.16.
Soit $\mathbb{K} = \mathbb{F}_{29}$. Le nombre 29 étant premier, $\mathbb{K}$ est un corps. C'est le corps des entiers modulo 29. Nous avons donc

$$-142 = -113 = -84 = -55 = -26 = 3 = 32 = 61 = 90 = 119. \tag{6.19}$$

Le petit théorème de Fermat nous permet aussi de calculer des exposants et des inverses. En effet, puisque $1 = x^{28}$ pour tout $x \in \mathbb{F}_{29}^*$, nous avons $x^{-1} = x^{27}$, et par suite, pour tout entier a,

$$x^{-a} = (x^a)^{27} = x^{27a}. \tag{6.20}$$

Le nombre $27a$ peut être grand par rapport à 29. Mais en réutilisant le fait que $1 = x^{28}$, on obtient

$$x^{-a} = x^{[27a]_{28}}. \tag{6.21}$$

Cette expression doit être comprise comme disant que pour tout $k \in [27a]_{28}$ nous avons $x^{-a} = x^k$.

Chose à retenir : dans les exposants nous calculons modulo 28. △

6.1.5 Nombres de Sophie Germain

Définition 6.17.
*Un nombre premier p est de **Sophie Germain** si $p > 2$ et si le nombre $q = 2p + 1$ est également premier.*

Notons qu'il existe des nombres premiers de Sophie Germain. Par exemple $p = 3$ donne $q = 2 \times 3 + 1 = 7$. Comme 7 est un nombre premier, le nombre 3 est de Sophie Germain. D'après Wikipédia[133], l'existence d'une infinité de nombres premiers de Sophie Germain est encore une question ouverte[6].

Lemme 6.18 ([134, 135])**.**
Soit p, un nombre premier de Sophie Germain. Si $m \in \mathbb{Z}$ n'est pas divisible par q, alors $m^p \in [\pm 1]_q$ où $q = 2p + 1$.

Démonstration. Le petit théorème de Fermat 6.15 nous dit que $m^{q-1} \in [1]_q$. Mais $m^{q-1} = m^{2p} = (m^p)^2$. Donc

$$(m^p)^2 \in [1]_q. \tag{6.22}$$

Puisque q est premier, l'anneau $\mathbb{Z}/q\mathbb{Z}$ est intègre par le corolaire 1.248. Comme $1 \neq -1$, la proposition 3.151 s'applique et nous avons $m^p \in [\pm 1]_q$. □

Théorème 6.19 ([134, 135, 136])**.**
Soit un nombre premier de Sophie Germain (définition 6.17). Il n'existe pas de solution (x, y, z) dans $\mathbb{Z}^3$ au système

$$\begin{cases} x^p + y^p + z^p = 0 \\ [xyz]_p \neq 0. \end{cases} \tag{6.23}$$

Démonstration. Nous supposons que $(x, y, z) \in \mathbb{Z}^3$ est une solution telle que $\operatorname{pgcd}(x, y, z) = 1$.

6. Si vous lisez ce paragraphe dans un futur où la question est tranchée, n'hésitez pas à m'écrire pour mettre à jour.

(i) **x, y et z sont premiers deux à deux** Supposons que k soit un diviseur commun à x et y. Alors il existe $\alpha, \beta \in \mathbb{Z}$ tels que $x = k\alpha$ et $y = k\beta$. Du coup on aurait $z^p = -k^p(\alpha^p + \beta^p)$. Donc k divise z^p. La proposition 3.19 nous indique qu'alors k divise z. Donc $k = 1$ parce que k diviserait x, y et z.

Ce qui est vrai pour le couple (x, y) est encore vrai pour (x, z) et pour (y, z).

(ii) **Le fameux u** Nous introduisons à présent une nouvelle variable intermédiaire qui va beaucoup nous aider. En utilisant le lemme 3.42,

$$-x^p = y^p + z^p = (y + z)\sum_{k=0}^{p-1} y^k(-z)^{p-1-k} = (y + z)u \tag{6.24}$$

où nous avons posé $u = \sum_{k=0}^{p-1} y^k(-z)^{p-1-k}$.

(iii) $\text{pgcd}(u, y + z) = 1$ Soit α un nombre premier qui divise u et $y + z$.

D'abord α divise x^p et donc x (proposition 3.19). Autrement dit : $[x]_\alpha = 0$.

Ensuite, $0 = [x + z]_\alpha$, de telle sorte que

$$[y]_\alpha = -[z]_\alpha \tag{6.25}$$

Enfin, en prenant la classe de u modulo α, et en substituant (6.25) :

$$0 = [u]_\alpha = \Big[\sum_{k=0}^{p-1} y^k(-z)^{p-1-k}\Big]_\alpha \tag{6.26a}$$

$$= \sum_{k=0}^{p-1} [y^k]_\alpha [y^{p-1-k}]_\alpha \tag{6.26b}$$

$$= \sum_{k=0}^{p-1} [y]_\alpha^{p-1} \tag{6.26c}$$

$$= p[y]_\alpha^{p-1}. \tag{6.26d}$$

Nous avons donc $p[y]_\alpha^{p-1} = 0$. Mais $\mathbb{Z}/\alpha\mathbb{Z}$ est un anneau intègre[7]. Donc la règle du produit nul s'applique[8] et nous avons deux possibilités : $[p]_\alpha = 0$ ou $[y]_\alpha^{p-1} = 0$.

D'abord α divise déjà x. Or $\text{pgcd}(x, y) = 1$. Donc α ne peut pas diviser y (à part si $\alpha = 1$ évidemment), et à fortiori pas y^{p-1} non plus. Nous en déduisons que $[p]_\alpha = 0$. Puisque p est premier, nous avons $p = \alpha$.

Mais nous avons déjà vu que α divisait x. Donc $0 = [x]_\alpha = [x]_p$. Cela n'est pas possible parce que $[xyz]_p \neq 0$.

Nous avons donc prouvé que $\text{pgcd}(u, y + z) = 1$.

(iv) **Utilisation d'un lemme** Pour rappel, p est impair. Nous avons $(-x)^p = (y + z)u$. Le lemme 3.23 nous dit qu'il existe $a, \alpha \in \mathbb{Z}$ tels que

$$\begin{cases} y + z = a^p \\ u = \alpha^p. \end{cases} \tag{6.27}$$

Ce que nous venons de faire pour $y + z$ peut être fait pour les autres couples. Donc il existe $a, b, c \in \mathbb{Z}$ tels que

$$y + z = a^p \tag{6.28a}$$

$$x + z = b^p \tag{6.28b}$$

$$x + y = c^p. \tag{6.28c}$$

7. Corolaire 1.248.
8. Proposition 1.221.

(v) **q divise un et un seul des x, y ou z** Supposons que q ne divise ni x, ni y ni z. Alors le lemme 6.18 dit que $x^p \in [\pm 1]_q$, $y^p \in [\pm 1]_q$ et $z^p \in [\pm 1]_q$. Donc les valeurs possibles pour $[x^p + y^p + z^p]_q$ sont $[\pm 3]_q$ et $[\pm 1]_q$.

Mais nous avons supposé que $x^p + y^p + z^p = 0$ et que $q \geqslant 5$ de telle sorte que $[\pm 3]_q \neq 0$. Contradiction. Donc q divise au moins un des trois x, y ou z.

Nous avons déjà montré que x, y et z étaient deux à deux premiers entre eux. Donc q ne peut diviser qu'un seul des trois.

(vi) **q divise x** Jusqu'à présent x, y et z avaient des rôles symétriques. Maintenant nous supposons que q divise x.

(vii) $[y^p]_q = [\pm 1]_q$ Nous savons que $c^p = x + y$. En passant aux classes modulo q, nous avons

$$[c^p]_q = [x]_q + [y]_q = [y]_q \tag{6.29}$$

parce que nous avons choisi que q divise x. Nous avons maintenant les implications suivantes :

q ne divise pas y.
$\Rightarrow$ q ne divise pas c^p parce que $[y]_q = [c^p]_q$.
$\Rightarrow$ q ne divise pas c par le lemme 3.20.
$\Rightarrow$ $c^p \in [\pm 1]_q$ par le lemme 6.18.
$\Rightarrow$ $y \in [\pm 1]_q$ toujours parce que $[y]_q = [c^p]_q$.

(viii) $[z^p]_q = [\pm 1]_q$ Parce que les rôles de y et z sont symétriques. Donc ce que nous avons fait pour obtenir $[y^p]_q = [\pm 1]_q$ tient pour avoir $[z^p]_q = [\pm 1]_q$.

(ix) **q divise $y + z$** Nous avons dit que q divisait x. Il divise donc aussi $2x$. Nous avons :

$$2x = (x + y) + (x + z) - (y + z) = c^p + b^p - a^p. \tag{6.30}$$

En passant au modulo q, à gauche nous avons $[2x]_q = 0$. Donc

$$\begin{aligned}
[a^p]_q &= [c^p]_q + [b^p]_q && (6.31a)\\
&= [y]_q + [z]_q && (6.31b)\\
&= [\pm 1]_q + [\pm 1]_q && (6.31c)\\
&\in \{[2]_q, [-2]_q, [0]_q\}. && (6.31d)
\end{aligned}$$

En tout état de cause, $[a^p]_q$ n'est pas $[\pm 1]_q$. La contraposée du lemme 6.18 nous dit alors que q divise $a^p = y + z$.

(x) $[y]_q = [-z]_q$ Il s'agit seulement du fait que q divise $y + z$.

(xi) $[u]_q = [p]_q$ Par définition, $u = \sum_{k=0}^{p-1} y^k (-z)^{p-1-k}$. On passe au modulo q, en substituant $[-z]_q$ par $[y]_q$:

$$[u]_q = \sum_{k=0}^{p-1} [y]_q^k [y]_q^{p-1-k} = \sum_{k=0}^{p-1} [y]_q^{p-1} = p[y]_q^{p-1} = [p]_q \tag{6.32}$$

La dernière égalité est le fait que $[y]_q = [\pm 1]_q$ et que $p - 1$ est pair.

(xii) **La contradiction** Vous vous souvenez de (6.27) qui disait que $u = \alpha^p$ pour un certain p ? Nous avons les implications suivantes :

$[\alpha^p]_q = [p]_q \neq 0 \Rightarrow q$ ne divise pas α^p.
$\Rightarrow q$ ne divise pas α.
$\Rightarrow$ $\alpha^p \in [\pm 1]_q$.
$\Rightarrow$ $[p]_q \in [\pm 1]_q$.

La dernière ligne est clairement fausse parce que $q > p$ et p n'est certainement pas égal à 1 ou -1.

Nous avons prouvé le théorème pour les triples (x, y, z) sans facteur commun. Si k était un diviseur commun de x, y et z, alors $(x/k, y/k, z/k)$ serait encore une solution, ce qui n'est pas le cas. □

6.2 Théorème des deux carrés

Proposition 6.20.
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. □

6.2.1 Un peu de structure dans $\mathbb{Z}[i]$

Lemme 6.21.
L'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a + bi &\mapsto a^2 + b^2 \end{aligned} \tag{6.33}$$

est un stathme euclidien pour $\mathbb{Z}[i]$.

Démonstration. Soient $t, z \in \mathbb{Z}[i]\backslash\{0\}$ dont le quotient s'écrit

$$\frac{z}{t} = x + iy \tag{6.34}$$

dans $\mathbb{C}$. Nous considérons $q = a + bi$ où a et b sont les entiers les plus proches de x et y. Si il y a *ex aequo*, on prend au hasard [9]. Alors nous avons

$$|\frac{z}{t} - q| \leqslant \frac{|1+i|}{2} = \frac{\sqrt{2}}{2} < 1. \tag{6.35}$$

On pose $r = z - qt$ qui est bien un élément de $\mathbb{Z}[i]$. De plus

$$|r| = |z - qt| = |t||\frac{z}{t} - q| < |t|, \tag{6.36}$$

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 3.107).

Lemme 6.22.
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

$$1 = N(zz') = N(z)N(z'), \tag{6.37}$$

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

$$\mathbb{Z}[i]^* = \{\pm 1, \pm i\}. \tag{6.38}$$

□

Définition 6.23 ([137]).
*Un **monoïde** est un triplet $(E, *, e)$ où E est un ensemble, e est un élément de E et $*\colon E \times E \to E$ est une loi de composition telle que pour tout $x, y \in E$,*

9. Dans l'exemple 3.106, nous prenions toujours l'inférieur parce que le stathme tenait compte de la positivité.

(1) $x * (y * z) = (x * y) * z$ *(associativité)*

(2) $e * x = x * e = x$ *(e est un neutre).*

Nous notons $\Sigma = \{a^2 + b^2 \text{ tel que } a, b \in \mathbb{N}\}$.

Lemme 6.24.
L'ensemble $\Sigma = \{a^2 + b^2 \text{ tel que } a, b \in \mathbb{N}\}$ *est un sous-monoïde*[10] *de* $\mathbb{N}$.

Démonstration. Il suffit de prouver que si $m, n \in \Sigma$, alors le produit mn est également dans Σ. Soit N, le stathme euclidien sur $\mathbb{Z}[i]$ (celui donné par le lemme 6.21). Vu que $n \in \Sigma$, il existe $z \in \mathbb{Z}[i]$ tel que $N(z) = n$. Idem pour m : il existe $z' \in \mathbb{Z}[i]$ tel que $N(z') = m$. Nous avons évidemment $zz' \in \mathbb{Z}[i]$, et

$$N(zz') = N(z)N(z') = nm. \tag{6.39}$$

Donc nm est l'image de zz' par N, ce qui prouve que $nm \in \Sigma$. □

Théorème 6.25 (Théorème des deux carrés, version faible)**.**
Un nombre premier est somme de deux carrés si et seulement si $p = 2$ *ou* $p \in [1]_4$.

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

$$\mathbb{Z}[i] = \{a + bi \text{ tel que } a, b \in \mathbb{Z}\}. \tag{6.40}$$

puis l'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a + bi &\mapsto a^2 + b^2. \end{aligned} \tag{6.41}$$

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 6.22).

Le lemme 6.21 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 3.107).

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

$$p^2 = N(p) = N(z)N(z'). \tag{6.42}$$

Comme p est premier par hypothèse, c'est uniquement possible avec $N(z) = N(z') = p$ (avoir $N(z) = 1$ est impossible parce que cela signifirait 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 3.81 s'applique donc et p sera non irréductible si et seulement si l'idéal (p) est 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

10. Définition 6.23.

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 $(\mathbb{A}/I)/J \simeq (\mathbb{A}/J)/I$. Par conséquent, en tenant compte du fait que $\mathbb{Z}[i] = \mathbb{Z}[X]/(X^2+1)$, nous avons

$$\mathbb{Z}[i]/(p) = (\mathbb{Z}[X]/(p))/(X^2+1) = \mathbb{F}_p[X]/(X^2+1). \tag{6.43}$$

Nous avons donc équivalence des propositions suivantes :

$$p \in \Sigma \tag{6.44a}$$
$$\mathbb{F}_p[X]/(X^2+1) \text{ n'est pas intègre} \tag{6.44b}$$
$$X^2+1 \text{ n'est pas irréductible dans } \mathbb{F}_p \tag{6.44c}$$
$$X^2+1 \text{ admet une racine dans } \mathbb{F}_p \tag{6.44d}$$
$$-1 \in (\mathbb{F}_p^*)^2 \tag{6.44e}$$
$$\exists y \in \mathbb{F}_p^* \text{ tel que } y^2 = -1. \tag{6.44f}$$

Le point (6.44c) vient de la proposition 6.20. 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,

$$(-1)^{(p-1)/2} = (y^2)^{(p-1)/2} = y^{p-1} = 1 \tag{6.45}$$

parce que dans $\mathbb{F}_p$ nous avons $y^{(p-1)} = 1$ par le petit théorème de Fermat (théorème 6.15). Ainsi p doit vérifier

$$1 = (-1)^{(p-1)/2}, \tag{6.46}$$

c'est-à-dire $\frac{p-1}{2} = 0 \mod 2$ ou encore $p = 1 \mod 4$. □

Remarque 6.26.
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 $[1]_4$.

Théorème 6.27 (Théorème des deux carrés[89])**.**
Soit $n \geqslant 2$ un nombre dont nous notons

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)} \tag{6.47}$$

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. (i) **Condition suffisante.** Le produit (6.47) 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

$$\prod_{p \in \mathcal{P} \cap [1]_4} p^{v_p(n)} \in \Sigma. \tag{6.48}$$

— Le seul nombre premier dans $[2]_4$ est 2. C'est un élément de Σ.

— Le produit

$$\prod_{p \in \mathcal{P} \cap [3]_4} p^{v_p(n)} \tag{6.49}$$

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 6.24.

(ii) **Condition nécessaire.** Soit p, un nombre premier. Nous voulons montrer que

$$\{v_p(n) \text{ tel que } n \in \Sigma\} \subset [2]_2. \tag{6.50}$$

Pour montrer cela nous allons procéder par récurrence sur les ensembles

$$E_k = \{v_p(n) \text{ tel que } n \in \Sigma\} \cap \{0, \dots, k\}. \tag{6.51}$$

Il est évident que les éléments de E_0 sont pairs, puisqu'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

$$n = a^2 + b^2 = (a + bi)(a - bi) \tag{6.52}$$

Comme $\mathbb{Z}[i]$ est principal, le lemme de Gauss 3.93 nous dit que si p divise n, alors il doit diviser soit $a + bi$, soit $a - bi$ (et en fait, les deux). Nous avons alors $p \mid a$ et $p \mid b$ en même temps. Et donc

$$p^2 \mid a^2 + b^2 = n. \tag{6.53}$$

Posons $a = pa'$ et $b = pb'$ avec $a', b' \in \mathbb{N}$. Nous avons

$$\frac{n}{p^2} = \frac{p^2a'^2 + p^2b'^2}{p^2} = a'^2 + b'^2 \in \Sigma. \tag{6.54}$$

Mais par construction,

$$v_p\left(\frac{n}{p^2}\right) = v_p(n) - 2 < k. \tag{6.55}$$

Donc $v_p(\frac{n}{p^2})$ est pair et par conséquent, $v_p(n)$ doit également être pair.

□

6.2.2 Résultats chinois

Nous allons maintenant parler du système d'équations

$$\begin{cases} x = a_1 \mod p & (6.56a) \\ x = a_2 \mod q & (6.56b) \end{cases}$$

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 corolaire qui établira l'isomorphisme d'anneaux $\mathbb{Z}/pq\mathbb{Z} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$. Voir [138].

Notons que le système (6.56), en suivant les notations usuelles du Frido s'écrit plutôt

$$x \in [a_1]_p \cap [a_2]_q. \tag{6.57}$$

Lemme 6.28 (Lemme chinois [139])**.**
Soient n_1, n_2 deux entiers premiers entre eux[11]*. Soient $a_1, a_2 \in \mathbb{Z}$.*

(1) Il existe $u_1, u_2 \in \mathbb{Z}$ tels que

$$u_1n_1 + u_2n_2 = 1. \tag{6.58}$$

(2) Nous posons

$$a = a_1u_2n_2 + a_2n_1u_1. \tag{6.59}$$

Nous avons

$$[a_1]_{n_1} \cap [a_2]_{n_2} = [a]_{n_1n_2}. \tag{6.60}$$

11. Définition 1.227.

Démonstration. En plusieurs points.

(i) **Pour (1)** Il suffit d'utiliser le théorème de Bézout 1.228

(ii) **Pour (2) première inclusion** Soit $x \in [a]_{n_1n_2}$, et vérifions que $x \in [a_1]_{n_1}$. Il existe $k \in \mathbb{Z}$ tel que

$$x = a_1u_2n_2 + a_2n_1u_1 + kn_1n_2. \tag{6.61}$$

En remplaçant u_2n_2 par $1 - u_1n_1$, nous avons

$$x = a_1(1 - u_1n_1) + a_2n_1u_1 + kn_1n_2 \in [a_1]_{n_1} \tag{6.62}$$

parce qu'il n'y a que le tout premier terme qui ne contient pas de facteur n_1.
Le fait que $x \in [a_2]_{n_2}$ se vérifie de même[12].

(iii) **Pour (2) seconde inclusion** Soit $x \in [a_1]_{n_1} \cap [a_2]_{n_2}$. Il existe $k \in \mathbb{Z}$ tel que $x = a_1 + kn_2$. Donc

$$\begin{aligned} x - a &= a_1 + kn_1 - (a_1u_2n_2 + a_2n_1u_1) && (6.63a)\\ &= a_1 - a_1u_2n_2 + k'n_1 && (6.63b)\\ &= a_1\underbrace{(1 - u_2n_2)}_{=u_1n_1} + k'n_1 && (6.63c)\\ &= a_1u_1n_1 + k'n_1, && (6.63d) \end{aligned}$$

ce qui signifie que n_1 divise $x - a$..
De même nous prouvons que n_2 divise $x - a$. Il existe $k, l \in \mathbb{Z}$ tels que

$$x - a = kn_1 = ln_2. \tag{6.64}$$

En particulier n_1 divise ln_2 alors que n_1 et n_2 sont premiers entre eux. Le lemme de Gauss 3.93 dit alors que n_1 divise l. Il existe donc $s \in \mathbb{Z}$ tel que $l = sn_1$. En remettant ce l dans (6.64), nous trouvons

$$x - a = ln_1 = sn_1n_2, \tag{6.65}$$

et donc $x - a$ est divisible en n_1n_2, c'est à dire

$$x \in [a]_{n_1n_2}, \tag{6.66}$$

ce qu'il fallait prouver.

□

Théorème 6.29 (Théorème chinois)**.**
Soient p, q deux naturels premiers entre eux. Si $p, q \geqslant 2$ alors l'application

$$\begin{aligned} \phi\colon \mathbb{Z}/pq\mathbb{Z} &\to \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}\\ [x]_{pq} &\mapsto \big([x]_p, [x]_q\big) \end{aligned} \tag{6.67}$$

est un isomorphisme d'anneaux.

Démonstration. Nous devons prouver que l'application ϕ respecte la somme, le produit et qu'elle est bijective. En ce qui concerne la somme,

$$\begin{aligned} \phi([q]_{pq} + [y]_{pq}) &= \phi\big((x + y) \mod pq\big) && (6.68a)\\ &= \big([x + y]_p, [x + y]_q\big) && (6.68b)\\ &= \big([x]_p + [y]_p, [x]_q + [y]_q\big) && (6.68c)\\ &= \big([x]_p, [x]_q\big) + \big([y]_p, [y]_q\big) && (6.68d)\\ &= \phi(x) + \phi(y). && (6.68e) \end{aligned}$$

12. Je dis ça au hasard ; je n'ai pas vérifié. Faites-moi signe si c'est bon.

En ce qui concerne le produit, c'est le même jeu : nous obtenons

$$\phi\big([xy]_{pq}\big) = \phi([x]_{pq})\phi([y]_{pq}) \tag{6.69}$$

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

$$\phi([x]_{pq}) = \big([y_1]_p, [y_2]_q\big) \tag{6.70}$$

revient à demander que $[x]_p = [y_1]_p$ et $[x]_q = [y_2]_q$, c'est-à-dire que x résolve le système

$$\begin{cases} x = y_1 \mod p & (6.71a) \\ x = y_2 \mod q. & (6.71b) \end{cases}$$

Le lemme chinois 6.28 nous assure qu'une solution existe.

En ce qui concerne l'injectivité, nous supposons que x et y soient deux entiers tels que

$$\phi([x]_{pq}) = \phi([y]_{pq}). \tag{6.72}$$

Nous en déduisons le système

$$\begin{cases} x \mod p = y \mod p & (6.73a) \\ x \mod q = y \mod q & (6.73b) \end{cases}$$

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

$$kp + lq = 0. \tag{6.74}$$

Étant donné que p et q sont premiers entre eux, la seule possibilité est $k = l = 0$, c'est-à-dire $x = y$. □

Théorème 6.30 (Théorème chinois[1])**.**
Soit A un anneau commutatif. Soient $n \geqslant 2$, 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. En plusieurs points.

(i) **Définition de J_k** Pour $k \in \{1, \dots, n\}$ nous notons J_k le produit $J_k = \prod_{l \neq k} I_l$.

(ii) **Si $k \neq i$ alors $J_k \subset I_i$** Un élément de J_k est de la forme

$$\prod_{l \neq k} a_j = a_i \prod_{\substack{l \neq k \\ l \neq i}} a_l \tag{6.75}$$

avec $a_l \in I_l$. Vu que a_i est dans I_i qui est un idéal, tout le produit (6.75) est dans I_i.

(iii) $I_i + J_i = A$ En effet, soit $k \neq i$. Nous avons $J_i \subset I_k$ et donc

$$I_i + J_i \subset I_i + I_k = A. \tag{6.76}$$

(iv) **Des α et des β** Soit $i \in \{1, \dots, n\}$. Puisque $I_i + J_i = A$, nous pouvons donc prendre $\alpha_i \in I_i$ et $\beta_i \in J_i$ tels que

$$\alpha_i + \beta_i = 1. \tag{6.77}$$

(v) **Et enfin** Nous posons $x = \sum_{i=1}^{n} \beta_i x_i$. Pour chaque k nous avons

$$x - x_k = \sum_{i=1}^{n} \beta_i x_i - x_k \tag{6.78a}$$
$$= \sum_{i \neq k} \beta_i x_i + (\beta_k - 1) x_k \tag{6.78b}$$
$$= \sum_{i \neq k} \beta_i x_i - \alpha_k x_k. \tag{6.78c}$$

Les deux termes sont dans I_k. En effet, d'une part $\beta_i \in J_i \subset I_k$ et I_k est un idéal, donc $\beta_i x_i \in I_k$, et d'autre part $\alpha_k \in I_k$, donc $\alpha_k x_k \in I_k$.
Un idéal étant stable par somme, la somme du tout est dans I_k.

□

Remarque 6.31.
Ce théorème chinois est bien une généralisation du lemme chinois 6.28. 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 corolaire 1.230.

6.3 Polynômes à coefficients dans un corps

Nous supposons que $\mathbb{K}$ est un corps commutatif, et nous étudions l'anneau $\mathbb{K}[X]$, défini en 1.303.

Proposition 6.32.
L'anneau $\mathbb{K}[X]$ des polynômes sur un corps commutatif $\mathbb{K}$ est factoriel.

Le théorème suivant est un cas particulier pour $\mathbb{K}[X]$ du théorème chinois 3.88.

Théorème 6.33 (Théorème chinois)**.**
Si P et Q sont deux polynômes premiers entre eux, alors nous avons l'isomorphisme

$$\mathbb{K}[X]/(P, Q) \simeq \mathbb{K}[X]/(P) \times \mathbb{K}[X]/(Q). \tag{6.79}$$

6.3.1 Irréductibilité

Définition 6.34 ([140])**.**
Un polynôme à coefficients dans un anneau commutatif est irréductible si il

(1) n'est pas inversible,

(2) n'est pas le produit de deux non inversibles.

Un polynôme est irréductible dans $\mathbb{K}[X]$ au sens de la définition 3.67 si et seulement si il est irréductible au sens de la définition 6.34 parce que seules les constantes (non nulles) sont inversibles dans $\mathbb{K}[X]$.

Exemple 6.35.
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

$$P = (X^2 + 1)(X^2 + 2), \tag{6.80}$$

mais n'a pas de racines dans $\mathbb{Z}$. Si on veut réduire plus, il faut sortir de $\mathbb{Z}$.

△

Définition 6.36 (Polynôme scindé)**.**
*Nous disons que $P \in \mathbb{K}[X]$ est **scindé** sur $\mathbb{K}$ si il est produit dans $\mathbb{K}[X]$ de polynômes de degré 1.*

*Soient $P \in \mathbb{K}[X]$ un polynôme irréductible, et $\mathbb{L}$ un corps, extension du corps $\mathbb{K}$. On dit que P est **scindé** dans $\mathbb{L}$ si P se décompose en un produit de polynômes de degré 1 dans $\mathbb{L}[X]$.*

Note : les constantes ne sont donc pas des polynômes scindés.

Proposition 6.37 (Critère d'Eisenstein)**.**
Soit le polynôme $P = \sum_{k=0}^{n} a_n X^n$ 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 3.138) 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 (3.93) 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 soit primitif au sens du pgcd. Il est donc irréductible et primitif sur $\mathbb{Q}[X]$ et le corolaire 3.152 nous dit alors que P est irréductible sur $\mathbb{Z}[X]$. □

Exemple 6.38.
Soit le polynôme $P = 3X^4 + 15X^2 + 10$. Pour appliquer 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. On aura vite vu que $p = 5$ fait l'affaire. Le polynôme P est donc irréductible sur $\mathbb{Q}[X]$. △

6.3.2 Idéaux

Soit $P \in \mathbb{K}[X]$ un polynôme. Nous notons (P) l'idéal engendré par P :

$$(P) = \{PR \text{ tel que } R \in \mathbb{K}[X]\}. \tag{6.81}$$

Lemme 6.39.
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 6.40.
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 3.132 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 3.129 de la division euclidienne[13], 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 6.40 que nous verrons plus bas est que tout polynôme annulateur d'un endomorphisme est divisé par le polynôme minimal (proposition 9.95).

Corolaire 6.41.
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 6.40, par conséquent la proposition 3.83(2) déduit que $\mathbb{K}[X]/(P)$ est un corps.

Une base de $\mathbb{L}$ est donnée par les projections de $1, X, X^2, \dots, X^{n-1}$. En effet ces éléments 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

$$\sum_{k=0}^{n-1} a_k X^k = SP, \tag{6.82}$$

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

$$\bar{X}^{n+l} = \bar{X}^n \bar{X}^l = (\bar{P} - \bar{Q})\bar{X}^l = \bar{Q}\bar{X}^l. \tag{6.83}$$

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. □

6.42.
Ce corolaire prendra une nouvelle jeunesse lorsque nous parlerons de polynômes d'endomorphismes, en particulier la proposition 9.105 va donner des précisions.

Lemme 6.43 ([141])**.**
Soit un isomorphisme de corps $\tau\colon \mathbb{K} \to \mathbb{K}'$. Alors

(1) L'application étendue

$$\begin{aligned} \tau\colon \mathbb{K}[X] &\to \mathbb{K}'[X] \\ \sum_i a_i X^i &\mapsto \sum_i \tau(a_i) X^i \end{aligned} \tag{6.84}$$

est un isomorphisme d'anneaux ;

(2) pour tout $P \in \mathbb{K}[X]$, le passage au quotient

$$\begin{aligned} \phi_\tau\colon \mathbb{K}[X]/(P) &\to \mathbb{K}'[X]/\big(\tau(P)\big) \\ \bar{Q} &\mapsto \overline{\tau(Q)} \end{aligned} \tag{6.85}$$

est un isomorphisme d'anneaux (et d'abord est bien définie).

13. Ici $\mathbb{K}$ est un corps et donc l'hypothèse d'inversibilité est automatiquement vérifiée.

Démonstration. Nous n'allons pas faire explicitement toutes les vérifications, mais tout de même les principales. Montrons que τ respecte le produit entre $\mathbb{K}[X]$ et $\mathbb{K}'[X]$. Nous rappelons que ce produit est défini par la formule (1.435). En notant P_i les coefficients de P et Q_i ceux de Q et en remarquant que la définition de τ est essentiellement que $\tau(P)_i = \tau(P_i)$, nous avons :

$$\tau(PQ) = \tau\Big(\sum_k \big(\sum_{l=0}^{k} P_l Q_{k-l}\big) X^k\Big) \tag{6.86a}$$
$$= \sum_k X^k \sum_{l=0}^{k} \tau(P_l Q_{k-l}) \tag{6.86b}$$
$$= \sum_k X^k \sum_{l=0}^{k} \tau(P_l)\tau(Q_{k-l}) \tag{6.86c}$$
$$= \sum_k X^k \sum_{l=0}^{k} \tau(P)_l \tau(Q)_{k-l} \tag{6.86d}$$
$$= \sum_i \big(\tau(P)_i X^i\big) \sum_j \big(\tau(Q)_j X^j\big) \tag{6.86e}$$
$$= \tau(P)\tau(Q). \tag{6.86f}$$

Passons à l'isomorphisme d'anneaux donné par ϕ_τ.

(i) **Bien définie** Si $\bar{Q}_1 = \bar{Q}_2$ alors $Q_2 = Q_1 + RP$ pour un certain $R \in \mathbb{K}[X]$. Dans ce cas,

$$\phi_\tau(Q_2) = \overline{\tau(Q_2)} = \overline{\tau(Q_1) + \tau(RP)} \tag{6.87a}$$
$$= \overline{\tau(Q_1) + \tau(R)\tau(P)} \tag{6.87b}$$
$$= \overline{\tau(Q_1)}. \tag{6.87c}$$

L'application ϕ_τ est donc bien définie.

(ii) **Injection** Si $\phi_\tau(\bar{Q}_1) = \phi_\tau(\bar{Q}_2)$ alors $\overline{\tau(Q_1)} = \overline{\tau(Q_2)}$, ce qui signifie que

$$\tau(Q_1) = \tau(Q_2) + R\ \tau(P) \tag{6.88}$$

pour un certain $R \in \mathbb{K}'[X]$. Puisque $\tau\colon \mathbb{K}[X] \to \mathbb{K}'[X]$ est un isomorphisme, nous pouvons y appliquer τ^{-1} pour trouver :

$$Q_1 = Q_2 + \tau^{-1}(R)P, \tag{6.89}$$

ce qui signifie que $\bar{Q}_1 = \bar{Q}_2$.

(iii) **Surjection** Un élément de $\mathbb{K}'[X]/\big(\tau(P)\big)$ est de la forme $\bar{Q}$ avec $Q \in \mathbb{K}'[X]$. C'est l'image par ϕ_τ de l'élément $\overline{\tau^{-1}(Q)} \in \mathbb{K}[X]/(P)$.

(iv) **Morphisme** Nous vous laissons vérifier que l'application ϕ_τ est un morphisme d'anneaux.

□

6.3.3 Identité de Bézout

Théorème 6.44 (Bézout)**.**
Les polynômes $P_1, \dots, P_n$ dans $\mathbb{K}[X]$ sont étrangers entre eux si et seulement si il existe des polynômes $Q_1, \dots, Q_n \in \mathbb{K}[X]$ tels que

$$P_1 Q_1 + \cdots + P_n Q_n = 1. \tag{6.90}$$

Deux polynômes P et Q ne sont donc pas premiers entre eux si il existe des polynômes x et y tels que l'identité de Bézout soit vérifiée :

$$xP + yQ = 0; \tag{6.91}$$

cette dernière pourra être écrite en termes de la matrice de Sylvester, voir sous-section 9.1.7.

Lemme 6.45.
Soient $(P_i)_{i=1,\dots,n} \in \mathbb{K}[X]$ *des polynômes étrangers deux à deux. Alors les polynômes*

$$Q_i = \prod_{j\neq i} P_j \tag{6.92}$$

sont étrangers entre eux[14].

Lemme 6.46 ([142])**.**
Soit $\mathbb{K}$ *un corps commutatif et* $\mathbb{A} \subset \mathbb{K}$ *un sous anneau de* $\mathbb{K}$*. Alors* $\mathbb{A}[X]$*, vu comme idéal de* $\mathbb{K}[X]$*, est un idéal premier.*

En d'autres termes, si $\phi \in \mathbb{K}[X]$*, et si il existe* $Q \in \mathbb{K}[X]$ *unitaire tel que* $\phi Q \in \mathbb{A}[X]$*, alors* $\phi \in \mathbb{A}[X]$*.*

6.3.4 Lemme et théorème de Gauss

Théorème 6.47 (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[15] 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

$$R = PUR + QVR = PUR + VPS = P(UR + VS), \tag{6.93}$$

ce qui signifie que P divise R. □

Le lemme suivant est une généralisation du lemme de Gauss dans $\mathbb{Z}$ (lemme 3.93).

Lemme 6.48 (Lemme de Gauss[101])**.**
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

$$\pi_p \colon \mathbb{Z}[X] \to (\mathbb{Z}/p\mathbb{Z})[X]. \tag{6.94}$$

Par définition $abPQ = P_1Q_1 \in \mathbb{Z}[X]$; en prenant la projection,

$$\pi_p(P_1)\pi_p(Q_1) = \pi_p(P_1Q_1) = \pi_P(ab)\pi_p(PQ) = 0 \tag{6.95}$$

parce que $\pi_p(ab) = 0$. Étant donné que $(\mathbb{Z}/p\mathbb{Z})[X]$ est intègre (théorème 3.126), 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. □

14. Et non seulement deux à deux.
15. théorème 6.44.

6.3.5 Polynômes sur un corps et pgcd

Nous savons qu'un corps est un anneau intègre (lemme 1.242). De plus l'ensemble des polynômes sur un anneau intègre est lui-même un anneau intègre (théorème 3.126). Donc la notion de pgcd à utiliser dans le cas de $\mathbb{K}[X]$ est celle de la définition 1.215.

Lemme 6.49 (Unicité du pgcd à inversibles près)**.**
Soit un corps commutatif $\mathbb{K}$ *et* $S \subset \mathbb{K}[X]$*. Si* δ_1 *et* δ_2 *sont des pgcd*[16] *de* S*, alors* $\delta_1 = k\delta_2$ *avec* $k \in \mathbb{K}$*.*

Démonstration. Nous savons que δ_1 est un pgcd de S, mais que δ_2 divise S. Donc $\delta_2 \mid \delta_1$. De la même manière, $\delta_1 \mid \delta_2$. Il existe donc $A, B \in \mathbb{K}[X]$ tels que $\delta_1 = A\delta_2$ et $\delta_2 = B\delta_1$. En substituant,

$$\delta_1 = AB\delta_1. \tag{6.96}$$

Mais $\mathbb{K}[X]$ possède la propriété de simplification par la proposition 1.221(3). Donc $AB = 1$. Cela signifie entre autres que A et B sont des inversibles de $\mathbb{K}[X]$.

Or les seuls inversibles dans $\mathbb{K}[X]$ sont les éléments de $\mathbb{K}$; si vous en doutez, pensez que le degré de AB est supérieur ou égal à celui de A. □

6.50.
En général, lorsque nous dirons « le » pgcd d'un ensemble, nous parlerons du pgcd unitaire, qui existe et est bien défini par le lemme 6.49.

Lemme 6.51 ([143])**.**
Soit un corps commutatif $\mathbb{K}$*, deux polynômes quelconques* $A, B \in \mathbb{K}[X]$ *et un polynôme unitaire* G*.*

Nous avons $G = \operatorname{pgcd}(A, B)$ *si et seulement si les deux conditions suivantes sont satisfaites :*

(1) Il existe $U, V \in \mathbb{K}[X]$ *tels que* $AU + BV = G$*,*

(2) G *divise* A *et* B*.*

Démonstration. Une implication dans chaque sens.

(i) $\Rightarrow$ Si G est le pgcd de A et B, il est clair que $G|A$ et $G|B$. Il reste donc à montrer l'existence des polynômes U et V vérifiant $AU + BV = G$. Puisque G divise A et B, il existe des polynômes A_1, B_1 tels que $A = GA_1$ et $B = GB_1$.

Nous montrons que les polynômes A_1 et B_1 sont premiers entre eux. S'ils ont un diviseur commun D, alors GD est un diviseur commun à A et B. Or, G est le pgcd de A et B donc $GD|G$; D ne peut être qu'un polynôme constant (c'est-à-dire un élément de $\mathbb{K}$). Mais comme G est unitaire, le coefficient du terme de plus haut degré de GD doit être 1. Donc $D = 1$. L'élément 1 est l'unique diviseur commun de A_1 et B_1 ; donc A_1 et B_1 sont bien premiers entre eux.

D'après le théorème de Bézout 6.44, il existe donc U et V tels que $A_1U + B_1V = 1$. En multipliant par G, nous obtenons l'égalité voulue : $AU + BV = G$.

(ii) $\Leftarrow$ Si G vérifie les deux conditions, montrons que G est le pgcd de A et B. Nous savons déjà (par hypothèse) que G divise A et B, il reste à montrer que tous les diviseurs commun à A et B divisent aussi G. Soit donc D un diviseur commun à A et B : il existe A_1 et B_1 tels que $A = DA_1$ et $B = DB_1$. Nous savons que $G = AU + BV$ donc $G = D(A_1U + B_1V)$, et $D|G$. Par définition, G est bien le pgcd de A et B.

□

Notons qu'en supprimant la condition d'unitarité de G, le résultat tient presque : il suffit de remplacer partout « le pgcd » par « un pgcd ».

16. Définition 1.215.

Lemme 6.52 ([143]).
Soient deux polynômes A, B premiers entre eux. Si le polynôme P est divisible par A et par B alors P est divisible par AB.

Démonstration. Comme $A \mid P$, il existe $Q_1 \in \mathbb{K}[X]$ tel que $P = AQ_1$. Mais B divise $P = AQ_1$ alors que B est premier avec A ; donc d'après le théorème de Gauss 6.47 : $B|Q_1$.

Il existe donc $Q_2 \in \mathbb{K}[X]$ tel que $Q_1 = BQ_2$. On a donc $P = ABQ_2$: P est bien divisible par AB. □

Lemme 6.53 ([143]).
Quelques propriétés du PGCD[17] *dans les polynômes. Soient des polynômes $P, Q, R \in \mathbb{K}[X]$.*

(1) Nous avons l'égalité[18]

$$\operatorname{pgcd}(P, PQ + R) = \operatorname{pgcd}(P, R). \tag{6.97}$$

(2) Si Q et R sont premiers entre eux,

$$\operatorname{pgcd}(P, QR) = \operatorname{pgcd}(P, Q)\operatorname{pgcd}(P, R) \tag{6.98}$$

(3) Si P et Q sont premiers entre eux,

$$\operatorname{pgcd}(P, QR) = \operatorname{pgcd}(P, R) \tag{6.99}$$

Démonstration. Dans la suite si A et B sont des polynômes, nous dirons « les diviseurs de $\{A, B\}$ » pour parler des diviseurs communs de A et B.

(1) Nous montrons que $\{P, PQ + R\}$ a les mêmes diviseurs que $\{P, R\}$.
D'une part, si $A \mid \{P, PQ + R\}$, alors il existe des polynômes B_1 et B_2 tels que $P = AB_1$ et $PQ + R = AB_2$. Donc

$$R = AB_2 - PQ = AB_2 - AB_1Q = A(B_2 - B_1Q), \tag{6.100}$$

et nous concluons que A divise R.
D'autre part, si $A \mid \{P, R\}$ alors il existe des polynômes B_1 et B_2 tels que $P = AB_1$ et $R = AB_2$. Donc

$$PQ + R = AB_1Q + AB_2 = A(B_1Q + B_2), \tag{6.101}$$

et A divise $PQ + R$.
Conclusion : les paires $\{P, PQ + R\}$ et $\{P, R\}$ ont même ensemble de diviseurs, et donc même pgcd.

(2) Nous avons trois polynômes P, Q, R et nous savons que Q et R sont premiers entre eux. Nous notons : $G_1 = \operatorname{pgcd}(P, Q)$ et $G_2 = \operatorname{pgcd}(P, R)$. Il faut montrer que G_1G_2 est le pgcd de P et QR ; pour cela nous allons utiliser le lemme 6.51.

(i) **$\exists U, V$ tels que $G_1G_2 = PU + QRV$** Puisque $G_1 = \operatorname{pgcd}(P, Q)$, il existe U_1 et V_1 tels que $G_1 = PU_1 + QV_1$ (lemme 6.51). On a de même : $G_2 = PU_2 + RV_2$. En prenant le produit :

$$G_1G_2 = (PU_1 + QV_1)(PU_2 + RV_2) = P(PU_1U_2 + RU_1V_2 + QV_1V_2) + QR(V_1V_2). \tag{6.102}$$

Donc c'est bon pour ce point.

(ii) **G_1 et G_2 sont premiers entre eux** Si D est un diviseur commun à G_1 et G_2, alors D divise Q et R qui sont premiers entre eux ; D ne peut être qu'un polynôme constant. Tous les diviseurs communs de G_1 et G_2 sont dans $\mathbb{K}$. Mais le pgcd est par définition un diviseur commun unitaire, donc $\operatorname{pgcd}(G_1, G_2) = 1$. Cela signifie que G_1 et G_2 sont premiers entre eux (définition 1.227).

17. Définition 1.215.
18. Notez l'analogie avec le lemme 3.6.

(iii) $\underline{G_1G_2 \mid QR}$ En effet : $G_1|Q$ et $G_2|R$ donc $G_1G_2|QR$.

(iv) $\underline{G_1G_2 \mid P}$ Le polynôme P est divisible par G_1 et par G_2, et de plus G_1 et G_2 sont premiers entre eux. Donc le lemme 6.52 conclut que P est divisible par G_1G_2.

(3) Supposons d'abord que $A \in \mathbb{K}[X]$ divise P et QR. Le théorème de Bézout 6.44 assure l'existence de polynômes U et V tels que $PU+QV=1$. Ensuite l'hypothèse de division nous donne des polynômes B_1 et B_2 tels que $P = AB_1$ et $QR = AB_2$. Nous avons :

$$1 = PU + QV = AB_1U + QV. \tag{6.103}$$

Cela prouve que A est premier avec Q grâce encore à Bézout, mais dans l'autre sens. Donc A est premier avec Q et $A \mid QR$. Donc $A|R$ par le théorème de Gauss 6.47.

Dans l'autre sens, si $A|R$ alors on a évidemment : $A|QR$.

Les diviseurs de $\{P, QR\}$ sont exactement les diviseurs de $\{P, R\}$. En conséquence, nous concluons que les paires $\{P, QR\}$ et $\{P, R\}$ ont le même pgcd.

□

6.4 Extension de corps

Lemme 6.54.
Soit $\mathbb{L}$ un corps[19] *fini et $\mathbb{K}$ un sous-corps de $\mathbb{L}$. Alors il existe $s \in \mathbb{N}$ tel que*

$$\mathrm{Card}(\mathbb{L}) = \mathrm{Card}(\mathbb{K})^s. \tag{6.104}$$

Démonstration. Le corps $\mathbb{L}$ est un $\mathbb{K}$-espace vectoriel de dimension finie. Si s est la dimension alors nous avons la formule (6.104) parce que chaque élément de $\mathbb{L}$ est un s-uple d'éléments de $\mathbb{K}$. □

Définition 6.55 ([144]).
Soit $\mathbb{K}$ un corps commutatif. Une ***extension*** *de $\mathbb{K}$ est un couple $(\mathbb{L}, j)$ où $\mathbb{L}$ est un corps et $j\colon \mathbb{K} \to \mathbb{L}$ est un morphisme de corps.*

Nous identifions le plus souvent $\mathbb{K}$ avec $j(\mathbb{K}) \subset \mathbb{L}$, mais il faut savoir que le corps $\mathbb{L}$ étendant $\mathbb{K}$ n'est pas toujours un sur-corps de $\mathbb{K}$. En particulier, l'ensemble $\mathbb{L}$ peut ne pas être une extension de l'ensemble $\mathbb{K}$.

Lemme-Définition 6.56.
Si $(\mathbb{L}, i)$ est une extension de $\mathbb{K}$, alors $\mathbb{L}$ devient un espace vectoriel sur $\mathbb{K}$ si nous posons

$$\lambda \cdot x = i(\lambda)x \tag{6.105}$$

pour tout $\lambda \in \mathbb{K}$ et $x \in \mathbb{L}$. La multiplication du membre de droite est celle du corps $\mathbb{L}$.

Définition 6.57.
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 6.58.
L'ensemble $\mathbb{C}$ est une extension de $\mathbb{R}$ et son degré est $[\mathbb{C} : \mathbb{R}] = 2$. △

Proposition 6.59 (Composition des degrés[145]).
Si $\mathbb{L}_2$ est une extension de $\mathbb{L}_1$ qui est elle-même une extension de $\mathbb{K}$, alors $\mathbb{L}_2$ est une extension de $\mathbb{K}$ et on a :

$$[\mathbb{L}_2 : \mathbb{K}] = [\mathbb{L}_2 : \mathbb{L}_1][\mathbb{L}_1 : \mathbb{K}]. \tag{6.106}$$

Dans ce cas, si $\{v_i\}_{i\in I}$ est une $\mathbb{K}$-base de $\mathbb{L}_1$ et si $\{w_\alpha\}_{\alpha\in A}$ est une $\mathbb{L}_1$-base de $\mathbb{L}_2$ alors $\{v_iw_\alpha\}_{\substack{i\in I\\ \alpha\in A}}$ est une $\mathbb{K}$-base de $\mathbb{L}_2$.

19. Définition 1.237.

Notons que la formule (6.106) n'est pas très instructive dans le cas des extensions non finies. La seconde partie, sur les bases, est en réalité nettement plus intéressante.

Démonstration. Soit $a \in \mathbb{L}_2$. Puisque les w_α forment une $\mathbb{L}_2$-base nous avons une décomposition

$$a = \sum_\alpha a_\alpha w_\alpha \tag{6.107}$$

pour des éléments $a_\alpha \in \mathbb{L}_1$. Mais les v_i forment une $\mathbb{K}$-base de $\mathbb{L}_1$, donc chacun des a_α peut être décomposé comme $a_\alpha = \sum_i a_{\alpha i} v_i w_\alpha$. Donc :

$$a = \sum_{\alpha i} a_{\alpha i} v_i w_\alpha, \tag{6.108}$$

qui donne une décomposition de a en éléments de $\{v_i w_\alpha\}$ à coefficients dans $\mathbb{K}$. La partie proposée est donc génératrice.

Pour prouver qu'elle est également libre, nous supposons avoir des éléments $a_{\alpha i} \in \mathbb{K}$ tels que

$$\sum_{\alpha i} a_{\alpha i} v_i w_\alpha = 0. \tag{6.109}$$

En récrivant sous la forme

$$\sum_\alpha \Big(\sum_i a_{\alpha i} v_i \Big) w_\alpha = 0, \tag{6.110}$$

nous reconnaissons une combinaison linéaire nulle des w_α à coefficients dans $\mathbb{L}_1$. Les coefficients sont donc nuls : $\sum_i a_{\alpha i} v_i = 0$. C'est une combinaison linéaire nulle des v_i à coefficients dans $\mathbb{K}$. Comme les v_i forment une base, les coefficients sont nuls : $a_{\alpha i} = 0$. □

6.4.1 Extension et polynôme minimal

Lemme-Définition 6.60 (Polynôme minimal)**.**
Soit $\mathbb{L}$ une extension de $\mathbb{K}$ et $a \in \mathbb{L}$. Nous considérons la partie

$$I_a = \{P \in \mathbb{K}[X] \text{ tel que } P(a) = 0\} \tag{6.111}$$

que nous supposons non réduite à $\{0\}$ [20]

(1) La partie I_a est un idéal dans $\mathbb{K}[X]$,

(2) la partie I_a est un idéal principal dans $\mathbb{K}[X]$,

(3) l'idéal I_a possède un unique générateur unitaire.

*Cet unique générateur unitaire est le **polynôme minimal** de a sur $\mathbb{K}$.*

Démonstration. En plusieurs parties.

(i) **Pour (1)** Soit $P \in I_a$: $P(a) = 0$. Si $Q \in \mathbb{K}[X]$ alors la proposition 1.307 nous indique que

$$(PQ)(a) = P(a)Q(a) = 0. \tag{6.112}$$

Donc $PQ \in I_a$. Comme de plus I_a est clairement vectoriel, I_a est un idéal.
Notez que nous avons utilisé la règle du produit nul justifiée par le fait que $\mathbb{K}$ soit un corps [21] et donc soumis au point (3) de la proposition 1.221.

(ii) **Pour (2)** Nous savons par le théorème 6.40 que $\mathbb{K}[X]$ est un anneau principal. En particulier, tous ses idéaux sont principaux, c'est dans la définition 3.75 d'un anneau principal.

20. La non trivialité de I_a est une vraie hypothèse. En effet si nous prenons $\mathbb{K} = \mathbb{Q}$ et l'extension $\mathbb{L} = \mathbb{R}$, alors il suffit de prendre un réel a non algébrique sur $\mathbb{Q}$ pour que I_a soit réduit au seul polynôme identiquement nul.

21. Si vous connaissez un contre-exemple à cette proposition dans le cas où $\mathbb{K}$ serait remplacé par un anneau, écrivez-moi.

(iii) **Pour (3)** Le théorème 6.40(3) nous informe alors que I_a possède un unique générateur unitaire.

□

Si nous avons un corps et un élément dans une extension du corps, il n'est pas autorisé de dire « soit le polynôme minimal de cet élément dans le premier corps » parce qu'il n'existe peut-être pas de polynôme annulateur.

6.61.
Dans le cas des opérateurs sur un espace de dimension finie (par exemple les matrices), il existe toujours un polynôme minimal, comme nous le verrons dans le lemme 9.90.

Exemple 6.62.
Le polynôme minimal dépend 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 6.63 ([1]).
Soit $\mathbb{L}$ une 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[22] *sur $\mathbb{K}$;*

(2) Le polynôme μ_a est premier[23] *avec tout polynôme de $\mathbb{K}[X]$ non annulateur de a.*

Démonstration. Une chose à la fois.

(1) D'abord le polynôme μ_a n'est pas inversible parce que seuls les éléments de $\mathbb{K}$ (ceux de degré zéro) peuvent être inversibles[24]. Mais ces polynômes sont constants et ne peuvent donc pas être des polynômes annulateurs de quoi que ce soit.

Ensuite, supposons la décomposition $\mu_a = PQ$ avec $P, Q \in \mathbb{K}[X]$. En évaluant cette égalité en a nous avons

$$0 = P(a)Q(a). \tag{6.113}$$

Puisque nous sommes sur un corps, nous avons la règle du produit nul[25] et nous déduisons que soit $P(a)$ soit $Q(a)$ est nul, ou les deux. Pour fixer les idées, nous supposons $P(a) = 0$.

Dans ce cas, P fait partie de l'idéal annulateur de a, lequel idéal est engendré par μ_a. Donc il existe $S \in \mathbb{K}[X]$ tel que $P = S\mu_a$. En récrivant $\mu_a = PQ$ avec cela nous avons :

$$\mu_a = S\mu_a Q \tag{6.114}$$

ou encore : $SQ = 1$, ce qui signifie que S et Q sont dans $\mathbb{K}$ et inversibles.

Nous concluons que μ_a ne peut pas être écrit sous forme de produit de deux non inversibles.

(2) Soit Q un polynôme non annulateur de a. Soit aussi un diviseur commun P de Q et μ_a dans $\mathbb{K}[X]$. Nous devons prouver que P est un inversible, c'est-à-dire un élément de $\mathbb{K}$ (le fait que P ne soit pas le polynôme nul est évident). Nous avons $\mu_a = PR_1$ et $Q = PR_2$ pour certains polynômes $R_1, R_2 \in \mathbb{K}[X]$. Puisque μ_a est irréductible par (1), il n'est pas produit de deux non inversibles. En d'autres termes, soit P soit R_1 est inversible. Si P n'est pas inversible, alors R_1 est inversible ; disons $R_1 = k \in \mathbb{K}$. Alors

$$0 = \mu_a(a) = P(a)k, \tag{6.115}$$

donc $P(a) = 0$. Mais alors

$$Q(a) = P(a)R_2(a) = 0, \tag{6.116}$$

ce qui est contraire à l'hypothèse selon laquelle Q n'était pas annulateur de a.

Nous retenons donc que P est inversible, ce qu'il fallait montrer.

22. Définition 6.34.
23. Définition 3.133.
24. Et d'ailleurs, le sont, mais ce n'est pas important ici.
25. Parce qu'un corps est un anneau intègre par le lemme 1.242 et qu'un anneau intègre est justement un anneau sur lequel nous avons la règle du produit nul, voir la définition 1.221.

□

Définition 6.64.
*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 $\mathbb{Q}$.*

6.4.2 Extensions algébriques et éléments transcendants

6.4.2.1 Éléments algébriques et transcendants

Définition 6.65.
L'ensemble $A[X]$ devient un $\mathbb{K}$-espace vectoriel avec la définition

$$(\lambda P)_k = \lambda P_k. \tag{6.117}$$

Voici une définition d'un élément algébrique sur un corps. Une caractérisation plus « pratique » sera donnée dans le lemme 6.68.

Lemme-Définition 6.66 (Élément algébrique et transcendant[146])**.**
Soit une extension $\mathbb{L}$ de $\mathbb{K}$ et $\alpha \in \mathbb{L}$. Nous considérons l'application

$$\begin{aligned} \varphi\colon \mathbb{K}[X] &\to \mathbb{L} \\ P &\mapsto P(\alpha). \end{aligned} \tag{6.118}$$

Alors

(1) L'application φ est un morphisme d'anneaux[26]*.*

(2) L'application φ est un morphisme de $\mathbb{K}$-espace vectoriel.

*Si φ est injective, nous disons que α est **transcendant**. Sinon, nous disons qu'il est **algébrique**.*

Démonstration. Le fait que φ soit un morphisme d'anneaux est le lemme 1.307 déjà prouvé.

Pour le morphisme de $\mathbb{K}$-espace vectoriel, il faut seulement ajouter le calcul

$$\varphi(\lambda P) = (\lambda P)(\alpha) = \lambda P(\alpha) = \lambda\varphi(P). \tag{6.119}$$

Notons la justification suivante qui n'est pas tout à fait triviale :

$$(\lambda P)(\alpha) = \sum_k (\lambda P)_k \alpha^k = \sum_k \lambda P_k \alpha^k = \lambda P(\alpha) \tag{6.120}$$

qui utilise la définition 6.65. □

Exemple 6.67.
L'injectivité de φ n'est pas automatique. Prenons par exemple $\mathbb{L} = \mathbb{Q}[\sqrt{2}]$ dans $\mathbb{R}$. Les polynômes dans $\mathbb{Q}[X]$ ont des degrés arbitrairement élevés en X, tandis que les éléments de $\mathbb{L}$ n'ont pas de degré très élevés en $\sqrt{2}$ parce que $\sqrt{2}\sqrt{2} = 2$. L'ensemble $\mathbb{Q}[\sqrt{2}]$ ne contient donc que des éléments de la forme $a + b\sqrt{2}$ avec $a, b \in \mathbb{Q}$.

Si par contre $x_0 \in \mathbb{R}$ n'est racine d'aucun polynôme (cela existe parce que $\mathbb{R}$ n'est pas dénombrable), alors $\mathbb{Q}[x_0]$ contient tous les $\sum_{k=0}^{N} a_k x_0^k$ avec N arbitrairement grand. Et tous ces nombres sont différents. △

Le lemme suivant donne une caractérisation d'élément algébrique moins abstraite que la définition 6.66.

Lemme 6.68.
Soit $\mathbb{K}$, un corps et $\mathbb{L}$, une extension de $\mathbb{K}$. Un élément $\alpha \in \mathbb{L}$ est algébrique sur $\mathbb{K}$ si et seulement si existe un polynôme non nul $P \in \mathbb{K}[X]$ tel que $P(\alpha) = 0$.

26. Définition 1.39.

Démonstration. Nous considérons l'application φ de la définition 6.66. Si φ n'est pas injective, c'est qu'il existe un polynôme P dans $\mathbb{K}[X]$ tel que $\varphi(P) = 0$. Dans ce cas, $P(\alpha) = 0$.

À l'inverse si il existe P non nul dans $\mathbb{K}[X]$ tel que $P(\alpha) = 0$, alors $\varphi(P) = 0$ et φ n'est pas injective. □

Définition 6.69.
*Un corps $\mathbb{K}$ est **algébriquement clos** si tout polynôme non constant à coefficients dans $\mathbb{K}$ contient au moins une racine dans $\mathbb{K}$.*

Nous verrons dans le théorème de d'Alembert 12.87 que $\mathbb{C}$ est un corps algébriquement clos.

Définition 6.70 (Extension algébrique, clôture algébrique)**.**
Soient un corps $\mathbb{K}$ et une extension $\alpha\colon \mathbb{K} \to \mathbb{L}$.

*(1) L'extension $\mathbb{L}$ est une extension **algébrique** de $\mathbb{K}$ si tous ses éléments sont algébriques*[27] *sur $\mathbb{K}$, c'est-à-dire sont racines de polynômes à coefficients dans $\alpha(\mathbb{K})$, voir le lemme 6.68.*

*(2) L'extension $\mathbb{L}$ est **algébriquement close** si le corps $\mathbb{L}$ est algébriquement clos (définition 6.69).*

*(3) L'extension $\mathbb{L}$ est une **clôture algébrique** du corps $\mathbb{K}$ si elle est une extension algébrique qui est algébriquement close.*

6.71.
Donc une extension est algébrique si elle contient seulement des racines de polynômes ; elle est close si elle contient au moins une racine de chaque polynôme. L'extension est une clôture algébrique si elle est les deux en même temps.

Exemple 6.72.
Le corps $\mathbb{R}$ 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. Voir aussi la proposition 6.124. △

Lemme 6.73.
Un corps est algébriquement clos si et seulement si tous ses polynômes sont scindés[28].

Démonstration. Si tout polynôme est scindé, tout polynôme possède des racines ; c'est l'autre sens qui est plus consistant.

Soit un corps algébriquement clos $\mathbb{K}$. Nous allons effectuer une récurrence sur le degré des polynômes. Si P est un polynôme de degré 1, alors il est scindé.

Nous supposons que tous les polynômes de degré $n-1$ sont scindés. Soit un polynôme P de degré n. Le corps étant algébriquement clos, le polynôme P a une racine que nous notons $a_n \in \mathbb{K}$. La proposition 3.145 nous explique qu'il existe un polynôme Q de degré $n-1$ tel que $P = (X - a_n)Q$.

Par hypothèse de récurrence, le polynôme Q est scindé : il existe $\{a_i\}_{i=1,\dots,n-1}$ dans $\mathbb{K}$ tels que $Q = \prod_{i=1}^{n-1}(X - a_i)$. Au final,

$$P = (X - a_n)Q = \prod_{k=1}^{n}(X - a_k) \tag{6.121}$$

et P est scindé. □

Lemme 6.74.
Soient un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$. Nous supposons que P est scindé :

$$P = \prod_{k=1}^{n}(X - a_k). \tag{6.122}$$

Si α est une racine de P, alors α est l'un des a_k.

27. Définition 6.66.
28. Définition 6.36

Démonstration. Dire que α est une racine de P revient à dire que

$$\prod_{k=1}^{n}(\alpha - a_k) = 0 \tag{6.123}$$

Un corps est toujours un anneau intègre (lemme 1.242), c'est-à-dire que la règle du produit nul est utilisable. Dans notre cas, le produit nul (6.123) nous indique que $\alpha - a_k = 0$ pour (au moins) un des k. Donc effectivement α est l'un des a_k. □

Lemme 6.75.
Soient un corps algébriquement clos $\mathbb{K}$ ainsi qu'une extension algébrique $\alpha\colon \mathbb{K} \to \mathbb{L}$. Alors $\alpha(\mathbb{K}) = \mathbb{L}$.

Démonstration. Nous allons montrer que tous les éléments de $\mathbb{L}$ sont dans l'image de α. Soit donc $l \in \mathbb{L}$. Puisque l'extension $\alpha\colon \mathbb{K} \to \mathbb{L}$ est une extension algébrique, il existe un polynôme $P \in \alpha(\mathbb{K})[X]$ tel que $P(l) = 0$.

Étant donné que α est injective, il est possible de considérer le polynôme $Q = \alpha^{-1}(P)$, c'est-à-dire que, si $P = \sum_k a_k X^k$, nous posons $Q = \sum_k \alpha^{-1}(a_k)X^k$.

Le corps $\mathbb{K}$ étant algébriquement clos, le polynôme Q est scindé (proposition 6.73) :

$$Q = \prod_{k=1}^{n}(X - b_k) \tag{6.124}$$

avec $b_k \in \mathbb{K}$. Nous avons alors aussi la factorisation

$$P = \prod_{k=1}^{n}(X - \alpha(b_k)) \tag{6.125}$$

dans $\mathbb{L}[X]$. Nous avons vu que l était une racine de P. Donc l est un des $\alpha(b_k)$ (lemme 6.74). Cela prouve que $l \in \alpha(\mathbb{K})$. □

6.4.3 Extension algébrique et polynôme minimal

Proposition 6.76 ([78]).
Soit une extension algébrique[29] *$\mathbb{L}$ du corps $\mathbb{K}$.*

(1) Pour tout $a \in \mathbb{L}$, il existe un polynôme $P \in \mathbb{K}[X]$ tel que $P(a) = 0$.

(2) Le polynôme minimal de a dans $\mathbb{K}[X]$ est l'unique polynôme unitaire irréductible annulant a.

Démonstration. Le premier point est seulement la définition 6.70 d'une extension algébrique.

L'idéal annulateur $I_a = \{P \in \mathbb{K}[X] \text{ tel que } P(a) = 0\}$ n'est pas réduit à $\{0\}$ parce que $\mathbb{L}$ est une extension algébrique. L'existence du polynôme minimal est le lemme 6.60 et le fait qu'il soit irréductible est la proposition 6.63(1).

Ce qui nous intéresse ici est l'unicité. Soit $\mu_1 \in \mathbb{K}[X]$, un polynôme annulateur de a irréductible et unitaire. Puisque $\mu_1 \in I_a$ et que par définition, $I_a = (\mu)$, il existe $P \in \mathbb{K}[X]$ tel que $\mu_1 = P\mu$. Comme μ n'est pas inversible et que μ_1 est irréductible, P doit être inversible : $\mu_1 = k\mu$ pour un certain $k \in \mathbb{K}$.

Puisque μ et μ_1 sont unitaires, $k = 1$. Donc $\mu_1 = \mu$. □

Lemme 6.77.
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ de $\mathbb{K}$ et $\alpha \in \mathbb{L}$, un élément algébrique[30] *sur $\mathbb{K}$. Si μ est le polynôme minimal de α sur $\mathbb{K}$ alors*

$$\begin{aligned} \varphi\colon \mathbb{K}[\alpha] &\to \mathbb{K}[X]/(\mu) \\ Q(\alpha) &\mapsto \bar{Q} \end{aligned} \tag{6.126}$$

29. Définition 6.70.
30. Définition 6.66.

avec $Q \in \mathbb{K}[X]$ est un isomorphisme de corps et de $\mathbb{K}$-espaces vectoriels.

Démonstration. D'abord, α est algébrique, donc l'idéal annulateur I_α n'est pas réduit à $\{0\}$, et l'existence d'un polynôme minimal est assurée par le lemme 6.60.

Ensuite, le fait que $\mathbb{K}[X]/(\mu)$ soit un corps est le corolaire 6.41. Nous montrons à présent que φ est un isomorphisme (d'anneaux) ; cela suffit pour en déduire que $\mathbb{K}[\alpha]$ est également un corps.

Ces préliminaires étant dits, nous commençons.

(i) **Bien définie** Nous devons prouver que φ est bien définie, c'est-à-dire que tout élément de $\mathbb{K}[\alpha]$ peut être écrit sous la forme $Q(\alpha)$ pour un $Q \in \mathbb{K}[X]$, et que si $Q_1(\alpha) = Q_2(\alpha)$ alors $\bar{Q}_1 = \bar{Q}_2$.

Le fait que tous les éléments de $\mathbb{K}[\alpha]$ peuvent être écrits sous la forme $Q(\alpha)$ est la proposition 3.125. Supposons que $Q_1(\alpha) = Q_2(\alpha)$. Alors nous définissons $R \in \mathbb{K}[X]$ par $Q_1 = Q_2 + R$, et en évaluant cette égalité en α nous avons

$$Q_1(\alpha) = Q_2(\alpha) + R(\alpha), \tag{6.127}$$

autrement dit $R(\alpha) = 0$. Donc R est dans l'idéal annulateur de α et est donc dans (μ), c'est-à-dire que dans le quotient $\mathbb{K}[X]/(\mu)$ nous avons $\bar{R} = 0$ et donc $\bar{Q}_1 = \bar{Q}_2$.

(ii) **Surjective** Tout élément de $\mathbb{K}[X]/(\mu)$ est de la forme $\bar{Q}$ pour un $Q \in \mathbb{K}[X]$. Or ces éléments sont ceux de l'ensemble d'arrivée de φ.

(iii) **Injective** Si $\bar{Q}_1 = \bar{Q}_2$, alors $Q_1 = Q_2 + R$ avec R dans l'idéal engendré par μ, c'est-à-dire entre autres $R(\alpha) = 0$. Donc $Q_1(\alpha) = Q_2(\alpha)$.

Nous devons encore montrer que nous avons là un morphisme de $\mathbb{K}$-espaces vectoriels.

(1) Si $k \in \mathbb{K}$ alors $\varphi\big(kQ(\alpha)\big) = \overline{kQ}$. Mais par définition de la structure d'espace vectoriel sur $\mathbb{K}[X]/(\mu)$, $\overline{kQ} = k\bar{Q}$ (vérifier que cette définition de la multiplication par un scalaire sur $\mathbb{K}[X]/(\mu)$ est correcte).

(2) Nous avons aussi $\varphi\big(Q_1(\alpha) + Q_2(\alpha)\big) = \varphi\big((Q_1 + Q_2)(\alpha)\big) = \overline{Q_1 + Q_2} = \bar{Q}_1 + \bar{Q}_2$.

□

6.4.4 Extensions et polynômes

Nous savons déjà depuis la définition 1.303 ce qu'est $A[X]$ pour tout anneau A et donc, à fortiori, pour un corps.

Définition 6.78.
Soit un corps commutatif[31]*. Nous notons $\mathbb{K}(X)$ le corps des fractions*[32] *de $\mathbb{K}[X]$.*

Lemme-Définition 6.79.
Si $R \in \mathbb{K}(X)$, avec $R = P/Q$ et si $\mathbb{L}$ est une extension[33] *de $\mathbb{K}$ contenant l'élément α, alors nous définissons*

$$R(\alpha) = P(\alpha)Q(\alpha)^{-1}. \tag{6.128}$$

Cela est une bonne définition au sens où elle ne dépend pas du choix du représentant (P, Q) pris dans la classe P/Q.

Démonstration. Supposons $R = P_1/Q_1 = P_2/Q_2$. Par définition des classes (définition 1.313) nous avons

$$P_1Q_2 = Q_1P_2. \tag{6.129}$$

Puisque l'évaluation est un morphisme $\mathbb{K}[X] \to \mathbb{K}$[34] nous pouvons évaluer l'équation (6.129) en α :

$$P_1(\alpha)Q_2(\alpha) = Q_1(\alpha)P_2(\alpha). \tag{6.130}$$

31. Sauf mention du contraire, tous les corps du Frido sont commutatifs.
32. Définition 1.313.
33. Définition 6.55.
34. Lemme 1.307. Certes ce lemme ne parle que d'anneaux, mais à y bien penser, dans le passage de (6.129) à (6.130), nous ne considérons que les structures d'anneaux sur $\mathbb{K}[X]$ et $\mathbb{K}$.

Cette dernière est une égalité dans le corps $\mathbb{K}$. Nous pouvons donc la multiplier par $Q_2(\alpha)^{-1}P_2(\alpha)^{-1}$ (et utiliser toutes les hypothèses de commutativité des anneaux et corps) pour obtenir

$$P_1(\alpha)Q_1(\alpha)^{-1} = P_2(\alpha)Q_2(\alpha)^{-1}, \tag{6.131}$$

c'est-à-dire

$$(P_1/Q_1)(\alpha) = (P_2/Q_2)(\alpha). \tag{6.132}$$

□

Proposition-Définition 6.80 ([1]).
Soient un corps $\mathbb{K}$, une extension $(\mathbb{L}, j_{\mathbb{L}})$ de $\mathbb{K}$ et un élément $\alpha \in \mathbb{L}$. Nous définissons $\mathbb{K}(\alpha)_{\mathbb{L}}$ comme étant l'intersection de tous les sous-corps de $\mathbb{L}$ contenant $j_{\mathbb{L}}(\mathbb{K})$ et α.

Alors

(1) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est un sous-corps de $\mathbb{L}$,

(2) $\mathbb{K}(\alpha)_{\mathbb{L}}$ est une extension[35] *de $\mathbb{K}$.*

Démonstration. Nous commençons par prouver que $\mathbb{K}(\alpha)_{\mathbb{L}}$ est bien un corps. Si $a, b \in \mathbb{K}(\alpha)_{\mathbb{L}}$ alors il suffit de calculer ab, $a+b$ et a^{-1} dans n'importe quel sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α ; nous avons une garantie que a, b, ab, $a+b$ et a^{-1} sont dans tous les tels sous-corps.

Pour prouver que $\mathbb{K}(\alpha)_{\mathbb{L}}$ est bien une extension, nous devons trouver un morphisme de corps $j\colon \mathbb{K} \to \mathbb{K}(\alpha)_{\mathbb{L}}$. On constate que prendre $j = j_{\mathbb{L}}$ fonctionne parce que par définition, $\mathbb{K}(\alpha)_{\mathbb{L}}$ est une partie de $\mathbb{L}$ contenant l'image de $j_{\mathbb{L}}$. □

Lemme 6.81.
Soit n tel que $\sqrt{n}$ ne soit pas un rationnel. Si $\alpha \in \{a + b\sqrt{n}\}_{a,b\in\mathbb{Q}}$, alors il existe un unique choix $(x, y) \in \mathbb{Q}^2$ tel que

$$\alpha = x + y\sqrt{n}. \tag{6.133}$$

Démonstration. L'existence est dans la définition de α. Il s'agit de voir l'unicité. Supposons $x + y\sqrt{n} = a + b\sqrt{n}$ avec $x, y, a, b \in \mathbb{Q}$. Si $b \neq y$ nous pouvons écrire

$$\sqrt{n} = \frac{x-a}{b-y}. \tag{6.134}$$

Comme $\sqrt{n}$ n'est pas un rationnel, une telle écriture est impossible. Donc $b = y$. Nous avons alors $x + y\sqrt{n} = a + y\sqrt{n}$ et du coup aussi $x = a$. □

Exemple 6.82.
Nous avons

$$\mathbb{Q}(\sqrt{2})_{\mathbb{R}} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}} \tag{6.135}$$

où à droite nous calculons les sommes et les produits dans $\mathbb{R}$. Le tout est un sous-ensemble de $\mathbb{R}$ qui se révèle être un corps contenant $\mathbb{Q}$ et $\sqrt{2}$.

En particulier, dans $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ nous avons $\sqrt{2}\sqrt{2} = 2$. △

Lemme 6.83.
Les corps $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ et $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ ne sont pas isomorphes.

Démonstration. Supposons l'existence d'un morphisme de corps[36]

$$\psi\colon \mathbb{Q}(\sqrt{2})_{\mathbb{R}} \to \mathbb{Q}(\sqrt{3})_{\mathbb{R}}. \tag{6.136}$$

Nous notons « 1 » à la fois le neutre de la multiplication dans $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ et $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ (qui s'avèrent être les mêmes en tant qu'élément de $\mathbb{R}$, mais ça n'a pas d'importance ici).

35. Définition 6.55.
36. Définition 1.39. Oui, c'est un bête morphisme d'anneaux. Il n'y a pas plus de structure dans un corps que dans un anneau.

Soit $\alpha \in \mathbb{Q}(\sqrt{2})_{\mathbb{R}}$ tel que $\alpha^2 - 1 = 0$. Alors nous avons aussi

$$\psi(\alpha)^2 - 1 = \psi(\alpha^2) - \psi(1) = \psi(\alpha^2 - 1) = \psi(0) = 0. \tag{6.137}$$

Donc $\psi(\alpha)$ est un élément de $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ qui est une racine de $X^2 - 1$.

Or un tel élément n'existe pas dans $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$ parce que nous savons que dans $\mathbb{R}$ entier, il n'y a que deux racines : $\pm\sqrt{2}$, et aucune des deux n'est dans $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$. □

Exemple 6.84.
Est-ce que $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend réellement de $\mathbb{L}$? Si $\mathbb{L}_2$ est une extension de $\mathbb{L}$ alors nous avons évidemment[37] $\mathbb{K}(\alpha)_{\mathbb{L}_2} = \mathbb{K}(\alpha)_{\mathbb{L}}$.

Nous commençons par construire un corps $\mathbb{K}$ un peu idiot qui, comme ensemble, est comme $\mathbb{Q}(\sqrt{2})_{\mathbb{R}}$, c'est-à-dire la partie

$$\{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}, \tag{6.138}$$

de $\mathbb{R}$.

Mais cette fois nous définissons la multiplication suivante :

$$(a + b\sqrt{2})(c + d\sqrt{2}) = ac + 3bd + (ad + bc)\sqrt{2}. \tag{6.139}$$

C'est un corps parce que tout élément non nul est inversible. En effet, l'équation

$$(a + b\sqrt{2})(x + y\sqrt{2}) = 1 \tag{6.140}$$

donne

$$\begin{pmatrix} a & 3b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{6.141}$$

Ce système a une unique solution si et seulement si $\det \begin{pmatrix} a & 3b \\ b & a \end{pmatrix} = 0$. Cela survient si et seulement si

$$a^2 - 3b^2 = 0. \tag{6.142}$$

Les solutions de cette équation dans $\mathbb{R}$ sont $a = \pm\sqrt{3}|b|$. Dès que a ou b est non nul, cela ne peut pas satisfaire $a, b \in \mathbb{Q}$. Donc le déterminant est toujours non nul et il existe $x, y \in \mathbb{Q}$ tels que (6.140) soit satisfaite.

Tout cela nous a donné un corps $\mathbb{K}$ dont $\mathbb{Q}$ est un sous-corps et qui contient l'élément $\sqrt{2}$ de $\mathbb{R}$. Il n'est cependant pas un sous-corps de $\mathbb{R}$.

Ce corps est isomorphe à $\mathbb{Q}(\sqrt{3})_{\mathbb{R}}$. En effet, nous montrons que

$$\begin{aligned} \psi\colon \mathbb{K} &\to \mathbb{Q}(\sqrt{3})_{\mathbb{R}} \\ a + b\sqrt{2} &\mapsto a + b\sqrt{3} \end{aligned} \tag{6.143}$$

est un isomorphisme de corps. Pour le produit, nous avons

$$\begin{aligned} \psi\big((a + b\sqrt{2})(c + d\sqrt{2})\big) &= \psi\big(ac + 3bd + (ad + bc)\sqrt{2}\big) && \text{(6.144a)} \\ &= ac + 3bd + (ad + bc)\sqrt{3} && \text{(6.144b)} \\ &= (a + b\sqrt{3})(c + d\sqrt{3}) && \text{(6.144c)} \\ &= \psi(a + b\sqrt{2})\psi(c + d\sqrt{2}). && \text{(6.144d)} \end{aligned}$$

Remarques :

— L'application ψ est bien définie grâce au lemme 6.81 couplé au théorème 3.36 appliqué à $n = 2$ et $n = 3$.

— Dans le membre de gauche de (6.144a), $b\sqrt{2}$ est un produit dans $\mathbb{R}$ (d'où l'importance du lemme 6.81 qui permet de re-séparer les éléments de $\mathbb{R}$ partie rationnelle et partie multiple de $\sqrt{2}$), et le produit entre $(a + b\sqrt{2})$ et $(c + d\sqrt{2})$ est un produit dans $\mathbb{K}$.

37. Vérifiez-le tout de même.

— Dans (6.144b) et (6.144c), tous les produits sont dans $\mathbb{R}$.

En comparant avec le lemme 6.83, nous avons alors

$$\mathbb{Q}(\sqrt{2})_{\mathbb{K}} = \mathbb{Q}(\sqrt{3})_{\mathbb{R}} \neq \mathbb{Q}(\sqrt{2})_{\mathbb{R}} \tag{6.145}$$

△

6.85.
Nous allons encore enfoncer le clou sur le fait que $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend de $\mathbb{L}$.

Le fait est que si on y pense, l'objet $\sqrt{2}$ n'a aucun rapport avec $\mathbb{Q}$. En effet les objets de $\mathbb{Q}$ sont des classes d'équivalence de couples d'éléments de $\mathbb{Z}$, alors que l'élément $\sqrt{2}$ est une classe d'équivalence de suites de Cauchy dans $\mathbb{Q}$.

Lorsque nous écrivons $\mathbb{Q}(\sqrt{2})$, nous associons des objets de nature complètement différentes, et il n'y a aucune raison à priori de définir la multiplication entre eux d'une façon plutôt qu'une autre.

Plus généralement, dans ZF (nous faisons semblant de suivre ZF tout en sachant que nous ne savons pas ce que c'est réellement[38]), tout est ensemble. Peut-on dire ce que serait $\mathbb{Q}(I)$ si I est un ensemble quelconque ? Attention : en écrivant $\mathbb{Q}(I)$, nous entendons un corps dont I est un élément, pas un corps qui contiendrait comme éléments tous les éléments de I.

Si I est juste un ensemble, quelle définition donner de I^2 ? Il y a plein de choix et rien ne se dégage clairement comme étant pertinent. Si par contre, en guise de I nous considérons l'ensemble $\sqrt{2}$ (oui, c'est un ensemble : un ensemble de suites de Cauchy dans $\mathbb{Q}$), alors tout de suite nous nous disons que la bonne façon de faire est $\sqrt{2}^2 = 2$. Ce réflexe est juste conditionné par le fait que nous connaissons déjà par ailleurs le corps $\mathbb{R}$. Rien de plus.

Donc oui, $\mathbb{K}(\alpha)_{\mathbb{L}}$ dépend de $\mathbb{L}$, mais dans les cas particuliers où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$, il y a une égalité implicite $\mathbb{L} = \mathbb{C}$. Cela étant dit, il n'y a plus d'ambiguïté en écrivant $\mathbb{Q}(\sqrt{2})$.

Définition 6.86.
Soit une extension[39] *de corps* $j\colon \mathbb{K} \to \mathbb{L}$. *Soit* $A \subset \mathbb{L}$.

(1) Nous notons $\mathbb{K}(A)_{\mathbb{L}}$ *le plus petit sous-corps de* $\mathbb{L}$ *contenant* $j(\mathbb{K})$ *et* A. *C'est l'intersection de tous les sous-corps de* $\mathbb{L}$ *contenant* A *et* $j(\mathbb{K})$.

(2) Nous notons $\mathbb{K}[A]_{\mathbb{L}}$ *le plus petit sous-anneau de* $\mathbb{L}$ *contenant* $j(\mathbb{K})$ *et* A. *C'est l'intersection de tous les sous-anneaux de* $\mathbb{L}$ *contenant* A *et* $j(\mathbb{K})$.

Nous disons que l'extension $\mathbb{L}$ *de* $\mathbb{K}$ *est* ***monogène*** *ou* ***simple*** *si 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.*

Le plus souvent, l'indice $\mathbb{L}$ dans $\mathbb{K}(A)_{\mathbb{L}}$ et $\mathbb{K}[A]_{\mathbb{L}}$ est omis parce que le contexte est clair[40], et nous avons même très souvent $\mathbb{K} \subset \mathbb{L}$ en tant qu'ensembles. Dans ce cas, l'application j est l'identité et elle sera omise.

Remarque 6.87.
Les ensembles $\mathbb{K}(A)$ et $\mathbb{K}[A]$ sont aussi appelés respectivement corps **engendré** et anneau engendré par A. Cependant il faut bien remarquer que ce sont les parties de $\mathbb{L}$ engendrées par A. Il n'est pas question à priori de parler de corps engendré par A sans dire dans quel corps plus grand nous nous plaçons.

Exemple 6.88.
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. △

38. En lisant quelques pages de Wikipédia, vous pourrez briller en société, mais ne tentez pas le coup à l'agrégation.
39. Définition 6.55.
40. Et je me demande si il est possible de trouver un cas tordu où $\mathbb{K}(A)_{\mathbb{L}} \neq \mathbb{K}(A)_{\mathbb{M}}$. Par exemple lorsque A est dans $\mathbb{L}$ et $\mathbb{M}$, mais que $\mathbb{L}$ n'est pas inclus dans $\mathbb{M}$, ni $\mathbb{M}$ dans $\mathbb{L}$.

Exemple 6.89.
Nous avons déjà vu à l'occasion de la définition 1.303 que $A[X]$ est l'anneau de tous les polynômes de degré fini en X. Cela rentre dans le cadre de la définition 6.86 parce qu'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 qu'un éventuel inverse de X n'est pas dedans [41]. Par contre, $\mathbb{K}(X)$ est un corps parce qu'il contient également les fractions rationnelles. △

Exemple 6.90.
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 étendu $\mathbb{F}_p$ 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 $x \in \mathbb{F}_p$. △

Dans l'énoncé suivant, la notation $R(\alpha)_{\mathbb{L}}$ signifie que l'évaluation de R sur α se fait en calculant dans le sur-corps $\mathbb{L}$ de $\mathbb{K}$. Cette proposition semble indiquer que $\mathbb{K}(\alpha)$ est donné en termes de $\mathbb{K}(X)$, lequel est défini de façon très intrinsèque sans faire appel implicitement à un sur-corps de $\mathbb{K}$.

Proposition 6.91 ([1])**.**
Soit une extension $\mathbb{L}$ du corps $\mathbb{K}$ et $\alpha \in \mathbb{L}$. Alors nous avons les isomorphismes de corps suivants :

(1) $\mathbb{K}(\alpha)_{\mathbb{L}} = \operatorname{Frac}\big(\mathbb{K}[\alpha]_{\mathbb{L}}\big)$,

(2) $\mathbb{K}(\alpha)_{\mathbb{L}} = \{R(\alpha)_{\mathbb{L}}$ *tel que* $R \in \mathbb{K}(X)\}$.

Démonstration. Le corps $\mathbb{K}(\alpha)$ est un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}[\alpha]$ comme sous-anneau. La proposition 1.314 nous dit alors que l'application suivante est un morphisme injectif de corps :

$$\begin{aligned} \epsilon\colon \operatorname{Frac}\big(\mathbb{K}[\alpha]\big) &\to \mathbb{K}(\alpha) \\ P/Q &\mapsto PQ^{-1}. \end{aligned} \tag{6.146}$$

Pour rappel, la notation P/Q est bien une notation pour la classe d'équivalence du couple (P, Q) pour la relation définie en 1.313.

Par ailleurs, la partie $\epsilon\Big(\operatorname{Frac}\big(\mathbb{K}[\alpha]\big)\Big)$ de $\mathbb{L}$ et est un corps contenant $\mathbb{K}$ et α. Donc ce corps fait partie des corps sur lesquels on prend l'intersection pour définir $\mathbb{K}(\alpha)$ [42]. Cela prouve que

$$\mathbb{K}(\alpha) \subset \epsilon\Big(\operatorname{Frac}\big(\mathbb{K}[\alpha]\big)\Big). \tag{6.147}$$

L'application ϵ est donc surjective sur $\mathbb{K}(\alpha)$. Comme elle était déjà injective, elle est bijective.

Pour la seconde partie, veuillez lire la définition 1.316 de l'évaluation d'une fraction rationnelle sur un élément de l'anneau. Si $R = P/Q \in \mathbb{K}(X)$ et si $\alpha \in \mathbb{L}$, nous avons

$$R(\alpha) = P(\alpha)Q(\alpha)^{-1}. \tag{6.148}$$

Tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α doit contenir en particulier $\{P(\alpha)$ tel que $P \in \mathbb{K}[X]\}$, les inverses $\{P(\alpha)^{-1}$ tel que $P \in \mathbb{K}[X]$, $P(\alpha) \neq 0\}$ et les produits de ceux-ci. Donc tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α contient $\{R(\alpha)$ tel que $R \in \mathbb{K}(X)\}$.

Nous avons donc

$$\{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\} \subset \mathbb{K}(\alpha). \tag{6.149}$$

Mais puisque $\mathbb{K}(\alpha)$ est lui-même un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, il est contenu dans $\{R(\alpha)$ tel que $R \in \mathbb{K}(X)\}$. D'où l'égalité. □

41. Lorsqu'on multiplie, les degrés montent toujours.
42. Pour rappel, la définition 6.86(1) donne $\mathbb{K}(\alpha)$ comme une intersection.

Pourquoi cela ne contredit pas l'exemple 6.84 ? Lorsque nous écrivons

$$\mathbb{K}(\alpha) = \{R(\alpha) \text{ tel que } R \in \mathbb{K}(X)\}, \tag{6.150}$$

certes $\mathbb{K}(X)$ est défini sans faire appel à un corps contenant $\mathbb{K}$. Mais l'évaluation $R(\alpha)$, oui. Pour calculer $R(\alpha)$, il faut écrire $R = P/Q$ et calculer $P(\alpha)Q(\alpha)^{-1}$. Tous les calculs de cette dernière expression doivent se faire dans un sur-corps de $\mathbb{K}$. Il suffit que le sur-corps en question soit un monceau de mauvaise foi comme celui de l'exemple 6.84, et en réalité $\mathbb{K}(\alpha)$ peut ne pas être ce que l'on croit.

Le corolaire suivant montre que les choses s'arrangent.

Corolaire 6.92.
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}_1$ de $\mathbb{K}$, un élément $\alpha \in \mathbb{L}_1$ et une extension $\mathbb{L}_2$ de $\mathbb{L}_1$. Alors

$$\mathbb{K}(\alpha)_{\mathbb{L}_1} = \mathbb{K}(\alpha)_{\mathbb{L}_2}. \tag{6.151}$$

Démonstration. La proposition 6.91 nous dit que

$$\mathbb{K}(\alpha)_{\mathbb{L}_1} = \{R(\alpha)_{\mathbb{L}_1} \text{ tel que } R \in \mathbb{K}(X)\} \tag{6.152a}$$
$$\mathbb{K}(\alpha)_{\mathbb{L}_2} = \{R(\alpha)_{\mathbb{L}_2} \text{ tel que } R \in \mathbb{K}(X)\}. \tag{6.152b}$$

Mais lorsque $R \in \mathbb{K}(X)$, le calcul de $R(\alpha)$ est exactement le même dans $\mathbb{L}_1$ et dans $\mathbb{L}_2$ parce que $\mathbb{L}_2$ est un sur-corps de $\mathbb{L}_1$ et que les calculs effectifs de $R(\alpha) = P(\alpha)Q(\alpha)^{-1}$ ne font intervenir que des quantités de $\mathbb{K}$ et des puissances de α. □

Ce que ce corolaire nous dit est que si le contexte fixe une extension de $\mathbb{K}$, nous pouvons faire tous les calculs dans cette extension, même si il y a des piles d'extensions à côté.

Typiquement, à chaque fois que nous considérons des sous-corps de $\mathbb{C}$, les extensions se feront dans $\mathbb{C}$: pour tout $\alpha \in \mathbb{C}$, les corps $\mathbb{Q}(\alpha)$, $\mathbb{R}(\alpha)$ se calculent dans $\mathbb{C}$.

Proposition 6.93.
Soit un corps $\mathbb{K}$, une extension $\mathbb{L}$ et un élément $\alpha \in \mathbb{L}$. Nous considérons l'application

$$\begin{aligned} \varphi\colon \mathbb{K}[X] &\to \mathbb{L} \\ P &\mapsto P(\alpha). \end{aligned} \tag{6.153}$$

(1) Si α est transcendant, alors $\mathbb{K}[\alpha] = \mathbb{K}[X]$ (isomorphisme d'anneaux).

(2) Si α est transcendant, alors $\mathbb{K}(\alpha)_{\mathbb{L}} = \mathbb{K}(X)$ (isomorphisme de corps),

(3) Si α est algébrique, alors $\ker(\varphi)$ est un idéal possédant un unique générateur unitaire, lequel est le polynôme minimal[43] *de α sur $\mathbb{K}$.*

Démonstration. Point par point.

(1) Nous savons que $\mathbb{K}[\alpha] = \{Q(\alpha) \text{ tel que } Q \in \mathbb{K}[X]\}$ (c'est la proposition 3.125). Donc φ est surjective sur $\mathbb{K}[\alpha]$, et est donc bijective. Elle est un isomorphisme[44] parce que le lemme 6.66 dit déjà que c'est un morphisme.

(2) Nous supposons encore que α est transcendant et nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{K}(X) &\to \mathbb{K}(\alpha) \\ P &\mapsto R(\alpha). \end{aligned} \tag{6.154}$$

Note : cette application n'est pas φ. En effet φ n'est définie que sur $\mathbb{K}[X]$; le corps des fractions $\mathbb{K}(X)$ est nettement plus grand (classes d'équivalence de couples).

43. Définition 6.60.

44. Les amateurs d'écriture inclusive ne seront, je l'espère, pas choqué par « *elle* est *un* isomorphisme » ; c'est une tournure que je propose ici sur le modèle de l'immonde « *elle* est *un* ministre » ou, à peine moins grave, « *il* est *une* sommité ».

Le fait que cette application soit surjective est la proposition 6.91(2). Pour l'injectivité nous supposons que $\psi(R) = 0$, c'est-à-dire que $R(\alpha) = 0$. Nous considérons un représentant (P, Q) de R ; c'est-à-dire $R = P/Q$. L'égalité $R(\alpha) = 0$ signifie $P(\alpha)Q(\alpha)^{-1} = 0$ (égalité dans $\mathbb{L}$). Puisque $\mathbb{L}$ est un corps, c'est un anneau intègre et nous avons la règle du produit nul ; soit $P(\alpha) = 0$, soit $Q(\alpha)^{-1} = 0$. La seconde possibilité est impossible parce que zéro n'est pas inversible. Donc $P(\alpha) = 0$. Donc $\varphi(P) = 0$ et φ étant injective, $P = 0$.

Lorsque $P = 0$, la classe P/Q est nulle dans $\mathbb{K}(X) = \text{Frac}\big(\mathbb{K}[X]\big)$.

(3) C'est le lemme-définition 6.60.

□

Proposition 6.94.
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[45] *tout polynôme annulateur.*

Autrement dit, l'idéal engendré par le polynôme minimal est l'idéal des polynômes annulateurs.

Démonstration. Nous considérons l'idéal

$$I = \{Q \in \mathbb{K}[X] \text{ tel que } Q(a) = 0\}. \tag{6.155}$$

Le fait que cela soit un idéal est simplement dû à la définition du produit : $(PQ)(a) = P(a)Q(a)$. Par le théorème 6.40, 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. □

6.4.4.1 Extension algébrique, degré

Proposition 6.95.
Toute extension finie est algébrique.

Démonstration. Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ de degré[46] n de $\mathbb{K}$ et $a \in \mathbb{L}$. Nous devons montrer qu'il existe un polynôme annulateur de a à coefficients dans $\mathbb{K}$.

Soit la partie $S = \{1, a, a^2, \dots, a^n\}$ de $\mathbb{L}$. Si cette partie contient des éléments non distincts, alors c'est plié. En effet, si $a^k = a^l$, alors le polynôme X^{k-l} est un polynôme annulateur de a.

Nous supposons donc que S contienne exactement $n+1$ éléments distincts. Le lemme 4.10 nous assure que S est une partie liée : il existe des éléments $k_i \in \mathbb{K}$ tels que $\sum_{i=0}^{n} k_i a^i = 0$.

Donc le polynôme $\sum_i a_i X^i$ est un polynôme annulateur de a. □

Proposition 6.96 (Propriétés d'extensions algébriques[1])**.**
Soit $\mathbb{K}$ un corps commutatif[47] *et a un élément algébrique*[48] *sur $\mathbb{K}$, de polynôme minimal μ_a de degré n. Alors*

(1) En considérant l'application d'évaluation

$$\begin{aligned} \varphi_a \colon \mathbb{K}[X] &\to \mathbb{L} \\ Q &\mapsto Q(a), \end{aligned} \tag{6.156}$$

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 $\{1, a, a^2, \dots, a^{n-1}\}$.

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

$$\big[\mathbb{K}[a] : \mathbb{K}\big] = n. \tag{6.157}$$

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

45. Définition 3.130.
46. Définition 6.57.
47. Juste en passant nous rappelons que tous les corps considérés ici sont commutatifs
48. Définition 6.66.

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

(6) Il existe un isomorphisme d'anneaux $\varphi\colon \mathbb{K}[a] \to \mathbb{K}[X]/(\mu_a)$ *tel que* $\varphi(k) = \bar{k}$ *pour tout* $k \in \mathbb{K}$. $\mathbb{K}[a] \simeq \mathbb{K}[X]/(\mu_a)$ *(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 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.

(2) La partie $\{1, a, a^2, \dots, a^{n-1}\}$ 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 de degré n.

Rappelons qu'en vertu de la définition 6.60, 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 :

$$Q(X) = Q_1(X) + X^n Q_2(X) \tag{6.158}$$

où Q_1 est de degré $n-1$ et Q_2 de degré $p-n$. Nous évaluons cette égalité en a :

$$Q(a) = Q_1(a) + S(a)Q_2(a). \tag{6.159}$$

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 6.63). Donc le théorème de Bézout 6.44 affirme l'existence d'éléments $U, V \in \mathbb{K}[X]$ tels que

$$U\mu_a + VQ = 1 \tag{6.160}$$

dans $\mathbb{K}[X]$. Nous évaluons cette égalité en a en tenant compte de $\mu_a(a) = 0$ dans $\mathbb{K}[a]$:

$$U(a)\mu_a(a) + V(a)Q(a) = 1 \tag{6.161}$$

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

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(\mu_a) &\to \mathbb{K}[a] \\ \bar{R} &\mapsto R(a) \end{aligned} \tag{6.162}$$

et nous montrons qu'elle convient. Pour cela, nous nous souvenons que la proposition 6.94 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 3.44 de l'idéal (μ_a). Cela étant mis au point, nous passons à la preuve.

(i) **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 $Q(a) = 0$.

(ii) **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 $x = \psi(\bar{Q})$.

(iii) **Injective** Si $\psi(\bar{R}) = 0$ alors $R(a) = 0$, mais comme mentionné plus haut, μ_a engendre l'idéal des polynômes annulateurs de a. Donc $R \in (\mu_a)$ et nous avons $\bar{R} = 0$ dans $\mathbb{K}[X]/(\mu_a)$.

□

Exemple 6.97.
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 6.96. Du point de vue de Bézout, $\mu_{\sqrt{2}}(X) = X^2 - 2$, et nous cherchons des polynômes U et V tels que

$$U(X^2 - 2) + VX = 1. \tag{6.163}$$

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 $\sqrt{2}$:

$$V(\sqrt{(2)}) = \frac{1}{2}\sqrt{2}. \tag{6.164}$$

△

Proposition 6.98 ([146])**.**
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ de $\mathbb{K}$ et un élément α de $\mathbb{L}$. Il y a équivalence entre les trois points suivants :

(1) α est algébrique sur $\mathbb{K}$,

(2) $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$,

(3) $\mathbb{K}[\alpha]$ est un $\mathbb{K}$-espace vectoriel de dimension finie.

Si ces affirmation sont vraies, alors $[\mathbb{K}(\alpha) : \mathbb{K}]$ est le degré du polynôme minimal de α sur $\mathbb{K}$.

Démonstration. Démonstration décomposée en plusieurs implications.

(i) **(1) implique (2)** Soit α algébrique sur $\mathbb{K}$. Nous considérons le polynôme minimal de α sur $\mathbb{K}$ (définition 6.60). Nous savons par le lemme 6.77 (qui fonctionne parce que α est algébrique) que $\mathbb{K}[\alpha] = \mathbb{K}[X]/(\mu)$ en tant qu'anneaux.
Mais $\mathbb{K}[X]$ est un anneau principal et μ en est un élément irréductible. Donc la proposition 3.81 dit que (μ) est un idéal maximum ; la proposition 3.83 avance encore un peu en disant que $\mathbb{K}[X]/(\mu)$ est un corps.
Donc $\mathbb{K}[X]/(\mu)$ est un corps isomorphe à $\mathbb{K}[\alpha]$ en tant qu'anneaux. En conséquence de quoi $\mathbb{K}[\alpha]$ est un corps.
Le corps $\mathbb{K}[\alpha]$ est un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α ; par définition nous avons donc $\mathbb{K}(\alpha) \subset \mathbb{K}[\alpha]$.
Mais d'autre part, $\mathbb{K}[\alpha]$ est contenu dans tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et α, donc il est inclus dans l'intersection de tout ces corps, donc $\mathbb{K}[\alpha] \subset \mathbb{K}(\alpha)$.
Nous avons donc l'égalité $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$.

(ii) **(2) implique (1)** Nous montrons que non-(1) implique non-(2). Nous disons donc que α est transcendant sur $\mathbb{K}$; cela implique par la proposition 6.93(1) que $\mathbb{K}[\alpha] = \mathbb{K}[X]$ en tant qu'anneaux. Donc $\mathbb{K}[\alpha]$ n'est pas un corps parce que $\mathbb{K}[X]$ ne l'est pas.
N'étant pas un corps, $\mathbb{K}[\alpha]$ ne peut pas être égal à $\mathbb{K}(\alpha)$ qui, lui, est un corps.

(iii) **(1) implique (3)** L'élément α est maintenant algébrique et nous considérons son polynôme minimal μ. Nous savons par le lemme 6.77 que $\mathbb{K}[\alpha] = \mathbb{K}[X]/(\mu)$ en tant qu'espaces vectoriels. Or $\mathbb{K}[X]/(\mu)$ est de dimension finie $\deg(\mu)$. Donc $\mathbb{K}[\alpha]$ est également de dimension finie.

(iv) **(3) implique (1)** Nous démontrons la contraposée. En supposant que α est transcendant nous avons $\mathbb{K}[\alpha] = \mathbb{K}[X]$ par la proposition 6.93. Or $\mathbb{K}[X]$ n'est pas de dimension finie sur $\mathbb{K}$, donc $\mathbb{K}[\alpha]$ non plus.

□

Lemme 6.99 ([147]).
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 $\mathbb{L}$.

Lemme 6.100.
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. C'est un corps par le corolaire 6.41. Nous identifions $\mathbb{K}$ avec $\phi(\mathbb{K})$ où

$$\phi\colon \mathbb{K}[X] \to \mathbb{L} \tag{6.165}$$

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 6.101 ([78]).
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

$$I_a = \{Q \in \mathbb{K}[X] \text{ tel que } Q(a) = 0\}; \tag{6.166a}$$
$$I_b = \{Q \in \mathbb{K}[X] \text{ tel que } Q(b) = 0\}. \tag{6.166b}$$

Même si $Q(a)$ et $Q(b)$ sont calculés dans $\mathbb{L}$, I_a, I_b sont des idéaux de $\mathbb{K}[X]$. Le polynôme μ_a est par définition le générateur unitaire de I_a, et comme a est une racine de P, nous avons $P \in I_a$ et il existe un polynôme $Q \in \mathbb{K}[X]$ tel que

$$P = \mu_a Q. \tag{6.167}$$

Montrons que $\mu_a(b) \neq 0$. Pour cela, nous supposons que $\mu_a(b) = 0$, c'est-à-dire que $\mu_a \in I_b$. Il existe alors $R \in \mathbb{K}[X]$ tel que $\mu_a = \mu_b R$. Mais par la proposition 6.63, le polynôme μ_a est irréductible, donc soit μ_b, soit R, est inversible. Comme les inversibles sont les éléments de $\mathbb{K}$ (polynômes de degré zéro), μ_b n'est pas inversible (sinon il serait constant et ne pourait pas être annulateur de b). Donc R est inversible. Disons $R = k$.

Donc $\mu_a = k\mu_b$. Mais puisque μ_a et μ_b sont unitaires, nous avons obligatoirement $k = 1$. Cela donnerait $\mu_a = \mu_b$, ce qui est contraire aux hypothèses. Nous en déduisons que $\mu_a(b) \neq 0$.

Étant donné que $\mu_a(b) \neq 0$, l'évaluation de (6.167) en b montre que $Q(b) = 0$, de telle sorte que $Q \in I_b$ et il existe un 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 6.102.
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 $\mu_b = X^2 + 2$. 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 $\mu_a = \mu_b = X^2 + 1$, tandis que le polynôme μ_a^2 ne divise pas P. △

6.4.5 Racines de polynômes

Corolaire 6.103 (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 $P(x) = (X - \alpha)Q$.

Démonstration. Il s'agit d'un cas particulier de la proposition 6.94 : 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. □

Avant de lire l'énoncé suivant, allez relire la définition 3.123 pour savoir ce qu'est un polynôme nul.

Théorème 6.104 (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 $P = 0$.

Démonstration. Si P est de degré 1, il s'écrit $P = aX + b$; si il a comme racines α et β, nous avons le système

$$\begin{cases} a\alpha + b = 0 & \text{(6.168a)} \\ a\beta + b = 0. & \text{(6.168b)} \end{cases}$$

La différence entre les deux donne $a(\alpha - \beta) = 0$. Puisque $\alpha \neq \beta$, la règle du produit nul (lemme 1.242) 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. Puisque $P(\alpha_1) = 0$, le corolaire 6.103 nous donne un polynôme Q de degré n tel que

$$P = (X - \alpha_1)Q. \tag{6.169}$$

Étant donne que pour tout $i \neq 1$ nous avons $\alpha_i \neq \alpha_1$,

$$0 = P(\alpha_i) = \underbrace{(\alpha_i - \alpha_1)}_{\neq 0} Q(\alpha_i), \tag{6.170}$$

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 (6.169) sont alors également nuls. □

Exemple 6.105.
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.

Nous verrons dans la proposition 6.173 une condition pour qu'un polynôme à plusieurs variables s'annule du fait qu'il ait « trop » de racines. △

Remarque 6.106.
L'intérêt du théorème 6.104 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 13.25.

6.4.6 Corps de rupture

Définition 6.107.
Soit $P \in \mathbb{K}[X]$ un polynôme irréductible. Une extension $\mathbb{L}$ de $\mathbb{K}$ est un ***corps de rupture*** *pour P si il existe $a \in \mathbb{L}$ tel que $P(a) = 0$ et $\mathbb{L} = \mathbb{K}(a)$.*

6.108.
Nous insistons sur le fait que nous ne définissons le concept de corps de rupture que pour un polynôme irréductible à coefficients dans un corps. Les deux points sont importants : irréductible et à coefficient dans un corps.

Nous discuterons brièvement le pourquoi de cela dans la section 6.4.11.

Exemple 6.109.
Soit $\mathbb{K} = \mathbb{Q}$ et $P = X^2 - 2$. On pose $a = \sqrt{2}$ et $\mathbb{L} = \mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$. De cette façon P est scindé dans $\mathbb{L}$:

$$P = (X - \sqrt{2})(X + \sqrt{2}). \tag{6.171}$$

Le corps $\mathbb{Q}(\sqrt{2})$ est donc un corps de rupture pour P. △

Exemple 6.110.
Dans l'exemple 6.109, nous avions un corps de rupture dans lequel le polynôme P était scindé. Il n'en est pas toujours ainsi. Prenons

$$P = X^3 - 2 \tag{6.172}$$

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 6.111.
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 degré 1 en $\sqrt{s}$. Le cardinal en est p^2. △

Comme nous allons abondamment parler du quotient $\mathbb{K}[X]/(P)$, nous nous permettons un petit lemme.

Lemme 6.112.
Soit un corps $\mathbb{K}$ et $P \in \mathbb{K}[X]$ non constant. Alors $\mathbb{K}[X]/(P)$ est un corps si et seulement si P est irréductible.

Démonstration. Nous utilisons le trio d'enfer dont il est question dans le thème 16. D'abord $\mathbb{K}[X]$ est un anneau principal par le lemme 3.132. Donc $\mathbb{K}[X]/(P)$ sera un corps si et seulement si (P) est un idéal maximum (proposition 3.52), et cela sera le cas si et seulement si (P) est engendré par un polynôme irréductible (proposition 3.81).

Il ne nous reste qu'à montrer que (P) est engendré par un irréductible si et seulement si P est irréductible. Il y a un sens dans lequel c'est évident.

Soit un irréductible μ tel que $(P) = (\mu)$. En particulier $\mu \in (P)$, c'est-à-dire qu'il existe Q tel que $\mu = PQ$. Puisque μ est irréductible, soit P, soit Q, est inversible. Si P est inversible, c'est-à-dire constant, c'est ce que nous avons exclu par hypothèse. Si par contre Q est inversible, alors $P = k\mu$ pour un certain $k \in \mathbb{K}$, ce qui montre que P est irréductible autant que μ. □

Proposition 6.113 (Existence d'un corps de rupture)**.**
Soit un corps $\mathbb{K}$ et un polynôme irréductible non constant P. Alors

(1) Le corps $\mathbb{L} = \mathbb{K}[X]/(P)$ est un corps de rupture pour P.

(2) L'élément $\bar{X}$ de $\mathbb{L}$ est une racine de P.

(3) $\mathbb{L} = \mathbb{K}(\bar{X})_{\mathbb{L}}$

Démonstration. Commençons par nous convaincre que $\mathbb{K}[X]/(P)$ est une extension de $\mathbb{K}$ (définition 6.55). Le fait que ce soit un corps est le lemme 6.112. Le morphisme $j\colon \mathbb{K} \to \mathbb{K}[X]/(P)$ est simplement $k \mapsto \bar{k}$ où à droite, $\bar{k}$ voit k dans $\mathbb{K}[X]$ comme étant le polynôme constant. Notez qu'il est automatiquement injectif (lemme 1.244).

Il faut maintenant voir que $\mathbb{K}[X]/(P) = \mathbb{K}(\alpha)$ pour un certain $\alpha \in \mathbb{K}[X]/(P)$. Grâce à notre compréhension des notations acquise dans 1.16.2.2, nous savons que $X \in \mathbb{K}[X]$ et qu'il est donc parfaitement légitime de poser $\alpha = \bar{X}$ dans $\mathbb{K}[X]/(P)$. Il s'agit simplement de l'ensemble $\bar{X} = \{X + QP \text{ tel que } Q \in \mathbb{K}[X]\}$ où X est une notation pour la suite $(0, 1, 0, 0, \ldots)$.

Bref, nous notons $\alpha = \bar{X}$ et nous démontrons que $P(\alpha) = 0$ et que $\mathbb{K}[X]/(P) = \mathbb{K}(\alpha)$ (isomorphisme de corps).

(i) $\underline{P(\bar{X}) = 0}$ C'est le moment de nous souvenir comment la notation des X fonctionne, et en particulier la pirouette autour de (1.444). D'abord la définition du produit sur $\mathbb{K}[X]/(P)$ est $\bar{P}\bar{Q} = \overline{PQ}$; en particulier si $P = \sum_k a_k X^k$, alors $P(\bar{X}) = \sum_k a_k \bar{X}^k = \sum_k a_k \overline{X^k}$, et

$$P(\bar{X}) = \overline{P(X)} = \bar{P} = 0. \tag{6.173}$$

(ii) **L'égalité** Nous montrons à présent que $\mathbb{K}(\bar{X})_{\mathbb{L}} = \mathbb{L}$. C'est-à-dire que $\mathbb{L}$ est bien engendrée par $\mathbb{K}$ et un seul élément. D'abord, $\mathbb{L} = \mathbb{K}[X]/(P)$ contient bien évidemment $\mathbb{K}$ et $\bar{X}$. Ensuite nous devons prouver que tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et $\bar{X}$ est en réalité $\mathbb{L}$ entier.

Soit $Q \in \mathbb{K}[X]$, et montrons que $\bar{Q}$ est dans tout sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et $\bar{X}$.

Par le lemme 3.121 nous avons $\bar{Q} = Q(\bar{X})$. Et si un corps contient $\mathbb{K}$ et $\bar{X}$, il doit contenir tous les polynômes en $\bar{X}$ à coefficients dans $\mathbb{K}$. Donc un tel corps doit contenir $Q(\bar{X})$ et donc $\bar{Q}$.

□

Exemple 6.114.
Soit le polynôme $P = X^2 + 1 \in \mathbb{Z}[X]$. Dans le quotient $\mathbb{Z}[X]/(P)$ nous avons $\bar{X}^2 + 1 = 0$ et donc $\bar{X}^2 = -1$. C'est-à-dire que $\mathbb{Z}[X]/(P)$ contient un élément dont le carré est -1. Avouez que c'est bien ce à quoi nous nous attendions.

Notons que $-\bar{X}$ est également une racine de P dans $\mathbb{Z}[X]/(P)$.

En calculant dans les polynômes à coefficients dans $\mathbb{Z}(\bar{X})$ nous avons :

$$(X + \bar{X})(X - \bar{X}) = X^2 - \bar{X}^2 = X^2 + 1, \tag{6.174}$$

c'est-à-dire que P est bien factorisé, et que nous avons retrouvé la multiplication $x^2 + 1 = (x + i)(x - i)$. △

6.115.
Il n'y a évidemment pas unicité d'un corps de rupture pour un polynôme donné. Une raison est qu'un polynôme peut accepter plusieurs racines complètement indépendantes. Le corps étendu par l'une ou l'autre racine donne deux corps de rupture différents. Par exemple dans $\mathbb{Q}[X]$, le polynôme

$$P = X^4 - X^2 - 2 \tag{6.175}$$

a pour racines (dans $\mathbb{C}$) les nombres $\sqrt{2}$ et i. Donc on a deux corps de rupture complètement différents : $\mathbb{Q}(\sqrt{2})$ et $\mathbb{Q}(i)$.

6.116.
La proposition suivante donne une unicité du corps de rupture dans le cas d'un polynôme irréductible. Et nous comprenons pourquoi : un polynôme irréductible n'a fondamentalement qu'une seule racine « indépendante ». Par exemple $X^2 - 2$ a pour racines $\pm\sqrt{2}$. Autre exemple, le polynôme $X^2 + 6X + 13$ a pour racines, dans $\mathbb{C}$, les nombres complexes conjugués $z = -3 + 2i$ et $\bar{z} = -3 - 2i$.

Proposition 6.117 ([141]).
Soient un corps $\mathbb{K}$ et un polynôme irréductible $P \in \mathbb{K}[X]$. Alors toute extension $\mathbb{L}$ contenant une racine α de P admet un unique morphisme de corps

$$\psi\colon \mathbb{K}[X]/(P) \to \mathbb{L} \tag{6.176}$$

tel que $\psi(\bar{X}) = \alpha$.

Dans un tel cas,

(1) l'image de ψ est $\mathbb{K}(\alpha)_{\mathbb{L}}$,

(2) si $\mathbb{L} = \mathbb{K}(\alpha)_{\mathbb{L}}$ alors ψ est un isomorphisme.

Démonstration. L'idéal annulateur de α parmi les polynôme de $\mathbb{K}[X]$ n'est pas réduit à $\{0\}$ parce qu'il contient P. Le lemme 6.60 s'applique donc et nous avons μ, le polynôme minimal de α dans $\mathbb{K}[X]$. Il divise P qui est irréductible, donc

$$P = \lambda\mu \tag{6.177}$$

pour un certain $\lambda \in \mathbb{K}$.

Nous posons

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(P) &\to \mathbb{L} \\ \bar{Q} &\mapsto Q(\alpha). \end{aligned} \tag{6.178}$$

(i) **Bien définie** Si $\bar{Q}_1 = \bar{Q}_2$ alors il existe un $R \in \mathbb{K}[X]$ tel que $Q_1 = Q_2 + RP$. Mais alors $\psi(\bar{Q}_1) = Q_1(\alpha) = Q_2(\alpha) + R(\alpha)P(\alpha) = Q_2(\alpha)$.

(ii) **Injective** Si $\psi(\bar{Q}_1) = \psi(\bar{Q}_2)$ alors $Q_1 - Q_2 = R$ pour un certain $R \in \mathbb{K}[X]$ vérifiant $R(\alpha) = 0$. Nous avons alors un polynôme S tel que $R = S\mu = \lambda^{-1}SP$. Donc $\bar{R} = 0$ et donc $\bar{Q}_1 = \bar{Q}_2$.

(iii) **Morphisme** Laissé comme exercice ; la paresse de l'auteur de ces lignes attend vos contributions.

(iv) **La condition** Le morphisme ψ respecte de plus la condition

$$\psi(\bar{X}) = X(\alpha) = \alpha. \tag{6.179}$$

En ce qui concerne l'unicité, fixer $\psi(\bar{X})$ est suffisant pour fixer un morphisme. En effet si $\psi(\bar{X}) = \alpha$, alors

$$\psi(\bar{Q}) = \psi\Big(\sum_k a_k \bar{X}^k\Big) = \sum_k a_k \psi(\bar{X})^k = \sum_k a_k \alpha^k. \tag{6.180}$$

Pour le second point de l'énoncé, il faut remarquer que α est algébrique et non transcendant. Donc en utilisant les propositions 3.125 et 6.96(5) nous trouvons

$$\text{Image}(\psi) = \{Q(\alpha) \text{ tel que } Q \in \mathbb{K}[X]\} = \mathbb{K}[\alpha] = \mathbb{K}(\alpha). \tag{6.181}$$

Et finalement pour le dernier point, un morphisme de corps est toujours injectif. Si il est également surjectif, il sera bijectif. □

6.4.7 Pile d'extensions

Lemme 6.118 ([1]).
Soient un corps $\mathbb{K}$, des extensions $\mathbb{L}_1, \ldots, \mathbb{L}_n$ et des éléments $\alpha_i \in \mathbb{L}_i$ tels que

$$\mathbb{L}_1 = \mathbb{K}(\alpha_1)_{\mathbb{L}_1} \tag{6.182}$$

et

$$\mathbb{L}_k = \mathbb{L}_{k-1}(\alpha_k)_{\mathbb{L}_k}. \tag{6.183}$$

Alors

$$\mathbb{L}_n = \mathbb{K}(\alpha_1, \ldots, \alpha_n)_{\mathbb{L}_n}. \tag{6.184}$$

Démonstration. Nous démontrons par récurrence sur n. Le cas $n = 1$ est simplement l'hypothèse (6.182).

Supposons donc que le lemme soit correct pour n, et étudions le cas $n + 1$. Nous avons, par définition et par hypothèse de récurrence :

$$\mathbb{L}_{n+1} = \mathbb{L}_n(\alpha_{n+1})_{\mathbb{L}_{n+1}} = \Big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\Big)(\alpha_{n+1})_{\mathbb{L}_{n+1}}. \tag{6.185}$$

Notre tâche sera donc de montrer que

$$\Big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\Big)(\alpha_{n+1}) = \mathbb{K}(\alpha_1, \dots, \alpha_{n+1}) \tag{6.186}$$

où nous n'écrivons plus les indices $\mathbb{L}_{n+1}$ partout.

Le membre de gauche est un sous-corps de $\mathbb{L}_{n+1}$ contenant à la fois $\mathbb{K}$ et tous les α_i, si bien que

$$\mathbb{K}(\alpha_1, \dots, \alpha_{n+1}) \subset \big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})_{\mathbb{L}_{n+1}}. \tag{6.187}$$

Il faut donc prouver l'inclusion inverse ; c'est-à-dire montrer que tout élément x du corps $\big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})$ est forcément dans tout sous-corps de $\mathbb{L}_{n+1}$ contenant $\mathbb{K}$ et les α_i. Un tel élément x est, par la proposition 6.91(2), de la forme $r(\alpha_{n+1})$ avec $r \in \mathbb{K}(\alpha_1, \dots, \alpha_n)(X)$, c'est-à-dire

$$P(\alpha_{n+1})Q(\alpha_{n+1})^{-1} \tag{6.188}$$

avec $P, Q \in \mathbb{K}(\alpha_1, \dots, \alpha_n)[X]$.

Prouvons d'abord que si $P \in \mathbb{K}(\alpha_1, \dots, \alpha_n)[X]$, alors $P(\alpha_{n+1})$ est dans tout sous-corps de $\mathbb{L}_{n+1}$ contenant $\mathbb{K}$ et les α_i. Nous pouvons écrire $P = \sum_i a_i X^i$ avec $a_i \in \mathbb{K}(\alpha_1, \dots, \alpha_n)$, et donc

$$P(\alpha_{n+1}) = \sum_i a_i \alpha_{n+1}^i. \tag{6.189}$$

Tout corps contenant $\mathbb{K}$ et les $\alpha_1, \dots, \alpha_n$ contient les a_i. Par produit, tout corps contenant $\mathbb{K}$, $\alpha_1, \dots, \alpha_{n+1}$ contient les termes $a_i \alpha_{n+1}^i$, et donc $P(\alpha_{n+1})$ par somme.

De la même façon, si un corps contient $\mathbb{K}$ et les α_i, $(i = 1, \dots, n+1)$, alors il contient $Q(\alpha_{n+1})$. Comme c'est un corps, il contient aussi son inverse $Q(\alpha_{n+1})^{-1}$, et il contient aussi le produit

$$r(\alpha_{n+1}) = P(\alpha_{n+1})Q(\alpha_{n+1})^{-1}. \tag{6.190}$$

On vient ainsi de montrer que tout élément $x \in \big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})$ était dans tout sous-corps de $\mathbb{L}_{n+1}$ qui contient $\mathbb{K}$ et les α_i, $(i = 1, \dots, n + 1)$; en d'autres termes :

$$\big(\mathbb{K}(\alpha_1, \dots, \alpha_n)_{\mathbb{L}_n}\big)(\alpha_{n+1})_{\mathbb{L}_{n+1}} \subset \mathbb{K}(\alpha_1, \dots, \alpha_{n+1}). \tag{6.191}$$

Les inclusions (6.187) et (6.191) prouvent l'égalité d'ensembles (6.186) que nous voulions montrer. □

6.4.8 Clôture algébrique

Le concept de clôture algébrique a été défini dans 6.70. Voici un lemme qui dit qu'une clôture algébrique est en quelque sorte une extension algébrique maximale.

Lemme 6.119 ([1]).
Soient un corps $\mathbb{K}$ et une extension algébrique $\mathbb{F}$ de $\mathbb{K}$. Nous supposons que pour toute extension algébrique de $\mathbb{L}$ nous avons $\mathbb{L} = \mathbb{F}$

Alors $\mathbb{F}$ est algébriquement clos[49].

Démonstration. Soit un polynôme $P \in \mathbb{K}[X]$. Nous voudrions prouver que P a des racines dans $\mathbb{F}$. Pour cela, nous voyons P comme un polynôme sur $\mathbb{F}[X]$ et, grace à la proposition 6.113 nous considérons un corps de rupture $\mathbb{L}$ pour P. Puisque $\mathbb{L}$ est une extension de $\mathbb{F}$, nous avons $\mathbb{L} = \mathbb{F}$. Donc $\mathbb{F}$ contient des racines de P. □

49. Définition 6.70(2).

6.120.
Nous avons défini le concept d'extension algébrique en 6.70. Nous allons en construire un petit exemple très piéton.

D'abord la proposition 1.404 nous donne l'existence et l'unicité d'un réel $\sqrt{2}$ strictement positif dont le carré est 2. Ce réel est irrationnel par la proposition 1.344. Cela étant posé, nous y allons.

Proposition 6.121 ([148])**.**
Soit $\mathbb{L} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}$.

(1) C'est un sous-corps de $\mathbb{R}$.

(2) Tout sous-corps de $\mathbb{R}$ *contenant* $\mathbb{Q}$ *et* $\sqrt{2}$ *contient* $\mathbb{L}$.

Démonstration. Nous devons d'abord prouver que $\mathbb{L}$ est un corps en vérifiant d'une part que c'est un anneau (définition 1.38) et d'autre part le fait que tous les éléments non nuls sont inversibles.

— La partie $\mathbb{L}$ de $\mathbb{R}$ est stable pour l'addition : dès que $a, b, a', b' \in \mathbb{Q}$,

$$(a + b\sqrt{2}) + (a' + b'\sqrt{2}) = (a + a') + (b + b')\sqrt{2} \in \mathbb{L}. \tag{6.192}$$

— Les neutres 0 et 1 sont dans $\mathbb{L}$.

— Si $\alpha \in \mathbb{L}$, alors $-\alpha \in \mathbb{L}$:

$$-(a + b\sqrt{2}) = -a - b\sqrt{2}. \tag{6.193}$$

— La partie $\mathbb{L}$ est stable pour le produit parce que

$$(a + b\sqrt{2})(a' + b'\sqrt{2}) = (aa' + 2bb') + (ab' + ba')\sqrt{2}. \tag{6.194}$$

— L'inverse d'un élément de $\mathbb{L}$ est dans $\mathbb{L}$. C'est le seul point pas tout à fait évident. D'abord, l'ensemble $\mathbb{R}$ est un corps par le théorème 1.352. Donc pour tout $a, b \in \mathbb{R}$, le nombre

$$\frac{1}{a + b\sqrt{2}} \tag{6.195}$$

existe dans $\mathbb{R}$.
D'abord $a - b\sqrt{2}$ n'est pas nul, parce que si il l'était, nous aurions $\sqrt{2} = -a/b \in \mathbb{Q}$ alors que $\sqrt{2}$ n'est pas rationnel par la proposition 1.344. Nous pouvons donc faire le coup de multiplier le numérateur et le dénominateur par le binôme conjugué :

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}. \tag{6.196}$$

Cela est un rationnel. Donc le éléments non nuls de $\mathbb{L}$ ont un inverse qui appartient également à $\mathbb{L}$.

Nous passons à la preuve du point (2). Si $\mathbb{L}'$ est un corps qui contient $\mathbb{Q}$ et $\sqrt{2}$, il doit contenir $b\sqrt{2}$ pour tout $b \in \mathbb{Q}$ et donc aussi tous les $a + b\sqrt{2}$ avec $a, b \in \mathbb{Q}$. Par conséquent, $\mathbb{L}'$ doit contenir au moins tout $\mathbb{L}$. □

Proposition 6.122.
Soit $\mathbb{L} = \{a + b\sqrt{2}\}_{a,b\in\mathbb{Q}}$.

(1) C'est un espace vectoriel de dimension 2 *sur* $\mathbb{Q}$.

(2) Si $\alpha \in \mathbb{L}$, *alors il existe un polynôme* $P \in \mathbb{L}[X]$ *de degré* 2 *ou moins tel que* $P(\alpha) = 0$.

(3) Le corps $\mathbb{L}$ *est une extension algébrique de* $\mathbb{Q}$.

Démonstration. En plusieurs parties.

(i) **(1)** Pour la dimension, notez que $\{1, \sqrt{2}\}$ est une partie libre et génératrice de $\mathbb{L}$.

(ii) **(2)** Soit $\alpha \in \mathbb{L}$. La partie $\{1, \alpha, \alpha^2\}$ est de cardinal 1, 2 ou 3. Si c'est 1 ou 2, c'est que $1 = \alpha$ ou $1 = \alpha^2$ ou $\alpha = \alpha^2$. Si par exemple $1 = \alpha$ alors avec $P = X - 1$ nous avons $P(\alpha) = 0$.

Si au contraire $\{1, \alpha, \alpha^2\}$ est de cardinal 3, alors c'est une partie liée par la proposition 4.6. Il existe donc des rationnels a, b, c tels que $a + b\alpha + c\alpha^2 = 0$, c'est-à-dire $P(\alpha) = 0$ avec $P = cX^2 + bX + a$.

(iii) **(3)** Nous venons de voir que tous les éléments de $\mathbb{L}$ sont des racines de polynômes de $\mathbb{Q}[X]$.

□

Lemme 6.123.
Si $\mathbb{K}$ est un corps infini, alors $\mathbb{K}[X]$ est équipotent[50] *à $\mathbb{K}$.*

Démonstration. Notons provisoirement $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré n. Nous avons une surjection

$$\begin{aligned} \varphi\colon \mathbb{K}^{n+1} &\to \mathbb{K}_n[X] \\ (k_0, \dots, k_n) &\mapsto \sum_{i=0}^{n} k_i X^i. \end{aligned} \tag{6.197}$$

Par récurrence sur le théorème[51] 1.155, nous avons $\mathbb{K}^{n+1} \approx \mathbb{K}$. La surjection (6.197) dit alors que

$$\mathbb{K}_n[X] \preceq \mathbb{K}^{n+1} \approx \mathbb{K}. \tag{6.198}$$

Mais puisqu'il y a une surjection $\mathbb{K} \to \mathbb{K}_n[X]$, nous avons aussi $\mathbb{K}_n[X] \succeq \mathbb{K}$. Le théorème 1.139 dit alors que $\mathbb{K}_n[X] \approx \mathbb{K}$.

Le lemme 1.156 nous permet alors de conclure que

$$\mathbb{K}[X] = \bigcup_{n=0}^{\infty} \mathbb{K}_n[X] \approx \mathbb{K}. \tag{6.199}$$

□

Proposition 6.124 ([1])**.**
Soit un corps $\mathbb{K}$. Une extension algébrique de $\mathbb{K}$ est

(1) au plus dénombrable si $\mathbb{K}$ est fini,

(2) équipotente à $\mathbb{K}$ si $\mathbb{K}$ est infini.

Démonstration. En deux parties.

(i) **Si $\mathbb{K}$ est fini** Un polynôme non nul possède toujours au maximum un nombre fini de racines (éventuellement zéro) par la proposition 6.104. Par ailleurs, chaque degré de polynôme ayant seulement un nombre fini de possibilités, l'ensemble $\mathbb{K}[X]$ est au plus dénombrable (proposition 1.132).

Pour $P \in \mathbb{K}[X]$ nous avons une surjection de $\mathbb{N}$ vers l'ensemble des racines de P. Nous la notons $\varphi_P\colon \mathbb{N} \to \mathbb{L}$, en posant par exemple $\varphi_P(n) = 1$ si P n'a pas de racines. Enfin nous posons

$$\begin{aligned} \varphi\colon \mathbb{K}[X] \times \mathbb{N} &\to \mathbb{L} \\ (P, n) &\mapsto \varphi_P(n). \end{aligned} \tag{6.200}$$

C'est la fonction qui à un polynôme P et un nombre n fait correspondre la n^{e} racine de P.

Comme $\mathbb{L}$ est une extension algébrique, φ est surjective.

En termes de cardinalité, que $\mathbb{K}[X]$ soit fini ou dénombrable, dans les deux cas, $\mathbb{K}[X] \times \mathbb{N}$ est dénombrable (proposition 1.129). Il existe donc une surjection d'un ensemble dénombrable vers $\mathbb{L}$. Le lemme 1.131 conclut que $\mathbb{L}$ est fini ou dénombrable.

50. Définition 1.109.

51. J'ai quand même du mal à croire qu'il faille vraiment le lemme de Zorn pour prouver que $\mathbb{K}[X]$ est équipotent à $\mathbb{K}$. Si vous connaissez un moyen plus direct, écrivez-moi.

(ii) **Si $\mathbb{K}$ est infini** Nous procédons de la même manière, mais nous devons faire appel à des résultats plus technologiques pour maitriser la cardinalité. Nous considérons à nouveau l'application

$$\begin{aligned}\varphi\colon \mathbb{K}[X]\times\mathbb{N}&\to\mathbb{L}\\ (P,n)&\mapsto\varphi_P(n).\end{aligned}\tag{6.201}$$

Cette application est encore surjective : $\mathbb{L}\preceq\mathbb{K}[X]\times\mathbb{N}$. Le lemme 6.123 nous assure que $\mathbb{K}[X]\approx\mathbb{K}$ parce que $\mathbb{K}$ est infini. Ensuite la proposition 1.149 nous dit que $\mathbb{K}[X]\times\mathbb{N}\approx\mathbb{K}[X]$. Donc

$$\mathbb{K}\approx\mathbb{K}[X]\approx\mathbb{K}[X]\times\mathbb{N}\succeq\mathbb{F}.\tag{6.202}$$

Mais $\mathbb{F}$ est une extension de $\mathbb{K}$. Donc il existe une injection $\mathbb{K}\to\mathbb{F}$, c'est-à-dire $\mathbb{K}\preceq\mathbb{F}$.

Ayant $\mathbb{K}\preceq\mathbb{F}\preceq\mathbb{K}$, le théorème 1.139 implique que $\mathbb{K}\approx\mathbb{F}$.

□

Lemme 6.125 ([1]).
Soient des corps $\mathbb{K}$ et $\mathbb{L}$ ainsi qu'un morphisme de corps $\rho\colon\mathbb{K}\to\mathbb{L}$. Si $P\in\mathbb{K}[X]$ a une racine dans $\mathbb{K}$, alors le polynôme $\rho(P)$ a une racine dans $\mathbb{L}$.

Démonstration. Nous notons $P=\sum_k P_kX^k$. Si $a\in\mathbb{K}$ est une racine de P, alors $\sum_k P_ka^k=0$. Nous appliquons ρ à cette égalité : $\sum_k\rho(P_k)\rho(a)^k=0$, c'est-à-dire $\rho(P)\big(\rho(a)\big)=0$. Donc $\rho(a)\in\mathbb{L}$ est une racine de $\rho(P)$. □

Lemme 6.126 ([149]).
Nous considérons un triplet $(\mathbb{K},\mathbb{L},\mathbb{F})$ où

(1) $\mathbb{K}$, $\mathbb{L}$ et $\mathbb{F}$ sont des corps

(2) il existe $a\in\mathbb{L}$ algébrique sur $\mathbb{K}$ tel que $\mathbb{L}=\mathbb{K}(a)$ et un morphisme de corps $\alpha\colon\mathbb{K}\to\mathbb{L}$.

(3) $\mathbb{F}$ est une extension algébriquement close de $\mathbb{K}$: il existe un morphisme $\beta\colon\mathbb{K}\to\mathbb{F}$.

Note : en pratique, les corps $\mathbb{L}$ et $\mathbb{F}$ sont le plus souvent des sur-corps de $\mathbb{K}$, de telle sorte que les applications α et β sont l'identité. En particulier, la conclusion de ce lemme s'écrit le plus souvent $\sigma|_{\mathbb{K}}=\mathrm{Id}$. Il faut juste savoir que le Frido est un névrosé des notations précises.

Démonstration. Comme $\mathbb{L}$ est monogène, si $\mu_a\in\mathbb{K}[X]$ est le polynôme minimal de $a\in\mathbb{L}$, alors les points (5) et (6) de la proposition 6.96 disent que $\mathbb{L}\simeq\mathbb{K}[a]\simeq\mathbb{K}[X]/(\mu_a)$. Pour référence ultérieure, nous considérons un isomorphisme

$$\varphi\colon\mathbb{L}\to\mathbb{K}[X]/(\mu_a).\tag{6.203}$$

Les coefficients de μ_a sont dans $\mathbb{K}$, donc nous pouvons voir $\mu_a\in\mathbb{F}[X]$. Plus précisément, si $\mu_a=\sum_k a_kX^k$, nous définissons

$$\mu'_a=\sum_k\beta(a_k)X^k\in\mathbb{F}[X].\tag{6.204}$$

Comme $\mathbb{F}$ est algébriquement clos, le polynôme μ'_a possède une racine (au moins) $b\in\mathbb{F}$: $\mu'_a(b)=0$.

Nous posons

$$\begin{aligned}\sigma'\colon\mathbb{K}[X]/(\mu_a)&\to\mathbb{F}\\ \overline{\sum_k s_kX^k}&\mapsto\sum_k\beta(s_k)b^k.\end{aligned}\tag{6.205}$$

(i) **σ' est bien définie** Si $P=\sum_k s_kX^k$ et $\mu_a=\sum_k a_kX^k$ $(a_k,s_k\in\mathbb{K})$, alors

$$\sigma'(\overline{P+\mu_a})=\sigma'\big(\overline{\sum_k(a_k+s_k)X^k}\big)=\sum_k\beta(a_k)b^k+\sum_k s_kb^k=\mu'_a(b)+\sigma'(\overline{P})=\sigma'(\overline{P}).\tag{6.206}$$

(ii) **σ' est un morphisme de corps** À justifier.

(iii) $\underline{\varphi\big(\alpha(k)\big) = \bar{k}}$ Quand nous parlons de $\bar{k}$, nous parlons de la classe du polynôme de degré zéro donné par $k \in \mathbb{K}$.

(iv) **La réponse** Nous posons

$$\sigma = \sigma' \circ \varphi. \tag{6.207}$$

Pour tout $k \in \mathbb{K}$,

$$(\sigma'\varphi\alpha)(k) = \sigma'(\bar{k}) = \beta(k), \tag{6.208}$$

c'est ce qu'il fallait.

□

Lemme 6.127 ([149]).
Soit un corps $\mathbb{K}$ muni de deux extensions $\alpha\colon \mathbb{K} \to \mathbb{L}$ et $\beta\colon \mathbb{K} \to \mathbb{F}$. Nous supposons que

(1) $\mathbb{L}$ est algébrique sur $\mathbb{K}$;

(2) $\mathbb{F}$ est algébriquement clos.

Alors il existe un morphisme de corps $\sigma\colon \mathbb{L} \to \mathbb{F}$ tel que $\sigma \circ \alpha = \beta$.

Démonstration. Nous allons utiliser le lemme de Zorn 1.22 sur l'ensemble

$$\mathcal{A} = \left\{ (\mathbb{M}, \varphi) \text{ tel que } \begin{cases} \mathbb{M} \text{ est un sous-corps de } \mathbb{L} \\ \alpha(\mathbb{K}) \subset \mathbb{M} \\ \varphi\colon \mathbb{M} \to \mathbb{F} \text{ est une extension de corps} \\ \varphi \circ \alpha = \beta \end{cases} \right\}. \tag{6.209}$$

Nous ordonnons (partiellement) cet ensemble en disant que $(\mathbb{M}_1, \varphi_1) < (\mathbb{M}_2, \varphi_2)$ si $\mathbb{M}_1 \subset \mathbb{M}_2$ et $\varphi_2|_{\mathbb{M}_1} = \varphi_1$. Il se fait que $\mathcal{A}$ est un ensemble inductif et que nous pouvons donc lui appliquer le lemme de Zorn.

Soit un élément maximal $(\mathbb{M}, \varphi)$. Nous allons montrer que $\mathbb{M} = \mathbb{L}$.

Soit $l \in \mathbb{L}$. Puisque $\mathbb{L}$ est une extension algébrique de $\mathbb{K}$, il existe un polynôme P à coefficients dans $\alpha(\mathbb{K})$ tel que $P(l) = 0$. Mais comme $\alpha(\mathbb{K}) \subset \mathbb{M}$, ce polynôme est également à coefficients dans $\mathbb{M}$. Donc l est un élément algébrique sur $\mathbb{M}$.

Nous pouvons donc considérer le triplet $\big(\mathbb{M}, \mathbb{M}(l), \mathbb{F}\big)$ qui vérifie les hypothèses du lemme 6.126. Il existe donc un morphisme de corps $\sigma\colon \mathbb{M}(l) \to \mathbb{F}$ tel que $\sigma|_{\mathbb{M}} = \varphi$.

Nous avons

$$\sigma \circ \alpha = \sigma|_{\sigma(\mathbb{K})} \circ \alpha = \sigma|_{\mathbb{M}} \circ \alpha = \varphi \circ \alpha = \beta. \tag{6.210}$$

Donc l'élément $\big(\mathbb{M}(l), \sigma\big)$ majore $(\mathbb{M}, \varphi)$ dans $\mathcal{A}$.

Par maximalité, nous déduisons que $\mathbb{M} = \mathbb{L}$. Donc le morphisme $\varphi\colon \mathbb{L} \to \mathbb{F}$ vérifie $\varphi \circ \alpha = \beta$, ce qu'il nous fallait. □

Théorème 6.128 (Steinitz[150, 151]).
À propos de clôture algébrique.

(1) Tout corps possède une clôture algébrique.

(2) Si $\alpha_1\colon \mathbb{K} \to \mathbb{F}_1$ et $\alpha_2\colon \mathbb{K} \to \mathbb{F}_2$ sont deux clôtures algébriques du même corps $\mathbb{K}$, alors il existe un isomorphisme de corps $\varphi\colon \mathbb{F}_1 \to \mathbb{F}_2$ tel que $\varphi \circ \alpha_1 = \alpha_2$.

Démonstration. Nous commençons par l'existence, en plusieurs points.

(i) **Un ensemble** Nous considérons un ensemble Ω qui contient $\mathbb{K}$, qui est strictement surpotent[52] à $\mathbb{K}$ et qui est infini non dénombrable si $\mathbb{K}$ est fini. Par exemple $\mathcal{P}(\mathbb{K}) \cup \mathbb{K}$ si $\mathbb{K}$ est infini et $\mathbb{R} \cup \mathbb{K}$ si $\mathbb{K}$ est fini (voir le théorème de Cantor 1.142).

52. Définition 1.109.

(ii) **L'ensemble pour Zorn** Nous considérons l'ensemble des extensions algébriques de $\mathbb{K}$ contenues dans Ω, c'est-à-dire

$$\mathcal{A} = \left\{ (\mathbb{L}, +_{\mathbb{L}}, \times_{\mathbb{L}}) \text{ tel que } \begin{cases} \mathbb{K} \subset \mathbb{L} \subset \Omega \\ (\mathbb{L}, +_{\mathbb{L}}, \times_{\mathbb{L}}) \text{ est une extension algébrique de } (\mathbb{K}, +, \times). \end{cases} \right\} \tag{6.211}$$

Nous ordonnons $\mathcal{A}$ par l'inclusion : nous disons que

$$(\mathbb{L}_1, +_1, \times_1) < (\mathbb{L}_2, +_2, \times_2) \tag{6.212}$$

lorsque $(\mathbb{L}_2, +_2, \times_2)$ est un sur-coprs de $(\mathbb{L}_1, +_1, \times_1)$ (en particulier $\mathbb{L}_1 \subset \mathbb{L}_2$).

(iii) **$\mathcal{A}$ est inductif** Soit une partie $\mathcal{F} = \{(\mathbb{L}_i, +_i, \times_i)\}_{i \in I}$ de $\mathcal{A}$ que nous supposons être totalement ordonnée. Nous allons lui trouver un majorant dans $\mathcal{A}$. Nous posons $\mathbb{L} = \bigcup_{i \in I} \mathbb{L}_i$, et si $a \in \mathbb{L}_1$, $b \in \mathbb{L}_j$, alors nous définissons

$$\begin{cases} a +_{\mathbb{L}} b = a +_k b & (6.213a) \\ a \times_{\mathbb{L}} b = a \times_k b & (6.213b) \end{cases}$$

où $k \in I$ est sélectionné de telle façon à avoir $(\mathbb{L}_i, +_i, \times_i) < (\mathbb{L}_k, +_k, \times_k)$ et $(\mathbb{L}_j, +_j, \times_j) < (\mathbb{L}_k, +_k, \times_k)$. Comme tous les corps L_i sont des sous-corps les uns des autres, c'est une bonne définition.

(iv) **Lemme de Zorn** Nous utilisons le lemme de Zorn 1.22. Nous notons $(\mathbb{F}, +, \times)$ un élément maximal de $\mathcal{A}$. Puisque $\mathbb{K}$ en est un sous-corps, il n'y a pas d'ambiguïté de noter $+$ et $\times$ ses opérations.

(v) **Stratégie pour la suite** Nous allons montrer que si $\mathbb{E}$ est une extension algébrique de $\mathbb{F}$, alors $\mathbb{E} = \mathbb{F}$ (le but est d'utiliser le lemme 6.119).

(vi) **Un peu de cardinalité** D'abord, comme $\mathbb{F}$ est algébrique sur $\mathbb{K}$, l'ensemble $\mathbb{F}$ est équipotent à $\mathbb{K}$ si $\mathbb{K}$ est infini, et au plus dénombrable, si $\mathbb{K}$ est fini ; c'est la proposition 6.124. En bref :

— Si $\mathbb{K}$ est infini, $\mathbb{K} \approx \mathbb{F} \approx \mathbb{E} \prec \Omega$.

— Si $\mathbb{K}$ est fini, $\mathbb{K} \preceq \mathbb{F} \preceq \mathbb{E} \prec \Omega$ où $\mathbb{E}$ est au maximum dénombrable et Ω est indénombrable.

Dans tous les cas, Ω est strictement surpotent à $\mathbb{F}$, et le lemme 1.151 permet de dire

$$\mathbb{E} \backslash \mathbb{F} \preceq \mathbb{E} \prec \Omega \approx \Omega \backslash \mathbb{F}. \tag{6.214}$$

(vii) **Quelques injections** Il existe donc une injection $\varphi \colon \mathbb{E} \backslash \mathbb{F} \to \Omega \backslash \mathbb{F}$. Nous posons

$$\begin{aligned} f \colon \mathbb{E} &\to \Omega \\ x &\mapsto \begin{cases} x & \text{si } x \in \mathbb{F} \\ \varphi(x) & \text{si } x \notin \mathbb{F}. \end{cases} \end{aligned} \tag{6.215}$$

Nous montrons que f est injective. Soient $x, y \in \mathbb{E}$ tels que $f(x) = f(y)$. Si $x, y \in \mathbb{F}$, alors $x = f(x) = f(y) = y$. Si $x \in \mathbb{F}$ et $y \notin \mathbb{F}$, alors $x = \varphi(y)$ alors que $x \in \mathbb{F}$ et $\varphi(y) \in \Omega \backslash F$; ce cas est impossible. Enfin si x et y sont hors de $\mathbb{F}$, alors $f(x) = \varphi(x)$ et $f(y) = \varphi(y)$; donc $\varphi(x) = \varphi(y)$ et $x = y$ par injectivité de φ.

Nous avons donc bien une injection $f \colon \mathbb{E} \to \Omega$.

(viii) **La maximalité** Nous pouvons mettre sur $f(\mathbb{E}) \subset \Omega$ la structure de corps venant de $\mathbb{E}$. Comme $f(\mathbb{F}) = \mathbb{F}$, le corps $f(\mathbb{E})$ est une extension algébrique de $\mathbb{F}$. Par maximalité, $f(\mathbb{E}) = \mathbb{F}$.

Mais si $x \in \mathbb{E} \backslash \mathbb{F}$, alors $f(x) \in \Omega \backslash \mathbb{F}$. Donc en réalité nous avons aussi $\mathbb{E} \subset \mathbb{F}$.

(ix) **Conclusion** En conclusion $\mathbb{E} = \mathbb{F}$ et le lemme 6.119 termine en disant que $\mathbb{F}$ est une clôture algébrique de $\mathbb{K}$.

Nous passons à la partie « unicité » de la clôture algébrique. Étant donné que $\mathbb{F}_1$ est une extension algébrique de $\mathbb{K}$ et que $\mathbb{F}_2$ est algébriquement clos, le lemme 6.127 nous donne un morphisme de corps $\sigma\colon \mathbb{F}_1 \to \mathbb{F}_2$ tel que $\sigma \circ \alpha_1 = \alpha_2$.

Nous sommes donc dans la situation où $\sigma\colon \mathbb{F}_1 \to \mathbb{F}_2$ est une extension de corps où $\mathbb{F}_1$ est algébriquement clos et $\mathbb{F}_2$ est algébrique. Le lemme 6.74 conclut que $\sigma(\mathbb{F}_1) = \mathbb{F}_2$, c'est-à-dire que σ est surjectif. En tant que morphisme de corps, σ était déjà injective ; elle est donc bijective.

Donc $\sigma\colon \mathbb{F}_1 \to \mathbb{F}_2$ est un isomorphisme de corps vérifiant $\sigma \circ \alpha_1 = \alpha_2$. □

6.129.
Bien que $\mathbb{C}$ soit une extension algébriquement close de $\mathbb{Q}$, l'ensemble $\mathbb{C}$ n'est pas une clôture algébrique de $\mathbb{Q}$. C'est ce que nous montrons maintenant.

Lemme 6.130.
Le corps $\mathbb{C}$ n'est pas une clôture algébrique[53] *de $\mathbb{Q}$.*

Démonstration. Nous montrons qu'il existe des éléments de $\mathbb{C}$ qui ne sont pas des racines de polynômes à coefficients rationnels. L'ensemble $\mathbb{Q}$ est dénombrable par la proposition 1.330. L'ensemble des polynômes de degré n à coefficients dans $\mathbb{Q}$ est en bijection avec les n-uples de rationnels, c'est-à-dire avec $\mathbb{Q}^n$ qui est également dénombrable par la proposition 1.134. Enfin l'ensemble des polynômes à coefficients sur $\mathbb{Q}$ est l'union des polynômes de degré fixés, donc dénombrable par la proposition 1.132.

Jusqu'ici nous avons prouvé que l'ensemble des polynômes à coefficients rationnels était dénombrable. Or chaque polynôme possède une quantité finie de racines par le corolaire 3.149. Donc la partie de $\mathbb{C}$ constituée des nombres qui sont des racines de polynômes à coefficients dans $\mathbb{Q}$ est dénombrable. Mais $\mathbb{C}$ n'est pas dénombrable, donc possède des éléments qui ne sont pas des racines de polynômes. □

6.4.9 Polynômes à plusieurs variables

Nous avons déjà vu $A[X, Y]$ lorsque A est un anneau en la définition 3.53.

Définition 6.131.
Soit un corps $\mathbb{K}$. Le corps $\mathbb{K}(X_1, \dots, X_n)$ est le corps des fractions de l'anneau $\mathbb{K}[X_1, \dots, X_n]$.

Définition 6.132.
Soient un corps $\mathbb{K}$ et une extension $\mathbb{L}$ de $\mathbb{K}$ contenant les éléments $\alpha_1, \dots, \alpha_n$ de $\mathbb{K}$. Nous définissons $\mathbb{K}(\alpha_1, \dots, \alpha_n)$ comme étant l'intersection de tous les sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ et les α_i.

La proposition suivante est analogue à 6.91(2).

Lemme 6.133 ([1])**.**
Soient un corps $\mathbb{K}$, une extension $\mathbb{L}$ et des éléments $\alpha_1, \dots, \alpha_n$ dans $\mathbb{L}$. Alors

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \{r(\alpha_1, \dots, \alpha_n) \text{ tel que } r \in \mathbb{K}(X_1, \dots, X_n)\}. \tag{6.216}$$

Démonstration. Ce que nous avons à droite est un corps : par exemple pour l'inverse, si $r = P/Q$ alors $r(\alpha_1, \dots, \alpha_n) = P(\alpha_1, \dots, \alpha_n)Q(\alpha_1, \dots, \alpha_n)^{-1}$. Cet élément a un inverse en la fraction $(Q/P)(\alpha_1, \dots, \alpha_n)$.

Donc à droite nous avons un sous-corps de $\mathbb{L}$ contenant $\mathbb{K}$ ainsi que les α_i. Donc

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) \subset \big\{r(\alpha_1, \dots, \alpha_n) \text{ tel que } r \in \mathbb{K}(X_1, \dots, X_n)\big\}. \tag{6.217}$$

53. Clôture algébrique, définition 6.70.

D'autre part, tout corps contenant $\mathbb{K}$ et les α_i doit contenir tous les $P(\alpha_1, \dots, \alpha_n)$ ($P \in \mathbb{K}[X_1, \dots, X_n]$), leurs inverses ainsi que leurs produits ; bref doit contenir tous les $r(\alpha_1, \dots, \alpha_n)$ avec $r \in \mathbb{K}[X_1, \dots, X_n]$. □

6.4.10 Corps de décomposition

Définition 6.134.
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 $\mathbb{L}$ étend $\mathbb{K}$ par toutes les racines de tous les polynômes de F.

Proposition 6.135 ([152])**.**
Tout polynôme admet un corps de décomposition. Plus précisément, soit un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$ de degré n. Il existe un corps de décomposition $\mathbb{D}$ de la forme $\mathbb{D} = \mathbb{K}(\alpha_1, \dots, \alpha_n)$ où les α_i sont des racines de P dans $\mathbb{D}$.

Notons que rien dans l'énoncé ne prétend que les α_i soient tous distincts, ni même que certains (voire tous) ne seraient pas dans $\mathbb{K}$.

Démonstration. Soient un corps $\mathbb{K}$ et un polynôme $P \in \mathbb{K}[X]$. Si le degré de P est 0 ou 1, alors $\mathbb{K}$ est un corps de décomposition pour P. Pour le reste nous faisons une récurrence sur le degré de P.

Il y a deux possibilités, soit il existe $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$, soit non.

(i) **Si racine dans $\mathbb{K}$** Alors le corolaire 6.103 nous permet de factoriser $X - \alpha$:

$$P = (X - \alpha)Q \tag{6.218}$$

avec $\deg(Q) = \deg(P) - 1$. Dans ce cas, $\mathbb{K}$ est un corps de rupture de P.

(ii) **Si pas de racines dans $\mathbb{K}$** Nous prenons alors un corps de rupture $\mathbb{L} = \mathbb{K}(\alpha)$ avec $P(\alpha) = 0$ (c'est la proposition 6.113 qui donne l'existence d'un corps de rupture). Dans $\mathbb{L}_1$ nous avons

$$P = (X - \alpha)Q \tag{6.219}$$

avec $Q \in \mathbb{L}_1[X]$ et $\deg(Q) = \deg(P) - 1$.

(iii) **Dans les deux cas** Dans les deux cas, par hypothèse de récurrence nous avons un corps de décomposition pour Q qui se présente sous la forme

$$\mathbb{L} = \mathbb{K}(\alpha_1, \dots, \alpha_{n-1}). \tag{6.220}$$

De plus, $\mathbb{L}$ est une extension de $\mathbb{L}_1$ parce que c'est une extension du corps sur lequel est Q.

Pour terminer la preuve nous prouvons que

$$\mathbb{D} = \mathbb{K}(\alpha_1, \dots, \alpha_{n-1}, \alpha) \tag{6.221}$$

est un corps de décomposition de P. Vu que $\mathbb{D}$ contient $\mathbb{K}(\alpha)$ (comme cas particulier du lemme 6.133), dans $\mathbb{D}$ nous avons l'égalité $P = (X - \alpha)Q$. Et vu que $\mathbb{D}$ contient également $\mathbb{K}(\alpha_1, \dots, \alpha_{n-1})$, toujours dans $\mathbb{D}$ nous avons aussi

$$Q = (X - \alpha_1) \dots (X - \alpha_{n-1}). \tag{6.222}$$

Donc nous avons dans $\mathbb{D}$ l'égalité

$$P = (X - \alpha)(X - \alpha_1) \dots (X - \alpha_{n-1}). \tag{6.223}$$

□

Lemme 6.136 ([1]).
Soit un polynôme $P \in \mathbb{K}[X]$, et un corps $\mathbb{L}$ dans lequel P est scindé. Si $P = P_1 \dots P_r$ est la décomposition de P en irréductibles dans $\mathbb{K}$, alors chacun des P_i est scindé dans $\mathbb{L}$.

Démonstration. Juste pour le mentionner, le fait que P ait une décomposition en irréductibles est le fait que $\mathbb{K}[X]$ soit factoriel, c'est-à-dire la proposition 6.32.

Le polynôme P est scindé dans $\mathbb{L}$, c'est-à-dire que, en notant n le degré de P,

$$P = \prod_{i=1}^{n} (X - \lambda_i) \tag{6.224}$$

avec $\lambda_i \in \mathbb{L}$.

Soit $\mathbb{L}_1$, une extension de $\mathbb{L}$ dans laquelle P_1 est scindé. Ensuite, $\mathbb{L}_2$ une extension de $\mathbb{L}_1$ dans laquelle P_2 est scindé et ainsi de suite. Nous avons construit $\mathbb{L}_r$, une extension de $\mathbb{L}$ dans laquelle tous les P_i sont scindés ainsi que P lui-même. Dans ce corps nous avons l'égalité

$$P = \prod_{k=1}^{n} (X - \mu_k) \tag{6.225}$$

où les μ_k sont des éléments des diverses extensions $\mathbb{L}_i$, et sont les racines des P_i. En tout cas, tous sont dans $\mathbb{L}_r$.

Les deux décompositions (6.224) et (6.225) sont des décompositions dans $\mathbb{L}_r[X]$ du polynôme P. Vu que ce dernier est factoriel, en réalité les deux décompositions sont identiques (se souvenir de la définition 3.71), et nous avons $\mu_k \in \mathbb{L}$ pour tout k. Toutes les extensions $\mathbb{L}_i$ sont en réalité triviales, et nous avons $\mathbb{L}_r = \mathbb{L}$.

Bref, tout cela pour dire que les P_i ont toutes leurs racines dans $\mathbb{L}$. □

Théorème 6.137 ([141]).
Soient :

- *un isomorphisme de corps $\tau \colon \mathbb{K} \to \mathbb{K}'$;*
- *un polynôme non constant $P \in \mathbb{K}[X]$ de degré n ;*
- *un corps de décomposition $\mathbb{L}$ de P sur $\mathbb{K}$;*
- *un corps de décomposition $\mathbb{L}'$ de P sur $\mathbb{K}'$;*

Alors τ se prolonge en un isomorphisme $\sigma \colon \mathbb{L} \to \mathbb{L}'$.

Démonstration. Soit m le nombre de racines de P dans $\mathbb{L} \backslash \mathbb{K}$. Nous faisons une récurrence sur m.

Si $m = 0$ alors $\mathbb{K}$ est un corps de rupture de P ; nous avons

$$P = (X - \lambda_1) \dots (X - \lambda_n) \tag{6.226}$$

avec $\lambda_i \in \mathbb{K}$. Dans ce cas nous avons aussi

$$\tau(P) = \big(X - \tau(\lambda_1)\big) \dots \big(X - \tau(\lambda_n)\big) \tag{6.227}$$

avec $\tau(\lambda_i) \in \mathbb{K}'$. Nous avons donc $\mathbb{L}' = \mathbb{K}'$ et prendre $\sigma = \tau$ fonctionne.

Nous supposons à présent que $m > 0$. Plus précisément nous considérons un polynôme possédant exactement m racines dans $\mathbb{L} \backslash \mathbb{K}$. Soit

$$P = P_1 \dots P_r \tag{6.228}$$

sa décomposition en irréductibles dans $\mathbb{K}[X]$ (notons que $r \leqslant n$ parce que chacun des facteurs est de degré au moins 1). Au moins un des P_i est de degré plus grand ou égal à 2. Nous savons de la proposition 6.32 que $\mathbb{K}[X]$ est un anneau factoriel. Le lemme 6.136 nous assure que les polynômes P_i sont également scindés dans $\mathbb{L}$. Et l'unicité de la décomposition fait en sorte que les racines des P_i sont celles de P.

Soit $\alpha \in \mathbb{L}$, une racine de P_1. Vu que P_1 est irréductible sur $\mathbb{K}$, l'application suivante est un isomorphisme de corps par le lemme 6.77 :

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(P_1) &\to \mathbb{K}[\alpha] \\ \bar{Q} &\mapsto Q(\alpha). \end{aligned} \tag{6.229}$$

Notons que le lemme parle du quotient par le polynôme minimal, mais ici nous avons un irréductible. Un polynôme annulateur irréductible est multiple du polynôme minimal, et l'idéal engendré est le même.

Nous avons aussi la décomposition

$$\tau(P) = \tau(P_1)\dots\tau(P_r), \tag{6.230}$$

et chacun des $\tau(P_i)$ a ses racines dans $\mathbb{L}'$. Soit β, une racine de $\tau(P_1)$ dans $\mathbb{L}'$. Alors nous avons l'isomorphisme

$$\psi'\colon \mathbb{K}'[X]/\big(\tau(P_1)\big) \to \mathbb{K}'[\beta]. \tag{6.231}$$

De plus, par le lemme 6.43, nous savons que τ passe aux classes :

$$\phi_\tau\colon \mathbb{K}[X]/(P_1) \to \mathbb{K}'[X]/\big(\tau(P_1)\big) \tag{6.232}$$

est un isomorphisme d'anneaux. Et enfin, dernier résultat externe à invoquer, la proposition 6.96 nous assure que $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$ et $\mathbb{K}'[\beta] = \mathbb{K}'(\beta)$. Posons pour l'occasion $\mathbb{K}_1 = \mathbb{K}(\alpha)$ et $\mathbb{K}'_1 = \mathbb{K}'(\beta)$.

Nous avons l'enchainement suivant d'isomorphismes de corps [54] :

$$\tau_1 = \psi' \circ \phi_\tau \circ \psi^{-1}\colon \mathbb{K}_1 \to \mathbb{K}[X]/(P_1) \to \mathbb{K}'[X]/\big(\tau(P_1)\big) \to \mathbb{K}'_1. \tag{6.233}$$

Cet isomorphisme $\tau_1\colon \mathbb{K}_1 \to \mathbb{K}_2$ prolonge τ. Si vous aimez les diagrammes, en voici un sur lequel les i représentent des inclusions et où τ et τ_1 sont des isomorphismes

$$\begin{tikzcd} \mathbb{K} \arrow[r, "i"] \arrow[d, "\tau"'] & \mathbb{K}_1 \arrow[r, "i"] \arrow[d, "\tau_1"] & \mathbb{L} \\ \mathbb{K}' \arrow[r, "i"'] & \mathbb{K}'_1 \arrow[r, "i"] & \mathbb{L}' \end{tikzcd} \tag{6.234}$$

Le corps $\mathbb{L}$ est un corps de décomposition de P sur $\mathbb{K}_1$, et le nombre de racines de P dans $\mathbb{L}\backslash\mathbb{K}_1$ est strictement plus petit que m parce qu'il y en a exactement m dans $\mathbb{L}\backslash\mathbb{K}$ et que $\mathbb{K}_1$ en a au moins une de plus que $\mathbb{K}$. Même raisonnement pour $\mathbb{K}'$, $\mathbb{K}'_1$ et $\mathbb{L}'$.

Résumons la situation :

— $\tau_1\colon \mathbb{K}_1 \to \mathbb{K}'_1$ est un isomorphisme de corps ;

— $P \in \mathbb{K}_1[X]$ est un polynôme non constant ;

— $\mathbb{L}$ est un corps de décomposition de P sur $\mathbb{K}_1$;

— $\mathbb{L}'$ est un corps de décomposition de P sur $\mathbb{K}'_1$;

— le nombre de racines de P dans $\mathbb{L}\backslash\mathbb{K}_1$ est strictement inférieur à m.

Donc, par hypothèse de récurrence sur m, il existe un isomorphisme $\sigma\colon \mathbb{L} \to \mathbb{L}'$ qui prolonge τ_1. Vu que τ_1 prolonge τ, nous avons également σ qui prolonge τ. □

L'énoncé le plus compact pour l'unicité du corps de décomposition (à isomorphisme près) est le suivant :

Proposition 6.138.
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 $f|_{\mathbb{K}} = \mathrm{Id}$.

54. En réalité il est plus exact de dire « isomorphisme d'anneaux », parce que la structure de corps n'est en réalité aucune nouvelle structure par rapport à l'anneau.

Démonstration. C'est un cas particulier du théorème 6.137, où nous considérons $\mathbb{K} = \mathbb{K}'$ muni de l'isomorphisme identité. □

Cependant le passage par l'énoncé plus compliqué 6.137 est nécessaire pour les besoins de la récurrence.

6.139.
À propos de terminologie. Lorsque nous disons « *un* corps de décomposition » nous référons à la définition 6.134 et il n'y a pas vraiment unicité. Si nous disons « *le* corps de décomposition » nous référons en général à celui construit dans la proposition 6.135 qui n'est en réalité même pas très explicite.

Quoi qu'il en soit, nous nous permettons de dire « le » corps de décomposition lorsque nous parlons de propriétés invariantes par isomorphisme.

6.140.
La construction du corps de décomposition d'un polynôme fonctionne en prenant successivement le corps de rupture des facteurs irréductibles. Nous insistons sur le fait que cette opération se fait sur chaque facteur irréductible séparément.

L'exemple suivant montre dans quel ordre se passent les choses.

Exemple 6.141.
Soit le polynôme $P = X^4 - 5X^2 + 6$. Sa factorisation en irréductibles est :

$$P = (X^2 - 2)(X^2 - 3). \tag{6.235}$$

Ce polynôme n'est pas irréductible sur $\mathbb{Q}$ et il ne s'agit donc pas de prendre brutalement un corps de rupture pour P. Il s'agit de poser $P = P_1P_2$ avec

$$P_1 = X^2 - 2 \tag{6.236a}$$
$$P_2 = X^2 - 3, \tag{6.236b}$$

de remarquer que P_1 et P_2 sont irréductibles sur $\mathbb{Q}$ et de chercher des corps de rupture pour eux. On commence par P_1. Nous avons le corps de rupture $\mathbb{L}_1 = \mathbb{Q}(\sqrt{2})$ et la factorisation

$$P_1 = (X + \sqrt{2})(X - \sqrt{2}). \tag{6.237}$$

Ensuite nous considérons P_2 dans $\mathbb{L}_1[X]$. Ce P_2 est encore irréductible. Nous lui cherchons un corps de rupture, et c'est $\mathbb{L}_2 = \mathbb{L}_1(\sqrt{3})$ dans lequel nous avons

$$P_2 = (X - \sqrt{3})(X + \sqrt{3}). \tag{6.238}$$

Nous savons (par le lemme 6.118) que

$$\mathbb{L}_2 = \mathbb{L}_1(\sqrt{3}) = \big(\mathbb{Q}(\sqrt{2})\big)(\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}). \tag{6.239}$$

Nous pouvons donc écrire en toute confiance, dans $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ la factorisation

$$P = (X + \sqrt{2})(X - \sqrt{2})(X + \sqrt{3})(X - \sqrt{3}). \tag{6.240}$$

Et nous notons que si nous avions commencé par P_2 au lieu de P_1, nous aurions eu le même résultat final. △

Corolaire 6.142 ([1]).
Le corps de décomposition d'un polynôme est une extension finie.

Démonstration. Soient un corps $\mathbb{K}$, un polynôme $P \in \mathbb{K}[X]$ et un corps de décomposition $\mathbb{D}$ de P de la forme $\mathbb{D} = \mathbb{K}(\alpha_1, \dots, \alpha_n)$ où les α_i sont les racines de P dans $\mathbb{D}$. Cela existe par la proposition 6.135.

Vu que le lemme 6.118 donne

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \big(\mathbb{K}(\alpha_1, \dots, \alpha_{n-1})\big)(\alpha_n), \tag{6.241}$$

le corps $\mathbb{D}$ se construit comme une pile d'extensions finies. Les degrés se composant par le lemme 6.59, au final ce corps de décomposition est une extension finie.

Soit maintenant un corps de décomposition quelconque $\mathbb{L}$. La proposition 6.138 donne un isomorphisme de corps[55] $f\colon \mathbb{L} \to \mathbb{D}$ tel que f soit l'identité sur $\mathbb{K}$.

Si $\{v_i\}_{i\in I}$ est une base de $\mathbb{D}$ comme espace vectoriel sur $\mathbb{K}$, êtes-vous prêts à parier que $\{f(v_i)\}_{i\in I}$ est une base de $\mathbb{L}$ comme espace vectoriel sur $\mathbb{K}$[56] ? □

6.4.11 Non irréductible ou pas corps ?

Nous avons déjà mentionné que nous ne définissons le corps de rupture d'un polynôme que dans le cas de polynôme irréductible à coefficients dans un corps.

D'abord si P n'est pas irréductible, la question de chercher un corps de rupture n'a pas beaucoup de sens.

Si A est un anneau intègre et si P est un polynôme irréductible sur A, nous pouvons considérer le corps des fractions de A, dire $P \in \mathrm{Frac}(A)[X]$ et continuer. Étendre la définition de corps de rupture de cette façon aux polynômes à coefficients dans un anneau intègre n'est pas une grande révolution.

Au lieu de cela, nous pouvons nous demander dans quel cas nous aurions que $A[X]/(P)$ est directement un corps.

Exemple 6.143.
Soit le polynôme constant $P = 2$ dans $\mathbb{Z}[X]$. Il y est irréductible parce qu'il ne peut pas être écrit comme produit de deux non inversibles. Ce polynôme a deux propriétés ennuyeuses :

— Il n'est plus irréductible sur $\mathbb{Q}$,

— Il n'existe aucun corps contenant une racine de P tout en contenant $\mathbb{Z}$ comme sous-anneau.

△

6.4.12 Clôture algébrique

Théorème 6.144.
Tout corps $\mathbb{K}$ possède une clôture algébrique[57] *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 6.145.
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. □

Exemple 6.146.
Soit p un nombre premier. Montrons que le polynôme

$$Q(X) = X^p - X + 1 \tag{6.242}$$

55. Un isomorphisme de corps est juste un isomorphisme d'anneaux.
56. Personnellement, je n'ai pas vérifié. Vérifiez vous-même et écrivez-moi pour dire si c'est bon ou non.
57. Définition 6.70.

est irréductible dans $\mathbb{F}_p$.

Nous supposons qu'il n'est pas irréductible, c'est-à-dire que

$$Q(X) = R(X)S(X) \tag{6.243}$$

avec R et S, des polynômes de degrés $\geqslant 1$ dans $\mathbb{F}_p[X]$

Soit $\bar{\mathbb{F}}_p$ une clôture algébrique[58] 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

$$\begin{aligned} Q(\alpha + a) &= (\alpha + a)^p - (\alpha + a) + 1 && (6.244a)\\ &= \alpha^p + a^p - \alpha - a + 1 && (6.244b)\\ &= \alpha^p - \alpha + 1 && (6.244c)\\ &= Q(\alpha) && (6.244d)\\ &= 0 && (6.244e) \end{aligned}$$

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

$$R(X) = \prod_{i=1}^{k} \big(X - (\alpha + a_i)\big) \tag{6.245}$$

où les a_i sont les éléments de $\mathbb{F}_p$. En développant un peu,

$$R(X) = X^k - \sum_{i=1}^{k}(\alpha + a_i^{k-1}) + \text{termes de degré plus bas en } X. \tag{6.246}$$

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 :

$$Q(\alpha) = \alpha^p - \alpha + 1 = 1 \neq 0. \tag{6.247}$$

Le polynôme $X^p - X + 1$ est donc irréductible sur $\mathbb{F}_p$. △

6.4.13 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 6.159.

Définition 6.147.
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, c'est-à-dire que si P est de degré n, alors il possède n racines distinctes dans un corps de décomposition.*

*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 6.148.
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 $\mathbb{L}'$.

58. Définition 6.70. Pour l'existence c'est le théorème 6.144.

Démonstration. L'ingrédient est la proposition 6.138 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

$$P = a(X - \alpha_1)\dots(X - \alpha_n) \tag{6.248}$$

avec $a \in \mathbb{K}$ et $\alpha_i \in \mathbb{L}$. Nous avons donc

$$P = \psi(P) = a(X - \psi(\alpha_1))\dots(X - \psi(\alpha_n)). \tag{6.249}$$

Donc les racines de P dans $\mathbb{L}'$ sont les éléments $\psi(\alpha_i)$ qui sont distincts. □

Exemple 6.149.
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

$$P = X^p - T^p \tag{6.250}$$

dans $\mathbb{K}[X]$. Par le morphisme de Frobenius nous avons

$$P = (X - T)^p \tag{6.251}$$

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 si 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 6.150.
Le polynôme $(X - 1)^3$ est séparable sur $\mathbb{Q}$ parce que ses facteurs irréductibles dans $\mathbb{Q}[X]$ sont $X - 1$ et $X^2 + X + 1$, et ces deux polynômes ont des racines simples (dans $\mathbb{Q}(i)$). △

Exemple 6.151.
Le polynôme $(X^2 + 1)^2$ est séparable dans $\mathbb{Q}[X]$. En effet, il a pour facteurs irréductibles le polynôme $X^2 + 1$ (en deux exemplaires), et ce polynôme a pour racines $\pm i$ dans l'extension $\mathbb{Q}(i)$, racines qui sont simples pour ce polynôme. △

Proposition 6.152 ([153]).
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. (i) **(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

$$P\big(\psi(a)\big) = \psi\big(P(a)\big). \tag{6.252}$$

(ii) **(2)⇒(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 $P = (X-a)^2 Q$ avec $Q \in \mathbb{L}[X]$. En dérivant,

$$P' = 2(X-a)Q + (X-a)^2 Q', \tag{6.253}$$

et donc a est également une racine de P'.

(iii) **(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

$$AP + BP' = D. \tag{6.254}$$

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 $\deg(D) \geqslant 1$.

(iv) **(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

$$P = (X-a)Q, \tag{6.255}$$

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 6.153.
Soit $P \in \mathbb{K}[X]$ irréductible. Le polynôme P est séparable si et seulement si $P' \neq 0$.

Démonstration. Soit $D = \operatorname{pgcd}(P, P')$ et nous voudrions prouver que $\deg(D) \geqslant 1$ si et seulement si $P' = 0$. Si $P' = 0$, alors $\operatorname{pgcd}(P, P') = P$ 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 et le polynôme constant 1. Ainsi, $D \in \mathbb{K}$, ou bien $P = \lambda D$ avec $\lambda \in \mathbb{K}$. et donc $\deg(P) \geqslant 1$. Dans les deux cas, P' est nécessairement non-nul. □

Corolaire 6.154.
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 6.152. □

Définition 6.155.
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 $\mathbb{K}$ (définition 6.147).*

*(2) L'extension $\mathbb{L}$ est **séparable** si tous ses éléments sont séparables.*

Proposition 6.156.
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. En plusieurs parties.

(i) **(1) implique (2)** Soit un polynôme irréductible P de $\mathbb{K}[X]$, et un corps de décomposition $\mathbb{L}$ de P. Cela est une extension algébrique par le corolaire 6.142. Elle est donc séparable par hypothèse.

Voilà une première chose de dite.

Maintenant, nous voudrions montrer que P est un polynôme séparable. Dans $\mathbb{L}$ nous avons

$$P = \prod_{i=1}^{n}(X - a_i), \tag{6.256}$$

et tout le défi est de prouver que les a_i sont tous distincts.

Soient deux racines $a, b \in \mathbb{L}$ de P. Nous considérons les polynômes minimaux μ_a et μ_b dans $\mathbb{K}[X]$. Ces deux polynômes divisent P parce que P est à la fois dans l'idéal annulateur de a et de b. Mais comme P est irréductible, il existe $k_a, k_n \in \mathbb{K}$ tels que $P = k_a\mu_a$ et $P = k_b\mu_b$. Donc les polynômes μ_a, μ_b et P sont multiples les uns des autres. Vu que μ_a et μ_n sont unitaires, $\mu_a = \mu_b$.

Nous avons :

$$P = k\mu = \prod_{i=1}^{n}(X - a_i). \tag{6.257}$$

Or le polynôme μ est irréductible par la proposition 6.63(1), et l'extension $\mathbb{L}$ est séparable, donc μ n'a que des racines simples, Donc tous les a_i sont distincts.

(ii) **(2) implique (1)** Soit une extension algébrique $\mathbb{L}$ de $\mathbb{K}$. Soit $a \in \mathbb{L}$. Nous devons prouver que le polynôme minimal de a dans $\mathbb{K}$ est séparable, c'est-à-dire qu'il n'a que des racines simples.

Le polynôme minimal $\mu_a \in \mathbb{K}[X]$ de a est irréductible et donc séparable. Donc $\mathbb{L}$ est séparable.

La dernière phrase est une conséquence du corolaire 6.154. □

Corolaire 6.157.
Toute les extensions algébriques de $\mathbb{Q}$ sont séparables.

Démonstration. Le corps $\mathbb{Q}$ est de caractéristique nulle (définition 1.294). Le corolaire 6.154 dit alors que tout polynôme sur $\mathbb{Q}$ est séparable. La proposition 6.156 conclut en disant que toutes les extensions algébriques de $\mathbb{Q}$ sont séparables. □

Théorème 6.158 ([86])**.**
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

$$x^\omega = 1 \tag{6.258}$$

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 3.33. □

Théorème 6.159 (Théorème de l'élément primitif[86])**.**
Toute extension de corps séparable finie admet un élément primitif[59].

Autrement dit, soient des éléments algébriques $\alpha_1, \dots, \alpha_n$ séparables[60] *sur $\mathbb{K}$, et soit l'extension engendrée $\mathbb{L} = \mathbb{K}(\alpha_1, \dots, \alpha_n)$. Alors $\mathbb{L}$ admet un élément primitif, c'est-à-dire un élément θ tel que $\mathbb{L} = \mathbb{K}(\theta)$.*

59. Définition 6.86.
60. Définition 6.155(1).

Démonstration. Si le corps $\mathbb{K}$ est fini, alors $\mathbb{L}$ est également fini. Donc $\mathbb{L}^*$ est cyclique par le théorème 6.158. 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

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \mathbb{K}(\alpha_1, \dots, \alpha_{n-1})(\alpha_n), \tag{6.259}$$

et donc si $\mathbb{K}(\alpha_1, \dots, \alpha_{n-1}) = \mathbb{K}(\theta)$, nous avons

$$\mathbb{K}(\alpha_1, \dots, \alpha_n) = \mathbb{K}(\theta, \alpha_n) \tag{6.260}$$

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

$$\alpha_1 = \alpha, \alpha_2, \dots, \alpha_r \tag{6.261}$$

de P dans $\mathbb{E}$ et les racines

$$\beta_1 = \beta, \beta_2, \dots, \beta_s \tag{6.262}$$

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[61] une solution donnée le cas échéant par

$$x = (\alpha_i - \alpha_1)(\beta_1 - \beta_k)^{-1} \tag{6.263}$$

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 (6.263) :

$$\alpha_i + c\beta_k \neq \alpha_1 + c\beta_1. \tag{6.264}$$

Nous posons $\theta = \alpha_1 + c\beta_1$ et nous prétendons que $\mathbb{L} = \mathbb{K}(\theta)$.

Pour cela, commençons par montrer que $\beta_1 \in \mathbb{K}(\theta)$. On considère, dans $\mathbb{K}(\theta)[T]$, les polynômes $Q(T)$ et $S(T) = P(\theta - cT)$, et on nomme R le PGCD de ces deux polynômes. Alors, une racine de R doit être une racine de Q, et est donc un des β_i. Or, d'une part, le choix de θ fait que β_1 est une racine de R parce que

$$S(\beta_1) = P(\theta - c\beta_1) = P(\alpha_1 + c\beta_1 - c\beta_1) = P(\alpha_1) = 0. \tag{6.265}$$

D'autre part, si $k \geqslant 2$, alors

$$S(\beta_k) = P(\alpha_1 + c\beta_1 - c\beta_k) = P\big(\alpha_1 + c(\beta_1 - \beta_k)\big) \neq 0 \tag{6.266}$$

parce que $\alpha_1 + c(\beta_1 - \beta_k)$ ne vaut ni α_1 (le second terme est non-nul), ni un autre α_i (à cause de (6.264)).

Il s'ensuit que Q et S n'ont qu'une racine commune $\beta_1 = \beta$, qui est donc l'unique racine de R. Ainsi,

$$R = X - \beta \in \mathbb{K}(\theta)[T], \tag{6.267}$$

et donc $\beta \in \mathbb{K}(\theta)$.

Dès lors $\alpha = \alpha_1 = \theta - c\beta$ est alors immédiatement dans $\mathbb{K}(\theta)$; puisque les deux éléments α et β sont dans $\mathbb{K}(\theta)$, nous avons obtenu $\mathbb{L} = \mathbb{K}(\alpha, \beta) = \mathbb{K}(\theta)$.

□

61. La solution (6.263) peut être dans $\mathbb{L}$ et non dans $\mathbb{K}$. L'équation peut donc très bien ne pas avoir de solutions $x \in \mathbb{K}$.

Exemple 6.160.
Le théorème de l'élément primitif 6.159 ne tient pas pour les corps non commutatifs. Considérons par exemple le corps $\mathbb{K}$ des quaternions et le groupe à 8 éléments $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Ce dernier groupe n'est pas cyclique alors qu'il est un groupe fini dans $\mathbb{K}^*$. △

Exemple 6.161.
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. △

6.5 Idéal maximum

6.5.1 Idéal maximum

Définition 6.162.
*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).*

Théorème 6.163 ([154])**.**
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 6.164 ([154])**.**
Si $\mathbb{K}$ est un corps algébriquement clos, les idéaux maximaux de $\mathbb{K}[X_1, \dots, X_n]$ sont de la forme

$$(X_1 - a_1, \dots, X_n - a_n) \tag{6.268}$$

où les a_i sont des éléments de $\mathbb{K}$.

Démonstration. Nous commençons par montrer que

$$J = (X_1 - a_1, \dots, X_n - a_n) \tag{6.269}$$

est un idéal maximum. Pour cela nous considérons le morphisme surjectif d'anneaux

$$\begin{aligned} \phi\colon \mathbb{K}[X_1, \dots, X_n] &\to \mathbb{K} \\ P &\mapsto P(a_1, \dots, a_n). \end{aligned} \tag{6.270}$$

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 :

$$P = (X - a_1)Q_1 + \cdots + (X_n - a_n)Q_n + R \tag{6.271}$$

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 (6.271) à $(a_1, \dots, a_n)$ et en nous rappelant que $P \in \ker(\phi)$ nous obtenons

$$0 = P(a_1, \dots, a_n) = R, \tag{6.272}$$

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

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{J} = \mathbb{K}. \tag{6.273}$$

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

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} \tag{6.274}$$

est une $\mathbb{K}$-algèbre de type fini (définition 6.162). De plus c'est un corps par la proposition 3.52. C'est donc une extension algébrique finie de $\mathbb{K}$ par le théorème 6.163. Mais $\mathbb{K}$ étant algébriquement clos, il est sa propre et unique extension algébrique ; nous en déduisons que

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} = \mathbb{K}. \tag{6.275}$$

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 $I = J$. □

Corolaire 6.165.
Soit $\mathbb{K}$ un corps algébriquement clos et I, un idéal de $\mathbb{K}[X_1, \dots, X_n]$. Si nous notons

$$V(I) = \{x \in \mathbb{K}^n \text{ tel que } P(x_1, \dots, x_n) = 0\} \tag{6.276}$$

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 6.164 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

$$P(a_1, \dots, a_n) = 0, \tag{6.277}$$

c'est-à-dire que $(a_1, \dots, a_n) \in V(I)$ et donc que $V(I) \neq \varnothing$. □

6.6 Polynômes symétriques et alternés

6.6.1 Polynômes symétriques, alternés ou semi-symétriques

Nous rappelons que le groupe symétrique S_n agit sur l'anneau des polynômes de n variables sur l'anneau A par le lemme 1.312.

Définition 6.166.
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.

6.6.2 Polynôme symétrique élémentaire

Définition 6.167.
Le k-ième ***polynôme symétrique élémentaire*** *à n inconnues est le polynôme*

$$\sigma_k(T_1, \dots, T_n) = \sum_{s \in F_k} \prod_{i=1}^{k} T_{s(i)} \tag{6.278}$$

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

$$\sigma_k = \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} X_{i_1}\ldots X_{i_k}. \tag{6.279}$$

Par exemple

$$\sigma_1(T_1,\ldots,T_n) = T_1+T_2+\cdots+T_n \tag{6.280a}$$
$$\sigma_2(T_1,\ldots,T_n) = T_1T_2+\cdots+T_1T_n+T_2T_3+\cdots+T_2T_n+\cdots+T_{n-1}T_n \tag{6.280b}$$
$$\sigma_n(T_1,\ldots,T_n) = T_1\ldots T_n. \tag{6.280c}$$

En particulier, $\sigma_2(x,y,z) = xy+yz+xz$.

Théorème 6.168 ([155]).
Si Q est un polynôme symétrique en $T_1,\ldots,T_n$, alors il existe un et un seul polynôme P en n indéterminées tel que

$$Q(T_1,\ldots,T_n) = P\big(\sigma_1(T_1,\ldots,T_n),\ldots,\sigma_n(T_1,\ldots,T_n)\big). \tag{6.281}$$

Exemple 6.169.
Nous voulons décomposer $P(x,y,z) = x^3+y^3+z^3$ en polynômes symétriques élémentaires, c'est-à-dire en

$$\begin{cases} \sigma_1 = x+y+z & (6.282a)\\ \sigma_2 = xy+xz+yz & (6.282b)\\ \sigma_3 = xyz. & (6.282c)\end{cases}$$

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 :

```
----------------------------------------------------------------------
| 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

$$P = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \tag{6.283}$$

△

Lemme 6.170 ([101]).
Soit $\mathbb{K}$ une extension de degré δ de $\mathbb{Q}$ et $P \in \mathbb{K}[T_1, \dots, T_m]$. Alors il existe $\bar{P} \in \mathbb{Q}[T_1, \dots, T_m]$ tel que

(1) $\deg \bar{P} = \delta \deg(P)$

(2) pour tout $(z_1, \dots, z_m) \in \mathbb{C}^m$ tel que $P(z_1, \dots, z_m) = 0$, on a $\bar{P}(z_1, \dots, z_m) = 0$.

Démonstration. En vertu de la proposition 6.156 et du corolaire 6.157, $\mathbb{K}$ est une extension séparable de $\mathbb{Q}$, et donc vérifie le théorème de l'élément primitif (6.159). 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

$$\{1, \theta, \dots, \theta^{\delta-1}\}. \tag{6.284}$$

Mais par ailleurs la proposition 6.96(2) nous indique qu'une base de $\mathbb{Q}(\theta)$ sur $\mathbb{Q}$ est donnée par

$$\{1, \theta, \dots, \theta^{n-1}\} \tag{6.285}$$

où n est le degré de P_θ. Donc P_θ est de degré δ. Nous nommons $\theta_1, \dots, \theta_\delta$ 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

$$\begin{aligned} \sigma_k \colon \mathbb{Q}(\theta) &\to \mathbb{Q}(\theta_k) \\ \sum_i q_i\theta^i &\mapsto \sum_i q_i\theta_k^i \end{aligned} \tag{6.286}$$

Nous avons $\sigma_1 \colon \mathbb{K} \to \mathbb{K}$ 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

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_{il} T_i^l \tag{6.287}$$

avec $c_{il} \in \mathbb{K}$. Nous voyons $c_{i,.}$ comme un élément de $\mathbb{K}^m$ et donc nous écrivons[62]

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_l(\theta)_i T_i^l \tag{6.288}$$

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

$$P^{\sigma_k} = \sum_{l,i} c_l(\theta_k)_i T_i^l \in \mathbb{Q}(\theta_k)[T_1, \dots, T_m], \tag{6.289}$$

et $\bar{P} = PP^{\sigma_2} \dots P^{\sigma_k}$. Le coefficient de T_i^l dans $\bar{P}$ est

$$\bar{c}_l(\theta_1, \dots, \theta_\delta)_i = \sum_{l_1 + \dots + l_\delta = l} c_{l_1}(\theta_1)_i \dots c_{l_\delta}(\theta_\delta)_i. \tag{6.290}$$

62. Il me semble qu'il manque la somme sur i dans [101].

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 6.168 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

$$\deg(\bar{P}) = \delta \deg(P) \tag{6.291}$$

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. □

6.6.3 Relations coefficients racines

Théorème 6.171 (Relations coeffitients-racines)**.**
Soit le polynôme $P = a_nX^n + \cdots + a_1X + a_0$ et r_i ses n racines. Alors nous avons pour chaque $1 \leqslant k \leqslant n$ la relation

$$\sigma_k(r_1, \dots, r_n) = (-1)^k \frac{a_{n-k}}{a_n} \tag{6.292}$$

où σ_k est le k^epolynôme symétrique défini en 6.167.

Exemple 6.172.
Soit le polynôme

$$P(x) = x^3 + 2x^2 + 3x + 4 \tag{6.293}$$

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[63] nous donnent $\sigma_1(a, b, c) = -2$ et $\sigma_2(a, b, c) = 3$, donc

$$a^2 + b^2 + c^2 = (-2)^2 - 2 \cdot 3 = -2. \tag{6.294}$$

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) :

```
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 6.173 (Annulation de fonctions polynomiales[156])**.**
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 $\mathrm{Card}(A_j) > \deg_{X_j}(P)$ pour tout j. Alors $P = 0$.

63. Théorème 6.171

De plus si $P = 0$ alors tous ses coefficients sont nuls[64].

Démonstration. Nous prouvons le résultat par récurrence sur le nombre n d'indéterminées. Si $n = 1$, cela est le théorème 6.104. Nous classons les monômes du polynôme P par ordre de puissance de X_n et nous le factorisons :

$$P = \sum_{i=1}^{m} P_i X_n^i \tag{6.295}$$

avec $P_i \in \mathbb{K}[X_1, \dots, X_{n-1}]$. Soit $(a_1, \dots, a_{n-1}) \in A_1 \times \dots \times A_{n-1}$ et posons

$$Q(T) = P(a_1, \dots, a_{n-1}, T) = \sum_{i=1}^{m} P_i(a_1, \dots, a_{n-1}) T^i. \tag{6.296}$$

Le polynôme Q s'annule sur A_n avec $\deg(Q) = \deg_{X_n}(P) < \mathrm{Card}(A_n)$ et le théorème 6.104 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 $A_1 \times \dots \times A_{n-1}$, mais nous avons

$$\deg_{X_j}(P_i) \leqslant \deg_{X_j} P < \mathrm{Card}(A_j), \tag{6.297}$$

donc l'hypothèse de récurrence donne $P_i = 0$. Par suite, $P = 0$ également. □

6.7 Minuscule morceau sur la théorie de Galois

Vous trouverez des détails et des preuves à propos de la théorie de Galois dans [157, 86].

Définition 6.174.
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 6.175.
*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*

$$\sum \alpha_{i_1 \dots i_n} b_1^{i_1} \dots b_n^{i_n} = 0 \tag{6.298}$$

avec $\alpha_{i_1 \dots i_n} \in \mathbb{K}$.

Nous disons que l'équation

$$x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0 \tag{6.299}$$

est l'**équation générale** de degré n si les coefficients a_i sont algébriquement indépendants sur $\mathbb{K}$.

Théorème 6.176.
Le groupe de Galois d'un polynôme de degré n est isomorphe au groupe symétrique S_n.

Corolaire 6.177 ([158])**.**
L'équation générale de degré n est résoluble par radicaux si et seulement si $n \leqslant 5$.

64. 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.

Chapitre 7

Topologie générale

7.1 Éléments généraux de topologie

7.1.1 Définitions et propriétés de base

Définition 7.1 ([159]).
Soient X, un ensemble et $\mathcal{T}$, une partie de l'ensemble de ses parties qui vérifie les propriétés suivantes.

(1) Les ensembles $\varnothing$ et X 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 X est une **topologie** sur X, et les éléments de $\mathcal{T}$ sont appelés des **ouverts**. On dit aussi que $(X, \mathcal{T})$ (voire simplement X lorsqu'il n'y a pas d'ambiguïté) est un **espace topologique**.*

Deux espaces topologiques sont isomorphes quand il existe une bijection continue d'inverse continue. Nous verrons ça en la définition 7.44.

7.1.2 Base de topologie

Proposition-Définition 7.2 (Base de topologie[160, 161]).
Soit un espace topologique $(X, \mathcal{T})$. Soit une partie $\mathcal{B}$ de $\mathcal{T}$. Les propriétés suivantes sont équivalentes :

(1) Tout élément de $\mathcal{T}$ est une réunion d'éléments de $\mathcal{B}$.

(2) Pour tout $x \in X$ et pour tout ouvert $\mathcal{O}$ contenant x, il existe $B \in \mathcal{B}$ tel que

$$x \in B \subset \mathcal{O}. \tag{7.1}$$

*Une partie $\mathcal{B}$ de $\mathcal{T}$ qui vérifie ces propriétés est une **base de topologie** pour X.*

Démonstration. En deux parties.

(i) **(1) implique (2)** Soient $x \in X$ et $\mathcal{O}$ un ouvert contenant x. Étant donné que $\mathcal{O}$ est une réunion d'éléments de $\mathcal{B}$, il y a au moins un $B \in \mathcal{B}$ contenant x. Ce B vérifie $x \in B \subset \mathcal{O}$.

(ii) **(2) implique (1)** Soit $\mathcal{O}$ un ouvert de X ; pour chaque $x \in \mathcal{O}$ nous considérons un ouvert $B(x) \in \mathcal{B}$ tel que $x \in B(x) \subset \mathcal{O}$. Nous avons alors $\mathcal{O} = \bigcup_{x \in \mathcal{O}} B(x)$.

□

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.

7.1.3 Fermés

Définition 7.3.
Si X est un espace topologique, un sous-ensemble F de X est dit ***fermé*** *si son complémentaire, $X \backslash F$, est ouvert.*

Définition 7.4.
Si $a \in X$, on dit que $V \subset X$ est un ***voisinage*** *de a si il existe un ouvert $\mathcal{O} \in \mathcal{T}$ tel que $a \in \mathcal{O}$ et $\mathcal{O} \subset V$.*

Définition 7.5 (Base de voisinage[162])**.**
Soient un espace topologique X ainsi que $a \in X$. Un ensemble $\{U_i\}_{i \in I}$ de voisinages de a est une ***base de voisinages*** *de a si pour tout voisinage V de a, il existe $k \in I$ tel que $U_k \subset V$.*

Lemme 7.6.
Union et intersection de fermés.

(1) Une intersection quelconque de fermés est fermée.

(2) Une union finie de fermés est fermée.

Démonstration. Soit $\{F_i\}_{i \in I}$ un ensemble de fermés ; nous avons

$$\left(\bigcap_{i \in I} F_i \right)^c = \bigcup_{i \in I} F_i^c. \tag{7.2}$$

Le membre de droite est une union d'ouverts, c'est donc un ouvert ; donc l'intersection qui apparaît dans le membre de gauche est le complémentaire d'un ouvert : c'est donc un fermé.

De la même manière, le complémentaire d'une union finie de fermés est une intersection finie de complémentaires de fermés, et est donc ouvert [2]. □

Dans un espace topologique, nous avons une caractérisation très importante des ouverts.

Théorème 7.7.
Une partie d'un espace topologique est ouverte si et seulement si elle contient un ouvert autour 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 dans X.

Pour le sens inverse, nous supposons que A contienne un ouvert autour de chacun de ses points. Pour chaque $x \in A$, choisissons $\mathcal{O}_x \subset A$ un ouvert autour de x. Alors,

$$A = \bigcup_{x \in A} \mathcal{O}_x \tag{7.3}$$

en effet, d'une part, $A \subset \bigcup_{x \in A} \mathcal{O}_x$ parce que chaque élément x de A est dans le $\mathcal{O}_x$ correspondant, par construction ; et d'autre part, $\bigcup_{x \in A} \mathcal{O}_x \subset A$ parce que chacun des $\mathcal{O}_x$ est inclus dans A.

Ainsi, A est égal à une union d'ouverts, cela prouve que A est un ouvert. □

Le lemme 7.197 est une version particulière de celui-ci, pour l'espace topologique $\mathbb{R}$. Une autre application typique est la proposition 7.2 et le théorème 7.170.

7.1.4 Quelques exemples

7.1.4.1 Une première vague

Exemple 7.8.
Soit un ensemble quelconque X. L'ensemble de parties $\mathcal{T} = \{\emptyset, X\}$ est une topologie sur X.

La topologie ainsi définie sur X est appelée **topologie grossière**. △

2. Un bon exercice est d'écrire ces unions et intersections, pour se convaincre que ça fonctionne.

Exemple 7.9.
Pour un ensemble X quelconque, on considère l'ensemble $\mathcal{T}$ constitué de toutes les parties de X. Avec cet ensemble, on confère à nouveau une structure d'espace topologique à X ; toutes les parties sont des ouverts, et aussi des fermés. La topologie ainsi posée sur X est appelée **topologie discrète**. △

Exemple 7.10 (Toutes les topologies d'un ensemble à 3 éléments)**.**
On pose $X = \{1, 2, 3\}$. Alors on peut munir X de 29 topologies différentes[163] ; saurez-vous les retrouver toutes ? △

7.1.4.2 Topologie engendrée

Proposition-Définition 7.11 (Topologie engendrée, prébase[164])**.**
Soient un ensemble X, et $\mathcal{T}_0$ un ensemble de parties de X. Nous définissons $\tau(\mathcal{T}_0)$ comme étant l'union quelconque d'intersections finies d'éléments de $\mathcal{T}_0$.

Plus précisément, nous faisons les constructions suivantes :

(1) Nous notons $\{\mathcal{O}_i\}_{i\in I}$ les éléments de $\mathcal{T}_0$ indexés par l'ensemble I.

(2) Soit $B(\mathcal{T}_0)$ l'ensemble des intersections finies d'éléments de $\mathcal{T}_0$:

$$B(\mathcal{T}_0) = \{\bigcap_{j\in J} \mathcal{O}_j\}_{J \text{ fini dans } I} \tag{7.4}$$

où nous convenons que $\bigcap_{j\in\emptyset} \mathcal{O}_j = X$ [3].

(3) Soit A un ensemble qui indexe $B(\mathcal{T}_0)$:

$$B(\mathcal{T}_0) = \{B_\alpha\}_{\alpha\in A}. \tag{7.5}$$

(4) Nous posons

$$\tau(\mathcal{T}_0) = \{\bigcup_{\alpha\in S} B_\alpha\}_{S\subset A}. \tag{7.6}$$

Alors :

(1) $\tau(\mathcal{T}_0)$ est une topologie sur X.

(2) Toute topologie sur X contenant $\mathcal{T}_0$ contient $\tau(\mathcal{T}_0)$.

*La topologie $\tau(\mathcal{T}_0)$ est appelée la **topologie engendrée** par $\mathcal{T}_0$. La partie $\mathcal{T}_0$ est appelée **prébase** de la topologie $\tau(\mathcal{T}_0)$.*

Démonstration. Pour (1), nous devons montrer les différents points de la définition 7.1 d'une topologie.

(1) L'ensemble vide est dans $\tau(\mathcal{T}_0)$ parce que $\emptyset = \bigcup_{\alpha\in\emptyset} B_\alpha$. L'ensemble X est également dans $\tau(\mathcal{T}_0)$ parce que $X \in B(\mathcal{T}_0)$.

(2) Soient $\{D_l\}_{l\in L}$ des éléments de $\tau(\mathcal{T}_0)$ indexés par un ensemble L. Pour chaque l nous avons un ensemble $S \subset A$ tel que $D_l = \bigcup_{\alpha\in S_l} B_\alpha$. En posant $S = \bigcup_{l\in L} S_l$ nous avons

$$\bigcup_{l\in L} D_l = \bigcup_{\alpha\in S} B_\alpha \in \tau(\mathcal{T}_0). \tag{7.7}$$

Donc $\tau(\mathcal{T}_0)$ est stable par union quelconque.

(3) Soient D_1 et D_2 des éléments de $\tau(\mathcal{T}_0)$. Nous posons $D_i = \bigcup_{\alpha\in S_i} B_\alpha$. Alors nous avons

$$\bigcup_{\alpha\in S_1} B_\alpha \cap \bigcup_{\beta\in S_2} B_\beta = \bigcup_{\alpha,\beta\in S_1\times S_2} (B_\alpha \cap B_\beta). \tag{7.8}$$

Mais B_α et B_β sont dans $B(\mathcal{T}_0)$. Donc $B_\alpha \cap B_\beta \in B(\mathcal{T}_0)$. Donc (7.8) est une union d'éléments de $B(\mathcal{T}_0)$.

3. Bref, nous mettons X dans $B(\mathcal{T}_0)$.

Au final nous avons prouvé que $\tau(\mathcal{T}_0)$ est une topologie sur X.

Nous démontrons à présent le point (2). Soit une topologie μ sur X contenant $\tau(\mathcal{T}_0)$. Puisque μ est une topologie, les intersections finies d'éléments de μ sont dans μ, donc, en suivant les notations de 7.11, $B(\mathcal{T}_0) \subset \mu$.

Comme toutes les unions d'éléments de μ sont dans μ, l'inclusion de $B(\mathcal{T}_0)$ dans μ implique celle de $\tau(\mathcal{T}_0)$. □

Dès que nous avons une topologie, nous avons une notion de convergence de suite.

Définition 7.12 (Convergence de suite)**.**
Une suite (x_n) d'éléments de X ***converge*** *vers un élément y de X si pour tout ouvert $\mathcal{O}$ contenant y , il existe un $K \in \mathbb{N}$ tel que $k \geqslant K$ implique $x_k \in \mathcal{O}$.*

La proposition suivante montre que vérifier la convergence d'une suite sur une prébase suffit pour vérifier la convergence.

Proposition 7.13.
Soit $\mathcal{T}_0$ un ensemble de parties de l'ensemble X. Soient une suite (x_n) dans X ainsi que $x \in X$. Nous supposons que la suite (x_n) satisfait la propriété suivante : pour tout $A \in \mathcal{T}_0$ tel que $x \in A$, il existe $K \in \mathbb{N}$ tel que $k \geqslant K$ implique $x_k \in A$.

Alors nous avons la convergence de suite[4]

$$x_n \xrightarrow{\left(X,\tau(\mathcal{T}_0)\right)} x. \tag{7.9}$$

Démonstration. Nous considérons la topologie $\tau(\mathcal{T}_0)$ sur X. Soit un ouvert $\mathcal{O}$ contenant x. Nous le décomposons en suivant (à l'envers) la construction de la définition 7.11 :

$$\mathcal{O} = \bigcup_{\alpha \in S} B_\alpha \tag{7.10}$$

avec $B_\alpha \in B(\mathcal{T}_0)$. Donc pour chaque α, il existe un ensemble fini J_α tel que

$$B_\alpha = \bigcap_{j \in J_\alpha} A_j \tag{7.11}$$

avec $A_j \in \mathcal{T}_0$. Puisque $x \in \mathcal{O}$, nous avons un α_0 tel que $x \in B_{\alpha_0}$. Donc $x \in A_j$ pour tous les $j \in J_{\alpha_0}$.

Pour chaque $j \in J_{\alpha_0}$, il existe $K_j \in \mathbb{N}$ tel que $k \geqslant K_j$ implique $x_k \in A_j$. Comme J_{α_0} est un ensemble fini, nous pouvons poser $K = \max_{j \in J_{\alpha_0}} K_j$.

Maintenant, si $k \geqslant K$, nous avons $x_k \in A_j$ pour tout j, et donc $x_k \in B_{\alpha_0}$. Par conséquent, $x_k \in \mathcal{O}$. □

7.1.5 Topologie produit

Définition 7.14 (Produit d'espaces topologiques, thème 26)**.**
Soient $\{(X_i, \tau_i)\}_{i=1,\dots,n}$ des espaces topologiques. Leur ***produit*** *est l'ensemble*

$$X = \prod_{i=1}^{n} X_i \tag{7.12}$$

muni de la topologie

$$\mathcal{T} = \{\mathcal{O} \subset X \text{ tel que } \forall x \in U, \exists U_i \in \tau_i \text{ tel que } x \in U_1 \times \dots \times U_n \subset \mathcal{O}\}. \tag{7.13}$$

Proposition 7.15 (Convergence composante par composante)**.**
Soient des espaces topologiques X_i $(i = 1, \dots, n)$ et une suite $(a_k^{(1)}, \dots, a_k^{(n)})$ dans $X_1 \times \dots \times X_n$. Nous avons la convergence

$$(a_k^{(1)}, \dots, a_k^{(n)}) \xrightarrow{X_1 \times \dots \times X_n} (a^{(1)}, \dots, a^{(n)}) \tag{7.14}$$

si et seulement si $a_k^{(i)} \to a^{(i)}$ pour chaque i.

4. Définition 7.12.

Démonstration. En deux parties.

(i) **Sens direct** Soient des ouverts $\mathcal{O}_i$ autour de $a^{(i)}$ dans X_i. Puisque $\mathcal{O}_1 \times \ldots \times \mathcal{O}_n$ est un ouvert autour de $(a^{(1)}, \ldots, a^{(n)})$, il existe $K \in \mathbb{N}$ tel que si $k \geqslant K$ nous avons $(a_k^{(1)}, \ldots, a_k^{(n)}) \in \mathcal{O}_1 \times \ldots \times \mathcal{O}_n$. Pour ce K nous avons séparément $a_k^{(i)} \in \mathcal{O}_i$ pour chaque i.

(ii) **Sens inverse** Une prébase de la topologie sur $X_1 \times \ldots \times X_n$ est donnée par les $\mathcal{O}_1 \times \ldots \times \mathcal{O}_n$ où $\mathcal{O}_i$ est un ouvert de X_i. Voir la définition 7.14 de la topologie produit et la définition 7.11 de ce qu'est une prébase.

La proposition 7.13 nous permet de ne vérifier la convergence de $(a_k^{(1)}, \ldots, a_k^{(n)})$ que sur la prébase. Soit donc $\mathcal{O} = \mathcal{O}_1 \times \ldots \times \mathcal{O}_n$ avec $(a^{(1)}, \ldots, a^{(n)}) \in \mathcal{O}$. Puisque $(a_k^{(i)})_{k\in\mathbb{N}} \to a^{(i)}$, pour chaque i, il existe $K_i \in \mathbb{N}$ tel que si $k \geqslant K_i$ alors $a_k^{(i)} \in \mathcal{O}_i$.

En posant $K = \max_i(K_i)$, nous avons $(a_k^{(1)}, \ldots, a_k^{(n)}) \in \mathcal{O}_1 \times \ldots \mathcal{O}_n$ pour tout $k \geqslant K$.

La proposition 7.13 permet de conclure.

□

7.2 Topologie rendant continue

Proposition 7.16 (Topologie qui rend continue[165]).
Soient un ensemble X, des espaces topologiques $(Y_i, \tau_i)_{i\in I}$, et des applications $\varphi_i \colon X \to Y_i$. Nous notons

$$\Lambda = \{(i, \omega_i) \text{ tel que } i \in I, \omega_i \in \tau_i\}. \tag{7.15}$$

Pour chaque Γ fini dans Λ, nous posons

$$\Phi_\Gamma = \bigcap_{(i,\omega_i)\in\Gamma} \varphi_i^{-1}(\omega_i). \tag{7.16}$$

Enfin nous posons

$$\tau_I = \bigcup_{\Gamma \text{ fini dans } \Lambda} \Phi_\Gamma. \tag{7.17}$$

Nous avons :

(1) τ_I est une topologie sur X.

(2) Toutes les applications φ_i sont continues pour cette topologie.

(3) La topologie τ_I est la plus faible topologie sur X pour laquelle toutes les φ_i sont continues.

La topologie ainsi définie est souvent référée comme la plus petite topologie qui rend les applications φ_i continues.

La topologie quotient d'un espace topologique par une relation d'équivalence est définie comme la plus petite topologie rendant continue la projection. Voir la définition 7.20.

Lemme 7.17 ([165]).
Soient des espaces topologiques $\{(Y_i, \tau_i)\}_{i\in I}$, et X, un ensemble. Nous considérons des applications $\varphi_i \colon X \to Y_i$ ainsi que τ, la topologie minimale sur X telle que les applications φ_i soient continues[5]

Soit

$$L = \{(i, V) \text{ tel que } i \in I, V \in \tau_i\}. \tag{7.18}$$

Alors l'ensemble

$$\Big\{ \bigcup_{(j,V)\in J} \varphi_j^{-1}(V) \Big\}_{J \text{ fini dans } I} \tag{7.19}$$

est une base de la topologie τ.

5. Proposition 7.16.

Lemme 7.18 ([165]).
Soient des espaces topologiques $\{(Y_i, \tau_i)\}_{i\in I}$, et X, un ensemble. Nous considérons des applications $\varphi_i\colon X \to Y_i$ ainsi que τ, la topologie minimale sur X telle que les applications φ_i soient continues.

Soit une suite (x_n) dans X. Nous avons $x_n \xrightarrow{\tau} x$ si et seulement si $\varphi_i(x_n) \xrightarrow{\tau_i} \varphi_i(x)$ pour tout $i \in I$.

Démonstration. Dans le sens direct, c'est seulement le fait que les φ_i sont continues.

Dans le sens réciproque, nous supposons que $\varphi_i(x_n) \xrightarrow{\tau_i} \varphi_i(x)$ pour tout i et nous devons prouver que $x_n \xrightarrow{\tau} x$.

Soit un voisinage U de x. Vue la base de topologie donnée dans le lemme 7.17, il existe un J fini dans I ainsi que des ouverts $\{V_j\}_{j\in J}$ tels que

$$x \in W = \bigcap_{j\in J} \varphi_j^{-1}(V_j) \subset U \tag{7.20}$$

Par hypothèse, $\varphi_i(x_n) \to \varphi_i(x)$. Mais V_j est un ouvert qui contient $\varphi_j(x)$. Donc il existe $N_j \in \mathbb{N}$ tel que $\varphi_j(x_n) \in V_j$ pour tout $n \geqslant N_j$. En posant[6] $N = \max\{N_j\}_{j\in J}$, nous avons que $\varphi_j(x_n) \in V_j$ pour tout $n \geqslant N$ et pour tout $j \in J$.

Dans ce cas nous avons aussi $x_n \in W \subset U$. La convergence est prouvée. □

Proposition 7.19 ([165]).
Soient des espaces topologiques (Y_i, τ_i), un ensemble X, et des applications $\varphi_i\colon X \to Y_i$. Nous considérons sur X la plus petite topologie rendant continues[7] *les φ_i.*

Soit un espace topologique Z. Une application $\psi\colon Z \to (X, \tau_I)$ est continue si et seulement si les applications $\varphi_i \circ \psi\colon Z \to Y_i$ sont continues pour tout $i \in I$.

Démonstration. Le sens direct est une simple composée de fonctions continues. Pour l'autre sens, nous supposons que les $\phi_i \circ \psi$ sont continues, nous considérons $U \in \tau_I$ et nous devons montrer que $\psi^{-1}(U)$ est un ouvert de Z.

En posant

$$\Lambda = \{(i, \omega) \text{ tel que } i \in I, \omega \in \tau_i\}, \tag{7.21}$$

il existe un Γ fini dans Λ tel que

$$U = \bigcap_{(i,\omega)\in\Gamma} \varphi_i^{-1}(\omega). \tag{7.22}$$

Nous avons donc

$$\psi^{-1}(U) = \bigcap_{(i,\omega_i)\in\Gamma} (\psi \circ \varphi_i^{-1})(\omega_i). \tag{7.23}$$

Vu que $\psi \circ \varphi^{-1}$ est continue, chacun des $(\psi \circ \varphi_i^{-1})(\omega)$ est ouvert. Donc $\psi^{-1}(U)$ est ouvert comme intersection finie d'ouverts. □

7.2.1 Topologie quotient

Définition 7.20 ([166]).
Soit un espace topologique X ainsi qu'une relation d'équivalence $\sim$ sur X. La ***topologie quotient*** *sur l'ensemble $X/\sim$ est la plus petite topologie qui rend continue*[8] *la projection canonique $p\colon X \to X/\sim$.*

Proposition 7.21.
Soit un espace topologique X muni d'une relation d'équivalence $\sim$. Une partie $\mathcal{O}$ est $X/\sim$ est ouverte[9] *si et seulement si $p^{-1}(\mathcal{O})$ est ouverte dans X.*

6. Notez l'utilisation du lemme 1.70 pour justifier que le maximum existe.
7. Proposition 7.16.
8. Définition 7.16.
9. La topologie est définie en 7.20.

Proposition 7.22 ([166]).
Soient des espaces topologiques X et Y ainsi qu'une relation d'équivalence $\sim$ sur X. Soit la projection canonique $p\colon X \to X/\sim$. Une application $f\colon X/\sim \to Y$ est continue si et seulement si l'application composée $f \circ p\colon X \to Y$ est continue.

Définition 7.23 (Passage d'une application aux classes).
Soient deux ensembles X et E ainsi qu'une relation d'équivalence $\sim$ sur X. Nous disons qu'une application $f\colon X \to E$ ***passe aux classes*** *si f est constante sur chaque classe d'équivalence de X. Dans ce cas nous considérons l'****application quotient***

$$\begin{aligned} \tilde{f}\colon X/\sim &\to E \\ [x] &\mapsto f(x). \end{aligned} \tag{7.24}$$

Lemme 7.24 ([167]).
Soient deux espaces topologiques X et Y. Soit une relation d'équivalence $\sim$ sur X. Nous considérons une application continue $f\colon X \to Y$ capable de descendre aux classes[10]*. Alors l'application quotient $\tilde{f}\colon X/\sim \to Y$ est continue.*

Démonstration. Nous considérons la projection canonique $p\colon X \to X/\sim$. Soit un ouvert $\mathcal{O}$ dans Y. La partie $\tilde{f}^{-1}(\mathcal{O})$ sera ouverte si $p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big)$ est ouverte (proposition 7.21). Nous avons

$$p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big) = (\tilde{f} \circ p)^{-1}(\mathcal{O}). \tag{7.25}$$

Mais $\tilde{f} \circ p = f$ parce que $(\tilde{f} \circ p)(x) = \tilde{f}([x]) = f(x)$. Donc

$$p^{-1}\big(\tilde{f}^{-1}(\mathcal{O})\big) = (\tilde{f} \circ p)^{-1}(\mathcal{O}) = f^{-1}(\mathcal{O}) \tag{7.26}$$

qui est ouvert parce que f est continue. □

Lemme 7.25 ([167]).
Soient deux espaces topologiques X et Y ainsi qu'une relation d'équivalence $\sim$ sur X. Nous considérons une application continue $g\colon X/\sim \to Y$. Alors

(1) L'application $g \circ p$ est continue.

(2) Si nous posons $f = g \circ p$, alors f descend aux classes et $\tilde{f} = g$.

Démonstration. La projection p est toujours continue par définition de la topologie quotient. Donc $g \circ p$ est continue par composition d'applications continues.

Voyons que f descend aux classes. Si $x \sim y$, alors $p(x) = p(y)$ et donc $f(x) = f(y)$. En ce qui concerne l'application $\tilde{f}$, nous avons

$$\tilde{f}\big([x]\big) = (g \circ p)(x) = g\big([x]\big), \tag{7.27}$$

donc $\tilde{f} = g$ et le lemme est démontré. □

7.2.2 Adhérence, fermeture, intérieur, point d'accumulation et point isolé

Définition 7.26.
Soient un espace topologique X et une partie A de X.

(1) Un point $x \in X$ est ***intérieur*** *à A si il existe un ouvert $\mathcal{O}$ tel que$x \in \mathcal{O} \subset A$.*

*(2) L'****intérieur*** *de A, notée $\mathrm{Int}(A)$, est l'union de tous les ouverts de X contenus dans A.*

Lemme 7.27.
Quelques propriétés en vrac.

(1) L'intérieur de A est l'ensemble de tous les points intérieurs de A.

10. Voir la définition 7.23.

(2) Pour tout $A \subset X$, l'ensemble $\mathrm{Int}(A)$ *est un ouvert.*

(3) On a $\mathrm{Int}(A) \subset A$

(4) Nous avons $\mathrm{Int}(A) = A$ *si et seulement si A est un ouvert.*

Démonstration. En plusieurs morceaux.

(i) **(1)** Si $a \in \mathrm{Int}(A)$, alors a est dans un ouvert contenu dans A, et donc a est un point intérieur à A. Dans l'autre sens, si a est un point intérieur à A, alors il existe un ouvert $\mathcal{O} \subset A$ contenant a. Puisque $\mathcal{O}$ est un ouvert dans A, nous avons $\mathcal{O} \subset \mathrm{Int}(A)$, et en particulier $a \in \mathrm{Int}(A)$.

(ii) **(2)** L'ensemble $\mathrm{Int}(A)$ est une union d'ouverts.

(iii) **(3)** L'ensemble $\mathrm{Int}(A)$ est une union d'ensembles contenus dans A.

(iv) **(4)** Supposons que $\mathrm{Int}(A) = A$. Puisque $\mathrm{Int}(A)$ est ouvert (point (3)), A est ouvert aussi. Dans l'autre sens, nous supposons que A est ouvert. Puisque A est un ouvert contenu dans A, nous avons $A \subset \mathrm{Int}(A)$. Mais comme $\mathrm{Int}(A) \subset A$, nous avons l'égalité.

□

7.2.2.1 Adhérence et fermeture

Disons-le tout de suite : « adhérence » et « fermeture » sont synonymes. Dans le Frido, nous allons utiliser les notations $\mathrm{Adh}(A)$ et $\bar{A}$ de façon opportuniste. La notation $\bar{z}$ définissant le complexe conjugué de z, si A est une partie de $\mathbb{C}$, il est plus sûr d'écrire $\mathrm{Adh}(A)$ pour la fermeture, plutôt que $\bar{A}$.

Au contraire, pour éviter une quantité excessive de parenthèses, nous écrirons $\overline{B(a,r)}$ pour la boule fermée.

Définition 7.28.
Soient un espace topologique X et une partie A de X. Un point $x \in X$ est ***adhérent*** *à A si tout ouvert de X contenant x a une intersection non vide avec A. L'ensemble des points d'adhérence de A est noté* $\mathrm{Adh}(A)$.

Lemme 7.29.
À propos d'adhérence.

(1) L'adhérence de A est l'intersection de tous les fermés de X contenant A.

(2) Nous avons l'égalité

$$\mathrm{Int}(A)^c = \mathrm{Adh}(A^c). \tag{7.28}$$

Démonstration. Commençons par prouver la dernière égalité d'ensembles. On a les équivalences entre les affirmations suivantes, pour tout $x \in X$:

— x n'est pas dans $\mathrm{Int}(A)$;

— il n'y a aucun ouvert contenant x et inclus dans A ;

— tout ouvert contenant x a une intersection non-vide avec A^c ;

— x est dans $\mathrm{Adh}(A^c)$.

Nous allons à présent montrer l'égalité d'ensembles $\mathrm{Int}(A)^c = \mathrm{Adh}(A^c)$ en prouvant la double inclusion.

(i) **Si** $\mathrm{Int}(A)^c \subset \mathrm{Adh}(A^c)$ Soit $x \in \mathrm{Int}(A)^c$. Nous devons prouver que $x \in \mathrm{Adh}(A^c)$. Soit un ouvert $\mathcal{O}$ contenant x. Vu que x n'est pas dans l'intérieur de A, l'ouvert $\mathcal{O}$ est pas inclus à A, et donc $\mathcal{O} \cap A^c$ est non vide.

Nous avons montré que tout ouvert contenant x intersecte A^c. Autrement dit : $x \in \mathrm{Adh}(A^c)$.

(ii) $\mathrm{Adh}(A^c) \subset \mathrm{Int}(A)^c$ Soit $x \in \mathrm{Adh}(A^c)$. Tout ouvert contenant x intersecte A^c, et ne peut donc pas être inclus à A. Si aucun ouvert contenant x n'est inclus à A, alors x n'est pas dans $\mathrm{Int}(A)$.

□

Remarque 7.30.
Comme corolaire du lemme 7.29, et grâce aux remarques faites pour les intérieurs, on obtient que pour $A \subset X$:

(1) l'ensemble $\bar{A}$ est fermé : c'est en effet le complémentaire d'un ouvert, précisément l'intérieur de A^c ;

(2) A est fermé si et seulement si $\bar{A} = A$: en effet, A est fermé si et seulement si A^c est ouvert, si et seulement si l'intérieur de A^c est A^c lui-même ; or, l'intérieur de A^c est le complémentaire de $\bar{A}$ par le lemme 7.29, si bien que A est fermé si et seulement si $(\bar{A})^c = A^c$, ou encore... ce qu'on affirmait au début.

Définition 7.31.
*Soit X un espace topologique. Un sous-ensemble A de X est **dense** dans X si $\bar{A} = X$.*

7.2.2.2 Frontière

Définition 7.32.
*Soit X un espace topologique, et $A \subset X$. La **frontière** de A, notée ∂A, est l'ensemble des points adhérents de A qui ne sont pas intérieurs à A. Ainsi,*

$$\partial A = \mathrm{Adh}(A) \backslash \mathrm{Int}(A). \tag{7.29}$$

7.2.2.3 Topologie induite

Proposition-Définition 7.33 (Topologie induite[168])**.**
Soit un espace topologique (X, τ_X), et soit $Y \subset X$. Nous définissons

$$\tau_Y = \{Y \cap \mathcal{O} \text{ tel que } \mathcal{O} \in \tau_X\}. \tag{7.30}$$

L'ensemble τ_Y est une topologie sur Y.
*Elle est la **topologie induite**.*

Démonstration. Il s'agit de vérifier les conditions de la définition 7.1.

(i) $\underline{Y \in \tau_Y}$ Parce que $Y = X \cap Y$ et que X est un ouvert de X.

(ii) $\underline{\varnothing \in \tau_Y}$ Parce que $\varnothing = Y \cap \varnothing$ et que $\varnothing$ est un ouvert de X.

(iii) **Union quelconque** Soient des ouverts A_i de X. Nous avons

$$\bigcup_{i \in I} Y \cap A_i = Y \cap \big(\bigcup_{i \in I} A_i\big). \tag{7.31}$$

Comme les A_i sont des ouverts de X, leur union est encore un ouvert de X. Donc (7.31) est encore dans τ_Y.

(iv) **Intersection finie** Nous avons

$$\bigcap_{i \in I} Y \cap A_i = Y \cap \left(\bigcap_{i \in I} A_i\right). \tag{7.32}$$

□

Lemme 7.34 ([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 $F \subset S$ et 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 F dans X nous avons

$$F^c = (S\backslash F) \cup (X\backslash S) = (S \cap \Omega) \cup S^c = (S \cap \Omega) \cup (S^c \cap \Omega) \cup S^c = \Omega \cup S^c. \tag{7.33}$$

Puisque 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. □

Lemme 7.35.
Si $B \subset A$ alors la fermeture de B pour la topologie de A (induite de X) que nous noterons $\tilde{B}$ est

$$\tilde{B} = \bar{B} \cap A \tag{7.34}$$

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. Comme $a \in \bar{B}$, l'ensemble $\mathcal{O}$ intersecte B et donc $(\mathcal{O} \cap A) \cap B \neq \emptyset$. Donc a est bien dans l'adhérence [11] de B au sens de la topologie de A.

Pour l'inclusion inverse, soit $a \in \tilde{B}$, 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 $a \in \bar{B}$. □

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 si il est ouvert dans X, mais certaines choses surprenantes se produisent tout de même.

Lemme 7.36.
La partie $\mathbb{Q}$ dans $\mathbb{R}$ est d'intérieur vide, et sa fermeture est $\mathbb{R}$.

Exemple 7.37.
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 7.38.
Quid de la boule ouverte $B(1, \epsilon)$ dans le fermé $[0, 1]$? Par définition c'est

$$B(1, \epsilon) = \{x \in [0, 1] \text{ tel que } |x - 1| < \epsilon\} =]1 - \epsilon, 1]. \tag{7.35}$$

Oui, c'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 dans $[0, 1]$. Bref, vous pouvez écrire

$$B(\frac{1}{2}, 10) \subset [0, 1], \tag{7.36}$$

mais vous avez intérêt à être très clair sur la topologie sous-entendue. △

11. Définition 7.28.

7.2.2.4 Points d'accumulation et isolés

Définition 7.39.
*Soient un espace topologique X et une partie A de X. Un point $s \in X$ est un **point d'accumulation** de A si tout ouvert de X contenant s contient au moins un élément de $A \backslash \{s\}$.*

Quelle est la différence entre un point d'accumulation et un point d'adhérence ? La différence est que tous les points de A sont des points d'adhérence de A, parce que tout voisinage de $a \in A$ contient au moins a lui-même, alors que certains points de A peuvent ne pas être des points d'accumulation de A. Voir l'exemple 7.202.

Notons qu'un point d'accumulation de A dans X n'est pas spécialement dans A.

Définition 7.40.
*Soient un espace topologique X et une partie A de X. Un point $s \in A$ est un **point isolé** de A si il existe un voisinage ouvert $\mathcal{O}$ de s dans X tel que $A \cap \mathcal{O} = \{s\}$.*

La définition suivante est *la* définition de la continuité dans tous les cas.

Définition 7.41 (Fonction continue[169])**.**
Deux définitions :

*(1) Soient une fonction $f\colon X \to Y$ entre les espaces topologiques X et Y et un point $a \in X$. Nous disons que f est **continue** en a si pour tout ouvert W contenant $f(a)$, il existe un voisinage V de a dans X tel que $f(V) \subset W$.*

*(2) Une fonction $f\colon X \to Y$ est **continue** sur X si pour tout ouvert $\mathcal{O}$ de Y, l'ensemble*

$$f^{-1}(\mathcal{O}) = \{x \in X \text{ tel que } f(x) \in \mathcal{O}\} \tag{7.37}$$

est ouvert dans X.

Lemme 7.42.
Soient deux espaces topologiques X et Y et une application $f\colon X \to Y$. Soit une base de topologie $\{A_i\}_{i\in I}$ de Y. Si $f^{-1}(A_i)$ est ouvert dans X pour tout $i \in I$ alors f est continue.

Exemple 7.43 ([1])**.**
Un truc bien avec la définition 7.41(1) est que la continuité de f en un point est définie pour tout point du domaine ; pas seulement les points d'accumulation. Soit par exemple une fonction simple

$$\begin{aligned} f\colon \{a\} &\to \mathbb{R} \\ a &\mapsto 4. \end{aligned} \tag{7.38}$$

Si W est un ouvert de $\mathbb{R}$ contenant 4, nous avons l'ouvert $V = \{a\}$ tel que $f(V) \subset W$. Donc f est continue au point 4.

Mais f est également continue sur $\{4\}$ en tant qu'espace topologique. En effet, si W est un ouvert de $\mathbb{R}$, l'ensemble $f^{-1}(W)$ est soit $\varnothing$ soit $\{a\}$. Dans les deux cas c'est un ouvert. △

7.2.2.5 Isomorphismes

Définition 7.44 (Isomorphisme d'espaces topologiques)**.**
*Un **isomorphisme** d'espaces topologiques est une application bijective continue*[12] *entre deux espaces topologiques dont la réciproque est continue. Un isomorphisme d'espaces vectoriels est également dit **homéomorphisme**.*

*Un isomorphisme d'un espace avec lui-même est un **automorphisme**.*

Lemme 7.45 ([1])**.**
Si $f\colon X \to Y$ est un homéomorphisme[13] *et si F est fermé dans X, alors $f(F)$ est fermé dans Y.*

12. Application continue, définition 7.41.
13. Définition 7.44.

Démonstration. Le complémentaire F^c est ouverte. Vu que f^{-1} est continue, la partie $f(F^c)$ est ouverte. Comme f est une bijection, nous avons $f(F^c) = f(F)^c$. D'où le fait que $f(F)^c$ est ouvert, et donc que $f(F)$ est fermé. □

7.3 Suites et convergence

7.46.
À propos de notations. La pire notation possible pour une suite est $(a_n)_n$. Mais que vient faire le second indice n ? Il peut être raisonnable d'écrire $(a_n)_{n\in I}$ lorsqu'on veut dire dans quel ensemble se déplace n. Si nous parlons de *suite*, il faut une sérieuse raison de prendre autre chose que $\mathbb{N}$ comme ensemble d'indices.

Une suite étant une fonction, de la même façon qu'on ne devrait pas dire « la fonction $f(x)$ », mais « la fonction f » ou « la fonction $x \mapsto f(x)$ », nous devrions simplement écrire a pour désigner la suite dont les éléments sont a_n.

Par conséquent, il est parfaitement légal, et même conseillé, d'écrire « $a + b$ » pour la somme des suites a et b. Et il est tout aussi légal d'écrire « $\lim a$ » au lieu de $\lim_{n\to\infty} a_n$.

Le hic est que nous écrivons souvent x la limite de la suite $n \mapsto x_n$. Dans ce cas, nous sommes évidemment obligé d'écrire l'indice n pour parler de la suite.

Tout cela pour dire qu'il faut être souple avec les notations.

7.3.1 Convergence dans un fermé

Proposition 7.47 ([1])**.**
Une suite contenue dans un fermé ne peut converger que vers un élément de ce fermé.

Démonstration. Soient un espace topologique X et un fermé F dans X. Nous supposons que la suite (x_k) soit contenue dans F. Nous allons prouver qu'aucun élément de F^c ne peut être limite.

Soit $a \in F^c$. Puisque le complémentaire de F est un ouvert, et d'après le théorème 7.7, il existe un ouvert $\mathcal{O}_a$ contenant a, et contenu dans F^c. Le voisinage $\mathcal{O}_a$ de a ne contient donc aucun élément de la suite (x_k), qui ne peut donc pas converger vers a. □

Corolaire 7.48.
Soit A un sous-ensemble d'un espace topologique X. Toute suite d'éléments de A qui converge, admet pour limite un élément de $\mathrm{Adh}(A)$.

Démonstration. Une fois la suite (x_n) fixée, il suffit de remarquer que tous les x_n sont dans $\mathrm{Adh}(A)$, et puis d'appliquer la proposition 7.47. □

Lemme 7.49.
Soit $A \subset X$ muni de la topologie induite de X et (x_n) une suite dans A. Si (x_n) converge vers un élément x dans A, alors elle converge aussi vers x dans 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 $x_n \in A \cap \mathcal{O} \subset \mathcal{O}$. □

7.3.2 Pour des limites uniques : séparabilité

Notons que l'on a parlé d'*une* limite de suite jusqu'à présent : en effet, si il existe deux éléments distincts x et y tels que tout ouvert contenant x contient y, alors la définition 7.12 dit que toute suite convergeant vers y converge aussi vers x...

Exemple 7.50.
Oui, il y a moyen de converger vers plusieurs points distincts si l'espace n'est pas super cool. Nous pouvons par exemple [170] 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 $0'$.

En fait, on « voit » le problème : on ne peut pas distinguer d'un point de vue topologique le 0 et le $0'$. △

Nous posons la définition suivante, qui nous permettra de donner une assez grande classe d'espaces topologiques dans lesquels nous avons unicité de la limite [14].

Définition 7.51 (Espace topologique Hausdorff)**.**
Si deux points distincts admettent toujours deux voisinages disjoints [15]*, nous disons que l'espace est* ***séparé*** *ou* ***de Hausdorff****.*

Attention, cette notion est à ne pas confondre avec :

Définition 7.52 (Espace topologique séparable)**.**
Un espace topologique est ***séparable*** *si il possède une partie dénombrable* [16] *dense* [17]*.*

Proposition 7.53.
Dans un espace topologique séparé, si une suite converge, alors sa limite est unique.

Démonstration. Supposons que la suite (x_k) converge vers deux éléments distincts x et y. L'espace étant séparé, il existe deux ouverts $\mathcal{O}_x$ et $\mathcal{O}_y$, disjoints, contenant respectivement x et y. La suite convergeant à la fois vers x et y, il existe k_x et k_y, tels que, si $k \geqslant \max\{k_x, k_y\}$, l'élément x_k est (à la fois) dans $\mathcal{O}_x$ et $\mathcal{O}_y$. Cela est en contradiction avec le fait que ces deux ensembles sont disjoints. □

7.54.
Donc, on pourra parler, avec des espaces séparés, de « la limite d'une suite ». On notera $x_n \to a$, ou $\lim_{n\to\infty} x_n = a$, pour signifier que la suite (x_n) converge vers a.

Lemme 7.55 ([1])**.**
Soit $a \neq 0$ dans un espace vectoriel topologique [18] *Hausdorff* [19]*. Il existe un voisinage V de 0 tel que $a \notin \bar{V}$.*

Démonstration. Étant donné que l'espace topologique est Hausdorff, nous pouvons considérer des voisinages V de 0 et W de a tels que $V \cap W = \varnothing$.

Dans ce cas nous avons $a \notin \bar{V}$ (voir la définition 7.28 de la fermeture de V). □

Proposition 7.56 ([1])**.**
La convergence de suite pour la topologie de l'espace produit [20] *est équivalente à la convergence des suites « composante par composante ».*

Démonstration. En deux parties

(i) **<u>Sens direct</u>** 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 $k \geqslant K$ 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 $k \geqslant K$ implique $x_k \in \mathcal{O}_1$. Donc $x_k \xrightarrow{X} x$. Idem pour y.

14. Voir la proposition 7.95.
15. Définition 1.3.
16. Définition 1.123.
17. Définition 7.31.
18. Définition 7.148.
19. Définition 7.51
20. Définition 7.14.

(ii) **Dans l'autre sens** Nous considérons l'espace produit[21] $X = \prod_{i=1}^{n} X_i$. Nous supposons pour chaque i, avoir une suite convergente $(x_i)_k \xrightarrow{X_i} x_i$.
Nous allons prouver que

$$\big((x_1)_k, \dots, (x_n)_k\big) \xrightarrow{X} (x_1, \dots, x_n). \tag{7.39}$$

Soit un ouvert $\mathcal{O}$ de X autour de $(x_1, \dots, x_n)$. Nous considérons des ouverts U_i de X_i tels que $x_i \in U_i$ et $U_1 \times \dots U_n \subset \mathcal{O}$.
Vu que $(x_i)_k \xrightarrow{X_i} x_i$, il existe $K_i \in \mathbb{N}$ tel que $k > K_i$ implique $(x_i)_k \in U_i$. Si $k \geqslant \max_i\{K_i\}$, alors $(x_i)_k \in U_i$ pour tout i et nous avons

$$\big((x_1)_k, \dots, (x_n)_k\big) \in U_1 \times \dots \times U_n \subset \mathcal{O}. \tag{7.40}$$

□

Lemme 7.57 ([1]).
Soit un espace topologique X. Soient dans X une suite (x_n) et un élément x tels que toute sous-suite de (x_n) contient une sous-suite convergente vers x. Alors $x_n \to x$.

Démonstration. Supposons que (x_n) ne converge pas vers x. Il existe alors un ouvert $\mathcal{O}$ autour de x tel que pour tout $N > 0$, il existe $n \geqslant N$ tel que x_n n'est pas dans $\mathcal{O}$.

Cela nous permet de construire une sous-suite de (x_n) composée d'éléments hors de $\mathcal{O}$. Aucune sous-suite de cette sous-suite ne peut converger vers x. □

7.3.3 Fonctions équivalentes

Proposition-Définition 7.58 ([171]).
Soit un espace topologique X et $D \subset X$. Soient encore des fonctions $f, g\colon D \to \mathbb{C}$ et un point $a \in \mathrm{Adh}(D)$[22].

Nous définissons sur $\mathrm{Fun}(D, \mathbb{C})$ la relation $f \sim g$ lorsque qu'il existe un voisinage V de a dans X et une fonction $\alpha\colon V \to \mathbb{R}$ telles que

(1) $\lim_{x\to a} \alpha(x) = 0$,

(2) pour tout $x \in (V \cap D)\backslash\{a\}$,

$$f(x) = \big(1 + \alpha(x)\big)g(x). \tag{7.41}$$

Cette relation est une relation d'équivalence.

*Lorsque $f \sim g$, nous disons que f et g sont **équivalentes** en a.*

Démonstration. Nous devons prouver les trois conditions de la définition 1.29 de relation d'équivalence.

(i) **Réflexive** Il suffit de poser $\alpha(x) = 0$.

(ii) **Symétrique** Si $f \sim g$, il existe une fonction α vérifiant ce qu'il faut telle que

$$f(x) = \big(1 + \alpha(x)\big)g(x). \tag{7.42}$$

Comme $\lim_{x\to a} \alpha(x) = 0$, il y a un voisinage de a sur lequel $|\alpha(x)| < 1$; il n'y a donc pas de problème de dénominateur en écrivant

$$g(x) = \frac{1}{1 + \alpha(x)} f(x). \tag{7.43}$$

Nous posons alors $\beta(x) = -\alpha(x)/\big(1 + \alpha(x)\big)$. Cela vérifie

$$g(x) = \big(1 + \beta(x)\big)f(x). \tag{7.44}$$

21. Pour les notations, ça va être le sport : $(x_i)_k$ désigne une suite dans X_i, mais x_i désigne la limite de cette suite.
22. Adhérence ou fermeture, c'est la même chose. Voir la définition 7.28 et le lemme 7.29.

Et

$$\lim_{x\to a} \beta(x) = 0 \tag{7.45}$$

parce que $\lim_{x\to a} \big(1 + \alpha(x)\big) = 1$ et $\lim_{x\to a} -\alpha(x) = 0$.

(iii) **Transitive** Soit $f \sim g$ et $g \sim h$. Sur un voisinage V de a nous avons

$$f(x) = \big(1 + \alpha(x)\big)g(x), \tag{7.46}$$

sur un voisinage W de a nous avons

$$g(x) = \big(1 + \beta(x)\big)h(x). \tag{7.47}$$

Sur le voisinage $V \cap W$ nous avons

$$f(x) = \big(1 + \beta(x) + \alpha(x) + (\alpha\beta)(x)\big)h(x). \tag{7.48}$$

Donc la fonction $\gamma(x) = \beta(x) + \alpha(x) + (\alpha\beta)(x)$ fait l'affaire.

□

Notons que la notion d'équivalence de fonctions, de même que la notion de limite, ne dépend pas des valeurs exactes atteintes par les fonctions au point.

Lemme 7.59.
Si f et g sont équivalentes en a, et si g ne s'annule pas sur un voisinage de a, alors pour tout $\epsilon > 0$, il existe r tel que

$$\frac{f(x)}{g(x)} \in B(1, \epsilon) \tag{7.49}$$

pour tout $x \in B(a, r)$.

Démonstration. Nous considérons un voisinage V de a sur lequel en même temps :

— la fonction α de la définition d'équivalence est définie,

— $|\alpha(x)| < \epsilon$ pour tout $x \in V$,

— $g(x) \neq 0$, pour tout $x \in V$.

Ensuite nous considérons $r > 0$ tel que $B(a, r) \subset V$. En divisant la condition (7.41) par $g(x)$ nous trouvons

$$\frac{f(x)}{g(x)} = 1 + \alpha(x). \tag{7.50}$$

Donc

$$|\frac{f(x)}{g(x)} - 1| = |\alpha(x)| \leqslant \epsilon, \tag{7.51}$$

ce qu'il fallait prouver. □

7.4 Connexité

L'idée de la connexité, c'est de s'assurer qu'un ensemble est « d'un seul tenant ».

Définition 7.60.
*Lorsque X 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*

$$A = (A \cap O_1) \cup (A \cap O_2), \tag{7.52}$$

*et tels que $A \cap O_1 \neq \emptyset$, et $A \cap O_2 \neq \emptyset$. Si un sous-ensemble n'est pas non-connexe, alors on dit qu'il est **connexe**.*

Une autre façon d'exprimer la condition (7.52) 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-Définition 7.61 ([172]).
Soient un espace topologique X et un point $x \in X$.

(1) La réunion de toutes les parties connexes de X contenant x est connexe.

(2) Cette réunion est la plus grande (au sens de la relation d'inclusion) de toutes les parties connexes de X contenant x.

La réunion de toutes les parties connexes de X contenant x est nommée ***composante connexe*** *de x dans X.*

Proposition 7.62 ([1]).
Soit X un espace topologique. Les conditions suivantes sont équivalentes.

(1) L'espace X est connexe.

(2) Si $X = O_1 \cup O_2$ avec O_1 et O_2 des ouverts disjoints, alors soit $O_1 = \varnothing$ soit $O_2 = \varnothing$.

(3) Si $X = F_1 \cup F_2$ avec F_1 et F_2 fermés disjoints dans X, alors $F_1 = \varnothing$ ou $F_2 = \varnothing$.

(4) Si $A \subset X$ avec A ouvert et fermé en même temps, alors $A = \varnothing$ ou $A = X$.

Démonstration. En quatre parties.

(i) **(1) implique (2)** Par rapport à la définition 7.60, nous prenons la partie X de l'espace X. Supposons que O_1 et O_2 sont tout deux non vides. Dans ce cas nous avons

$$X = O_1 \cup O_2 = (O_1 \cap X) \cup (O_2 \cap X), \tag{7.53}$$

ce qui prouverait que X est non connexe. Contradiction. Un des O_i est vide.

(ii) **(2) implique (3)** Soit une union disjointe de fermés $X = F_1 \cup F_2$. Puisque l'union est disjointe, nous avons $F_1 = X \backslash F_2$ et $F_2 = X \backslash F_1$, ce qui fait que F_1 et F_2 sont également ouverts. Nous en déduisons que $X = F_1 \cup F_2$ est une union disjointe d'ouverts. L'hypothèse indique que $F_1 = \varnothing$ ou $F_2 = \varnothing$.

(iii) **(3) implique (4)** Soit A une partie ouverte et fermée de X. Nous supposons que A est ouvert et fermé, donc $X \backslash A$ est également ouvert et fermé : c'est la définition 7.3 d'un fermé. Nous avons évidemment l'union $X = A \cup (X \backslash A)$ qui est une union disjointe de fermés. Par hypothèse nous avons soit $A = \varnothing$ soit $X \backslash A = \varnothing$.

(iv) **(4) implique (1)** Supposons que X ne soit pas connexe. Il existe donc des ouverts disjoints O_1 et O_2 tels que $X = O_1 \cup O_2$. Étant donné que $O_1 = X \backslash O_2$, la partie O_1 est fermée comme complémentaire d'ouvert. Donc O_1 est fermé et ouvert (et O_2 aussi d'ailleurs). Par hypothèse nous concluons que O_1 est soit X soit $\varnothing$.

□

Nous verrons plus tard (proposition 7.183) une autre caractérisation de la connexité basée sur la continuité des fonctions $X \to \mathbb{Z}$.

Proposition 7.63.
Si $A \subset X$ est connexe et si $A \subset B \subset \mathrm{Adh}(A)$, alors B est connexe.

Proposition 7.64.
Stabilité de la connexité par union.

(1) Une union quelconque de connexes ayant une intersection non vide est connexe.

(2) Pour tout $n \in \mathbb{N}, n > 0$, si $A_1, \dots, A_n$ sont des connexes de X avec $A_i \cap A_{i+1} \neq \varnothing$, alors l'union $\bigcup_{i=1}^n A_i$ est connexe.

Démonstration. Point par point.

(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

(1a) $C_j \subset A \cup B$ parce que $A \cup B$ recouvre tous les C_i,

(1b) $C_j \cap A \neq \emptyset$ parce que p est dans l'intersection,

(1c) $C_j \cap B \neq \emptyset$ 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 [23]. 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 $A_1 \cup A_2$ et A_3 ont un point d'intersection par hypothèse, et ainsi de suite.

□

7.5 Compacité

La compacité est le thème 31.

7.5.1 Définition et notions connexes

Soit E, un sous-ensemble de $\mathbb{R}$. Nous pouvons considérer les ouverts suivants :

$$\mathcal{O}_x = B(x, 1) \tag{7.54}$$

pour chaque $x \in E$. Évidemment,

$$E \subseteq \bigcup_{x\in E} \mathcal{O}_x. \tag{7.55}$$

Cette union contient en général de nombreuses redondances. Si par exemple $E = [-10, 10]$, l'élément $3 \in E$ est contenu dans $\mathcal{O}_{3.5}$, $\mathcal{O}_{2.7}$ et bien d'autres. Pire : même si on enlève par exemple $\mathcal{O}_2$ de la liste des ouverts, l'union de ce qui reste continue à être tout E. La question est : *est-ce qu'on peut en enlever suffisamment pour qu'il n'en reste qu'un nombre fini ?*

Définition 7.65.
Soit E, un sous-ensemble de $\mathbb{R}$. Une collection d'ouverts $\mathcal{O}_i$ est un ***recouvrement*** *de E si $E \subseteq \bigcup_i \mathcal{O}_i$.*

Définition 7.66.
Une partie A d'un espace topologique est ***compacte*** *si elle 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 extraire un recouvrement fini.*

Remarque 7.67.
Certaines sources (dont wikipédia) disent que pour être compact il faut aussi être séparé [24]. Pour ces sources, un espace qui ne vérifie que la propriété de Borel-Lebesgue est alors dit **quasi-compact**.

7.68.
La définition 7.66 en cache deux. En effet, si la partie A est l'espace topologique lui-même, cela définit un espace topologique compact. Un espace topologique est compact *en soi* lorsque de tout recouvrement par des ouverts, nous pouvons extraire un sous-recouvrement fini. Dans ce cas, si X est l'espace et si $\{A_i\}_{i\in I}$ est le recouvrement, nous avons $X = \bigcup_{i\in I} A_i$ et non une simple inclusion $X \subset \bigcup_{i\in I} A_i$.

23. Parce qu'on a la flemme de faire une preuve par récurrence !
24. Définition 7.51.

Lemme 7.69.
Si $a, b \in \mathbb{R}$, alors la partie $[a, b]$ est compacte[25] *dans $\mathbb{R}$.*

Lemme 7.70.
Si K est une partie compacte de l'espace topologique X, alors K est un espace topologique compact pour la topologie induite[26] *de X.*

Démonstration. Nous notons τ la topologie de X et τ_K la topologie induite de X vers K, c'est-à-dire

$$\tau_K = \{\mathcal{O} \cap K \text{ tel que } \mathcal{O} \in \tau\}. \tag{7.56}$$

Soient des ouverts $A_i \in \tau_K$ ($i \in I$ où I est un ensemble quelconque) tels que $\bigcup_i A_i = K$. Pour chaque $i \in I$, il existe un $\mathcal{O}_i \in \tau$ tel que $A_i = K \cap \mathcal{O}_i$. Nous avons

$$K = \bigcup_{i \in I} (K \cap \mathcal{O}_i) \subset \bigcup_{i \in I} \mathcal{O}_i. \tag{7.57}$$

Donc les $\mathcal{O}_i$ forment un recouvrement de K par des ouverts de X. Puisque K est une partie compacte de X, il existe un sous-ensemble fini J de I tel que

$$K \subset \bigcup_{j \in J} \mathcal{O}_j. \tag{7.58}$$

Nous avons donc aussi

$$K \subset \bigcup_{j \in J} K \cap \mathcal{O}_j = \bigcup_{j \in J} A_j. \tag{7.59}$$

Nous avons prouvé que $\{A_j\}_{j \in J}$ est un recouvrement fini de K par des ouverts de K. Donc K est un espace topologique compact. □

7.5.2 Espace localement compact

Définition 7.71.
*Une partie d'un espace topologique est **relativement compact** si sa fermeture est compacte.*

Définition 7.72.
*Un espace topologique est **localement compact** si tout élément possède un voisinage compact.*

Lemme 7.73.
Si X est un espace topologique localement compact et si K est compact dans X, il existe un ouvert V tel que $K \subset V$ et $\bar{V}$ est compact.

Lemme 7.74.
Soient un espace localement compact X, un compact K et un ouvert $\mathcal{O}$ tel que $K \subset \mathcal{O}$. Il existe un ouvert relativement compact V tel que

$$K \subset V \subset \bar{V} \subset \mathcal{O}. \tag{7.60}$$

7.5.3 Autres compacité

Définition 7.75 (Séquentiellement compact)**.**
*Nous disons qu'un espace topologique est **séquentiellement compact** si toute suite admet une sous-suite convergente.*

Définition 7.76.
*Un espace topologique est **dénombrable à l'infini** si il est réunion dénombrable de compacts.*

25. Définition 7.66.
26. Définition 7.33.

Définition 7.77.
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 7.78.
Soient X un espace topologique et $K \subset X$. Les propriétés suivantes sont équivalentes :

(1) K est compact.

(2) Si $\{F_i\}_{i\in I}$ est une famille de fermés telle que $\bigcap_{i\in I} F_i \cap K = \varnothing$, alors il existe une partie finie non vide A de I tel que $\bigcap_{i\in A} F_i \cap K = \varnothing$.

(3) Si $\{F_i\}_{i\in I}$ est une famille de fermés telle que pour tout choix de A fini dans I, $\bigcap_{i\in A} F_i \cap K \neq \varnothing$, alors l'intersection complète est non vide : $\bigcap_{i\in I} F_i \cap K \neq \varnothing$.

(4) Toute famille de fermés de X, à laquelle K est joint, et qui a 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[27] reformulation 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 :

$$K \subset \bigcup_{i\in A} \mathcal{O}_i \tag{7.61}$$

pour un certain sous-ensemble fini A de I. Pour ce même choix A, nous avons alors aussi

$$\bigcap_{i\in A} F_i \cap K = \varnothing. \tag{7.62}$$

L'implication (2) $\Rightarrow$ (1) est la même histoire de passage aux complémentaires. □

Le théorème 7.259 est en général celui qu'on nomme « théorème des fermés emboîtés », mais le corolaire suivant en mériterait également le nom.

Corolaire 7.79 ([1])**.**
Soient un espace topologique compact X et une suite $(F_i)_{i\in\mathbb{N}}$ de fermés emboîtés[28] *dans X telle que*

$$\bigcap_{i\in\mathbb{N}} F_i = \varnothing. \tag{7.63}$$

Alors il existe $j_0 \in \mathbb{N}$ tel que $F_i = \varnothing$ pour tout $i \geqslant j_0$.

Démonstration. La proposition 7.78 nous dit qu'il existe une partie finie non vide J de $\mathbb{N}$ telle que $\bigcup_{j\in J} F_j = \varnothing$. Si $j_0 = \min(J)$, alors $F_j \subset F_{j_0}$ pour tout $j \in J$ et nous avons

$$\varnothing = \bigcap_{j\in J} F_j = F_{j_0}. \tag{7.64}$$

Dès que $F_{j_0} = \varnothing$, tous les suivants sont également vides. □

7.5.4 Quelques propriétés

Lemme 7.80.
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.

Remarquons que la partie qui est réellement à prouver est que, si « ça marche » pour des ouverts d'une base de topologie, alors « ça marche » pour tous types d'ouverts.

27. Enfin, simple... il faut remarquer que dans la formulation de (4), les intersections peuvent ne pas faire intervenir K, mais, au final, on s'en moque.
28. C'est-à-dire que $F_{i+1} \subset F_i$.

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 7.2 : disons $\mathcal{O}_i = \bigcup_{j\in J_i} A_{(i,j)}$. Soit $J = \{j = (i, j_i) | i \in I, j_i \in J_i\}$; alors nous obtenons $\bigcup_{j\in J} A_j = \bigcup_{i\in I} \mathcal{O}_i$.

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 dans (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 $K \subset \bigcup_{j\in J_0} A_j \subset \bigcup_{i\in I_0} \mathcal{O}_i$. □

Exemple 7.81 (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. △

Lemme 7.82 (Compacts et fermés[173])**.**
À propos de parties fermées dans un compact.

(1) Une partie fermée d'un compact est compacte.

(2) Tout compact d'un espace topologique séparé est fermé.

Démonstration. En deux parties.

(i) **Pour (1)** Soient F fermé dans un compact K et $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de F par des ouverts. Puisque 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 $\{\mathcal{O}_i\}_{i\in I}$.

(ii) **Pour (2)** Soient X un espace séparé et K compact dans X. Nous considérons $y \in K^c$ et, par hypothèse de séparation, pour chaque $x \in K$ nous considérons un voisinage ouvert V_x de x et un voisinage ouvert[29] W_x de y tels que $V_x \cap W_x = \emptyset$. Bien entendu les V_x forment un recouvrement de K par des ouverts dont nous pouvons extraire un sous-recouvrement fini : soit S fini dans K tel que

$$K \subset \bigcup_{x\in S} V_x. \tag{7.65}$$

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.

Montrons que $W \cap K = \emptyset$. Soit $a \in K$; par définition de S, il existe $s \in S$ tel que $a \in V_s$. Par conséquent, a n'est pas dans W_s et donc pas non plus dans W.

L'ouvert W prouve que y est dans l'intérieur du complémentaire de K, et comme y est arbitraire, nous concluons que le complémentaire de K est ouvert (théorème 7.7), en d'autres termes, que K est fermé.

□

Lemme 7.83 ([1])**.**
Toute union finie de compacts est compacte.

Démonstration. Soient $(K_i)_{i=1,\dots,n}$ des compacts dans X. Si $\{\mathcal{O}_s\}_{i\in S}$ est un recouvrement de $\bigcup_{i\in I} K_i$ par des ouverts, à fortiori, ce sera un recouvrement de chacun des K_i. Pour chaque i, il existera donc une partie finie S_i de S telle que $\{\mathcal{O}_s\}_{s\in S_i}$ recouvre K_i.

L'union finie de parties finies S_i est une partie finie de S, et nous avons

$$\bigcup_i K_i \subset \bigcup_{s\in\bigcup_i S_i} \mathcal{O}_s. \tag{7.66}$$

□

29. Oui, la notation du voisinage peut surprendre, mais elle est quand même pratique pour ce qu'on veut en faire.

Proposition 7.84 ([174]).
Dans un espace séparé, toute intersection de compacts est compacte.

Démonstration. Soit un espace topologie séparé X et des compacts $\{K_i\}_{i\in I}$ dans X (I est un ensemble quelconque). Chacun des K_i est fermé par le lemme 7.82(2). Donc l'intersection

$$K = \bigcap_{i\in I} K_i \tag{7.67}$$

est un fermé de X par le lemme 7.6(1). Soit i dans I. Nous avons $K \subset K_i$. Donc K est un fermé dans le compact K_i ; il est donc compact par le lemme 7.82. □

Exemple 7.85 (Intersection de compacts non compacte[174]).
Un exemple d'intersection de compacts qui n'est pas compacte. Vu la proposition 7.84, il va falloir chercher un espace non séparé. Soit $X = \mathbb{N} \cup \{x_1, x_2\}$ où x_1 et x_2 sont deux éléments distincts hors de $\mathbb{N}$. Nous définissons une topologie sur X en disant que les ouverts sont les parties suivantes :

— les parties de $\mathbb{N}$,
— la partie $\mathbb{N} \cup \{x_1\}$,
— la partie $\mathbb{N} \cup \{x_2\}$,
— la partie $\mathbb{N} \cup \{x_1, x_2\}$.

Nous considérons les parties $K_1 = \mathbb{N} \cup \{x_1\}$ et $K_2 = \mathbb{N} \cup \{x_2\}$.

(i) **K_i est compact** Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de K_1 par des ouverts de X. Alors il existe $i_0 \in I$ tel que $x_1 \in \mathcal{O}_{i_0}$. Vue la liste des ouverts, $\mathcal{O}_{i_0}$ est soit $\mathbb{N} \cup \{x_1\}$ soit $\mathbb{N} \cup \{x_1, x_2\}$. Dans les deux cas, $\{\mathcal{O}_{i_0}\}$ est un sous-recouvrement fini de K_1.

(ii) **$K_1 \cap K_2 = \mathbb{N}$** C'est immédiat parce que x_1 et x_2 sont distincts.

(iii) **$\mathbb{N}$ n'est pas compact** Il peut être recouvert par les ouverts $\{\{i\}\}_{i\in\mathbb{N}}$ dont on ne peut pas extraire de sous-recouvrements finis.

△

Proposition 7.86.
Si V est une partie de l'espace topologique X muni de la topologie induite[30] *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. □

Proposition 7.87 ([1]).
Soient des espaces topologiques X et Y. Nous considérons des ouverts A de X et B de Y. Soit un compact M dans $A \times B$[31]*. Il existe des compacts K et L dans A et B tels que $M \subset K \times L$.*

Démonstration. Nous considérons les « projections » de M sur A et B :

$$K = \{a \in A \text{ tel que } \exists b \in B \text{ tel que } (a, b) \in M\}, \tag{7.68}$$

et

$$L = \{b \in B \text{ tel que } \exists a \in A \text{ tel que } (a, b) \in M\}. \tag{7.69}$$

Nous avons $M \subset K \times L$; il reste à montrer que K et L sont des compacts de leurs espaces respectifs. Soit un recouvrement $\{U_i\}_{i\in I}$ de K par des ouverts de X et $\{V_j\}_{j\in J}$ de L par des ouverts de Y. Alors

$$\{U_i \times V_j\}_{\substack{i\in I\\ j\in J}} \tag{7.70}$$

30. Définition 7.33.
31. Topologie produit, définition 7.14

est un recouvrement de M par des ouverts de $X \times Y$. Puisque M est un compact de $X \times Y$, nous pouvons en extraire un sous-recouvrement fini, c'est-à-dire I_0 fini dans I et J_0 fini dans J tels que

$$\{U_i \times V_j\}_{\substack{i\in I_0\\ j\in J_0}} \tag{7.71}$$

soit encore un recouvrement de $K \times L$. Nous prouvons à présent que $\{U_i\}_{i\in I_0}$ est un recouvrement de K, ce qui montrera que K est un compact.

Soit $a \in K$. Il existe $b \in B$ tel que $(a, b) \in M$. Donc il existe $i_0 \in I_0$ et $j_0 \in J_0$ tels que $(a, b) \in U_{i_0} \times V_{j_0}$. En particulier $a \in U_i$.

Le même raisonnement montre que $\{V_j\}_{j\in J_0}$ est un recouvrement de L. □

7.5.5 Compactifié d'Alexandrov

Proposition-Définition 7.88 ([175]).
Soit un espace topologique séparé localement compact[32] *X. Nous considérons un élément $\omega \notin X$ et l'ensemble $\hat{X} = X \cup \{\omega\}$. Nous nommons « ouverts de $\hat{X}$ » les parties suivantes :*

— *les ouverts de X,*

— *les parties de la forme $A \cup \{\omega\}$ avec $X\backslash A$ compact dans X.*

Alors $\hat{X}$ est un espace topologique compact (cela justifie le nom « ouvert » donné aux parties sus-définies).

Démonstration. La première chose à faire est de prouver que $\hat{X}$ est bien un espace topologique (définition 7.1). Nous notons τ la topologie sur X et $\hat{\tau}$ l'ensemble des « ouverts » de $\hat{X}$. Le but est de prouver que $\hat{\tau}$ est une topologie.

(i) **L'espace lui-même** $\hat{X} \in \hat{\tau}$ parce que $\hat{X} = X \cup \{\omega\}$ et que $X\backslash X = \emptyset$ est compact.

(ii) **Le vide** $\emptyset \in \tau \subset \hat{\tau}$.

(iii) **Union quelconque** Soient A_i $(i \in I)$ des éléments de $\hat{\tau}$. Nous posons $I_1 = \{i \in I$ tel que $A_i \subset X\}$ et $I_2 = I\backslash I_1$. Nous avons

$$\bigcup_{i\in I} A_i = \big(\bigcup_{i\in I_1} A_i\big) \cup \big(\bigcup_{i\in I_2} A_i\big) = B \cup \big(\bigcup_{i\in I_2} B_i \cup \{\omega\}\big) \tag{7.72}$$

où B et les B_i sont des ouverts de X tels que $X\backslash B_i$ est compact dans X. Nous récrivons ça sous la forme

$$\bigcup_{i\in I} A_i = B \cup \big(\bigcup_{i\in I_2} B_i\big) \cup \{\omega\}. \tag{7.73}$$

La question est de savoir si

$$X\backslash\Big(B \cup \big(\bigcup_{i\in I_2} B_i\big)\Big) \tag{7.74}$$

est compact dans X. Un peu de réécriture :

$$X\backslash\Big(B \cup \big(\bigcup_{i\in I_2} B_i\big)\Big) = (X\backslash B) \cap X\backslash\big(\bigcup_{i\in I_2} B_i\big) = (X\backslash B) \cap \big(\bigcap_{i\in I_2}(X\backslash B_i)\big). \tag{7.75}$$

La partie $X\backslash B$ est fermée dans X parce que B est ouverte. La proposition 7.84 dit qu'une intersection de compacts est compacte (parce que X est séparé). Nous sommes donc en présence de l'intersection entre un compact et un fermé.

Tout compact d'un espace séparé est fermé[33]. Donc nous sommes en présence de l'intersection de deux fermés. Donc $(X\backslash B) \cap \big(\bigcap_{i\in I_2}(X\backslash B_i)\big)$ est fermé. Mais c'est contenu dans le compact $\bigcap_{i\in I_2}(X\backslash B_i)$. Fermé dans un compact, donc compact (lemme 7.82).

32. Définition 7.72.
33. Lemme 7.82(2).

(iv) **Intersection finie** Nous considérons les « ouverts » $(A_i)_{i=1,\dots,n}$ de $\hat{X}$. Si ce sont tous des ouverts de X, l'intersection est un ouvert de X et on est bon.
Supposons que tous les A_i soient de la forme $A_i = B_i \cup \{\omega\}$ avec $X\backslash B_i$ compact. Alors

$$\bigcap_{i=1}^{n}(B_i \cup \{\omega\}) = \left(\bigcap_{i=1}^{n} B_i\right) \cup \{\omega\} \tag{7.76}$$

Mais le lemme 1.25 (appliqué un nombre fini de fois) donne

$$X\backslash\left(\bigcap_{i=1}^{n} B_i\right) = \bigcup_{i=1}^{n}(X\backslash B_i) \tag{7.77}$$

qui est compact en tant qu'union finie de compacts [34].
Enfin, nous supposons que les A_i sont un mélange des deux types, nous les séparons entre ceux qui sont directement des ouverts de X et les autres :

$$A_i = \begin{cases} B_i & \text{si } i \leqslant q \\ C_i \cup \{\omega\} & \text{si } q < i \leqslant n \end{cases} \tag{7.78}$$

où B_i sont ouverts et C_i sont des parties de X telles que $X\backslash C_i$ est compacte.
Nous avons

$$\bigcap_{i=1}^{n} A_i = \left(\bigcap_{i=1}^{q} B_i\right) \cap \left(\bigcap_{i=q+1}^{n} (C_i \cup \{\omega\})\right) \tag{7.79a}$$

$$= B \cap \left(\bigcap_{i=q+1}^{n} C_i\right). \tag{7.79b}$$

Justifications.

— Nous avons posé B est l'intersection des B_i.

— Vu que ω n'est pas dans B, nous pouvons l'oublier dans les $C_i \cup \{\omega\}$.

C'est le moment d'étudier $E = \bigcap_{i=q+1}^{n} C_i$. Nous avons

$$X\backslash E = X\backslash\left(\bigcap_{i=q+1}^{n} C_i\right) = \bigcup_{i=q+1}^{n}(X\backslash C_i). \tag{7.80}$$

Vu que $X\backslash C_i$ est compact, il est fermé [35]. La partie $X\backslash E$ est donc fermée comme union finie de fermés [36]. Et donc E est ouvert. Et finalement

$$\bigcap_{i=1}^{n} A_i = B \cap E \tag{7.81}$$

est un ouvert de X comme intersection d'ouverts. C'est donc aussi un ouvert de $\hat{X}$.

Nous avons fini de prouver que $(\hat{X}, \hat{\tau})$ est un espace topologique. Nous montrons à présent que $\hat{X}$ est compact.

Soit $\{A_i\}_{i\in I}$ un recouvrement de $\hat{X}$ par des ouverts. Pour au moins un $i_0 \in I$ nous avons $\omega \in A_{i_0}$. Nous posons $A_{i_0} = B \cup \{\omega\}$ avec $X\backslash B$ compact dans X.

Les ouverts $\{A_i\}_{i\in I}$ forment un recouvrement de $X\backslash B$ par des ouverts. Nous pouvons en extraire un sous-recouvrement fini :

$$X\backslash B \subset \bigcup_{i\in I_1} A_i. \tag{7.82}$$

34. Lemme 7.83.
35. Par le lemme 7.82(2) et le fait que nous considérons un espace séparé.
36. Par le lemme 7.6(2).

Nous avons alors

$$\hat{X} \subset \bigcup_{i \in I_1 \cup \{i_0\}} A_i. \tag{7.83}$$

Et voilà que $\hat{X}$ est recouvert par un nombre fini des A_i. Notez que (7.83) est une égalité, mais nous n'en avons pas besoin. □

7.89.
Oh bien entendu, les plus férus de questions embarrassantes demanderont, si X est l'espace considéré, où prendre ce ω ? Quel « objet » existe en-dehors de X ? Qui m'assure que X n'est pas tellement grand que tout est dedans ? Le fait est qu'il n'existe pas d'ensemble contenant tous les ensembles (c'est le corolaire 1.144). Nous pouvons donc toujours trouver un ensemble ω qui n'est pas dans X.

En ce qui concerne $\mathbb{R}$ auquel nous pouvons attacher deux infinis ($+\infty$ et $-\infty$), ce sera la définition 12.27.

Pour $\mathbb{C}$, nous donnerons une caractérisation de la limite en ∞ dans le lemme 12.82.

7.5.6 Propriété d'intersection finie

Définition 7.90 (Propriété d'intersection finie[176]).
*Soit un ensemble X. Une famille non vide $\mathcal{A}$ de parties de X a la **propriété d'intersection finie** si toutes les intersections finies d'éléments de $\mathcal{A}$ est non vide.*

Théorème 7.91 ([176]).
Un espace est compact si et seulement si toute famille de parties fermées ayant la propriété d'intersection finie[37] *a une intersection non vide.*

7.6 Limite de fonction

Définition 7.92 (Limite d'une fonction, thème 28[177]).
*Soient des espaces topologiques X et Y ainsi que $\Omega \subset X$ et $a \in \mathrm{Adh}(\Omega)$. Soit une application $f\colon \Omega \to Y$. Nous disons que l'élément ℓ de Y est une **limite** de f en a lorsque pour tout ouvert V contenant ℓ, il existe un voisinage ouvert U de a tel que*

$$f\big(U \cap \Omega \backslash \{a\}\big) \subset V. \tag{7.84}$$

Si un tel élément est unique[38]*, alors nous disons que cet élément est la **limite** de f et nous notons*

$$\lim_{x \to a} f(x) = \ell. \tag{7.85}$$

7.93.
Il aurait été tout aussi bien de définir la limite d'une fonction $f\colon X \to Y$ définie sur tout X, puis de considérer Ω avec la topologie induite depuis X.

Remarque 7.94.
Nous ne saurions trop insister sur le fait que la valeur de f en a n'intervient pas dans la définition de la limite de f en a. Il n'est même pas nécessaire que f soit définie en a pour que l'on puisse parler de limite de f en a. Par exemple nous avons

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = 2, \tag{7.86}$$

alors que la fonction n'est pas définie en $x = 1$.

37. Définition 7.90.
38. Rappelons que ce n'est pas toujours le cas, mais que ça l'est si l'espace topologique est séparé – définition 7.51.

Plus généralement, un peu par principe, toutes les fois que la notion de limite apporte une information, le point où l'on prend la limite est spécial. Sinon on ne calculerait pas la limite, mais on regarderait directement la valeur de la fonction. Cela est typiquement le cas lorsque nous aborderons les dérivées. En effet, regardons (en faisant semblant d'anticiper) la définition (12.160). Dans la formule

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}, \tag{7.87}$$

la fonction sur laquelle nous prenons la limite n'est *jamais* définie en $x = a$.

Proposition 7.95 (Unicité de la limite pour un espace séparé)**.**
Soient X un espace topologique, A une partie de X et Y un espace topologique séparé[39]*. Nous considérons une fonction $f\colon A \to Y$. Si $a \in \mathrm{Adh}(A)$, 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 :

$$\begin{cases} f(W \cap A) \subset V & \text{(7.88a)} \\ f(W' \cap A) \subset V'. & \text{(7.88b)} \end{cases}$$

Quitte à prendre des sous-ensembles nous pouvons supposer que W et W' sont ouverts. Il s'ensuit alors que :

— l'ensemble $W \cap W'$ est un ouvert contenant a et intersecte donc A ;

— l'ensemble $(W \cap W') \cap A$ est donc non vide ;

— et donc, $f(W \cap W' \cap A)$ est, lui aussi, non vide.

Mais

$$f(W \cap W' \cap A) \subset f(W \cap A) \subset V, \tag{7.89}$$

et

$$f(W \cap W' \cap A) \subset f(W' \cap A) \subset V', \tag{7.90}$$

d'où V et V' ont une intersection. Puisque ces ensembles sont arbitraires, nous avons prouvé que tout voisinage de y et tout voisinage de y' ont une intersection non vide ; étant donné que Y est séparé, nous devons avoir $y = y'$. □

Proposition 7.96.
À propos de séparation.

(1) Tout espace métrique est séparé.

(2) Si une suite dans un espace métrique possède une limite, alors elle est unique.

Démonstration. Si deux éléments x et y sont distincts, alors en posant $r = d(x, y)/3 > 0$, les boules $B(x, r)$ et $B(y, r)$ sont disjointes.

En ce qui concerne les limites, ce sont les propositions 7.53 et 7.95. □

7.7 Topologie, distances et normes

Certains ensembles ont plus de structures qu'une topologie. Nous fixons quelques bases maintenant, et nous détaillerons certains résultats plus tard.

39. Définition 7.51.

7.7.1 Distance et topologie métrique

Définition 7.97.
*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.

Théorème-Définition 7.98 (Topologie métrique)**.**
*Soit (E, d) un espace métrique. Nous définissons les **boules ouvertes** par*

$$B(a, r) = \{x \in E \text{ tel que } d(a, x) < r\}. \tag{7.91}$$

pout tout $a \in E$ et $r > 0$. Alors en posant

$$\mathcal{T} = \big\{\mathcal{O} \subset E \text{ tel que } \forall a \in \mathcal{O}, \exists r > 0 \text{ tel que } B(a, r) \subset \mathcal{O}\big\} \tag{7.92}$$

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 $\emptyset \in \mathcal{T}$ parce que tout élément de l'ensemble vide ...heu ...enfin parce que, d'accord hein [40]. Ensuite si les $\{A_i\}_{i\in I}$ 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 les $\{A_i\}_{i\in\{1,\dots,n\}}$ 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 $B(x, r) \subset \bigcap_{i=1}^n A_i$. □

Proposition 7.99.
La topologie sur un espace métrique [41] est la topologie engendrée [42] par ses boules ouvertes.

7.100.
Si vous avez un peu de temps, vous pouvez vérifier que si $\mathbb{K}$ est un corps totalement ordonné, alors avec toutes les définitions de 1.318, en posant $d(x, y) = |x - y|$ nous avons une distance sur $\mathbb{K}$.

De plus, les boules définies en 1.318 sont alors les mêmes que celles définies en (7.91), ce qui donne à tout corps totalement ordonné une structure d'espace topologique.

Proposition 7.101 ([1])**.**
Soient un espace métrique (E, d), ainsi qu'une suite convergente $a_n \xrightarrow{d} \ell$. Il existe $r > 0$ tel que pour tout n nous ayons $d(\ell, a_n) < r$.

Démonstration. Soit $r_1 > 0$ et $N \in \mathbb{N}$ tel que $n \geqslant N$ implique $d(\ell, a_n) < r_1$. Ensuite nous posons $r_2 = \max\{d(\ell, a_n)\}_{n=0,\dots,N}$.

Pour tout n nous avons $d(a_n, \ell) \leqslant r_1 + r_2$. □

40. Pour qui ne serait pas d'accord, ajoutez $\emptyset$ dans la définition des ouverts et puis c'est tout.
41. Définition 7.98.
42. Définition 7.11

7.7.2 Topologie métrique et induite

Lemme 7.102.
Soit un espace vectoriel normé $\big(V, \|.\|\big)$ muni d'un sous-espace vectoriel M. La topologie induite de M depuis V est la même que la topologie métrique $\big(M, \|.\|\big)$.

7.7.3 Intérieur, adhérence et frontière

7.103.
Choses déjà faites :

— Intérieur, définition 7.26.

— Adhérence, qui est la même chose que fermeture, définition 7.28, et précisé par le lemme 7.29.

Dans le cas de $\mathbb{R}^n$ dans lequel les boules forment une base de la topologie nous pouvons encore préciser de la façon suivante :

$$x \in \operatorname{Adh} A \overset{\text{def}}{\iff} \forall \epsilon > 0, B(x, \epsilon) \cap A \neq \varnothing \tag{7.93}$$

Proposition 7.104.
Pour $A \subset \mathbb{R}^n$, nous avons

$$\operatorname{Int} A \subseteq A \subseteq \operatorname{Adh} A$$

Définition 7.105.
*La **frontière** ou le **bord** de A est défini par $\partial A = \operatorname{Adh} A \backslash \operatorname{Int} A$.*

Lemme 7.106.
Une partie A d'un espace topologique est ouverte si $A = \operatorname{Int} A$, et fermée si $A = \operatorname{Adh} A$.

Lemme 7.107 (Caractérisation équivalente de la frontière)**.**
Soient X un espace topologique et $S \subset X$. 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 7.28), 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. □

Corolaire 7.108.
Un ensemble et son complémentaire ont même frontière.

Démonstration. Conséquence du lemme 7.107. 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 7.109.
Soit $X = [0, 1]$ muni de la topologie de la distance $|x - y|$ (définition 7.98). 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

$$B(1, r) = \{x \in X \text{ tel que } |x - 1| < r\} =]1 - r, 1] \tag{7.94}$$

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 7.110 (Passage de douane[178, 179])**.**
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 $\mathrm{Int}(A)$ 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 7.60 de la connexité). Donc γ doit contenir des points qui sont dans $\bar{A}$ mais pas dans $\mathrm{Int}(A)$. C'est-à-dire des points de ∂A. □

On vérifiera que les notations et les dénominations sont cohérentes en prouvant la proposition suivante.

Proposition 7.111.
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 7.112.
Si $A \subset \mathbb{R}^n$ et $A^c = \mathbb{R}^n\backslash A$, nous avons

(1) $(\mathrm{Int}\, A)^c = \mathrm{Adh}(A^c)$ et $(\mathrm{Adh}\, A)^c = \mathrm{Int}(A^c)$,

(2) A est ouvert si et seulement si A^c est fermé,

(3) $\mathrm{Int}\, A$ est le plus grand ouvert contenu dans A,

(4) $\mathrm{Adh}\, A$ est le plus petit fermé contenant A,

Exemple 7.113.
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 $\bar{A} \cap \bar{B} = \{1\}$. △

7.7.4 Boules ouvertes, fermées, sphères

Définition 7.114.
Soit un espace métrique (E,d).

*(1) Nous nommons **boule fermée** la fermeture de la boule ouverte, c'est à dire les parties de la forme $\overline{B(a,r)}$.*

*(2) La **sphère** de centre $a \in E$ et de rayon $r \in \mathbb{R}^+$ est la frontière[43] de la boule : $S(a,r) = \partial B(a,r)$.*

7.115.
Les différences entre boules ouverts, fermées et sphères 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.

Lemme 7.116.
Quelque liens entre les boules et les sphères.

(1) La sphère est donnée par $S(a,r) = \{x \in V \text{ tel que } d(x,a) = r\}$.

(2) La fermeture de la boule est $\overline{B(a,r)} = \{x \in V \text{ tel que } d(x,a) < r\}$;

(3) Nous avons $\overline{B(a,r)} = B(a,r) \cup S(a,r)$.

43. Frontière, définition 7.105.

7.7.5 Continuité séquentielle

Corolaire 7.117 (Caractérisation séquentielle de la continuité en un point[1])**.**
Une application entre deux espaces topologiques continue en un point y est séquentiellement continue.

Démonstration. Soit une application $f\colon X \to Y$ entre les espaces topologiques X et Y. Nous supposons que f est continue en $a \in X$. Soit une suite convergente $x_k \xrightarrow{X} a$. Nous devons prouver que $f(x_k) \to f(a)$.

Soit un voisinage V de $f(a)$ dans Y. Le fait que f soit continue en a signifie[44] que $f(a)$ est une limite de f en a, c'est-à-dire[45] qu'il existe un voisinage W de a tel que $f(W\backslash\{a\}) \subset V$.

Puisque $x_k \to a$, il existe N tel que $x_k \in W$ pour tout $k \geqslant N$. Pour ces valeurs de k, nous avons $f(x_k) \in V$.

Nous avons prouvé que pour tout voisinage V de $f(a)$ dans Y, il existe N tel que $f(x_k) \in V$ dès que $k \geqslant N$. Cela signifie exactement que $f(x_k) \to f(a)$. □

7.7.5.1 Les boules, une base de topologie

Proposition 7.118.
Un espace métrique séparable[46] *accepte une base de topologie*[47] *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

$$\mathcal{B} = \{B(a_i, r_j)\}_{i,j\in\mathbb{N}} \tag{7.95}$$

est une base de la topologie[48] *de E.*

Démonstration. Soient $x \in E$ 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 $r \in \mathbb{Q}$ (existence d'un tel rationnel par le lemme 1.375).

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

$$\|y - x\| \leqslant \|y - a\| + \|a - x\| < \frac{2}{3}r + \frac{1}{3}r = r. \tag{7.96}$$

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 dans V. Cela prouve que $\mathcal{B}$ est bien une base de la topologie de E. □

Remarque 7.119.
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 11.45).

Voir aussi le corolaire 14.223 qui donnera tout ouvert comme union de pavés presque disjoints.

Définition 7.120.
Soit (X, d) un espace métrique. Un sous-ensemble $A \subset X$ est ***borné*** *si il existe une boule de X contenant A.*

Proposition 7.121.
Toute réunion finie d'ensembles bornés est un ensemble borné. Toute partie d'un ensemble borné est un ensemble borné.

44. C'est la définition 7.41 de la continuité en un point.
45. Définition 7.92 d'une limite.
46. Qui possède une partie dense dénombrable, définition 7.52.
47. Base de topologie, définition 7.2.
48. Définition 7.2.

subsectionContinuité et compacité

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 maximums et minimums de fonctions. Le théorème exact est le suivant.

Lemme 7.122 (de Lebesgue[180])**.**
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. De ces 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 7.123 ([180])**.**
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 de parties finies de X autour des points desquels les boules de taille ϵ recouvrent X.

Nous construisons par récurrence une suite ne possédant pas de sous-suite convergente. Le premier terme, x_0 est pris arbitrairement dans X. Ensuite si nous avons déjà N termes de la suite, nous savons que les boules de rayon ϵ centrées sur 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 7.124 (Bolzano-Weierstrass[180], thème 31)**.**
Un espace métrique est compact si et seulement si toute suite admet une sous-suite qui converge à l'intérieur de l'espace.

Démonstration. Soient X un espace métrique compact et (x_n) une suite dans X. Nous considérons la suite de fermés emboîtés

$$X_n = \overline{\{x_k \text{ tel que } k > n\}}. \tag{7.97}$$

Ce sont des fermés ayant la propriété d'intersection finie non vide, et donc la proposition 7.78 nous dit qu'ils ont une intersection non vide. Un élément de cette intersection est automatiquement un point d'accumulation de la suite [49].

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 7.122, 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 7.123, nous considérons un ensemble fini $\{y_i\}_{i \in A}$ tel que les 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 7.125 (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-suite convergente. △

Le théorème de Bolzano–Weierstrass 7.124 a l'importante conséquence suivante.

49. Définition 7.39.

Théorème 7.126 (Weierstrass).
Une fonction continue à valeurs réelles définie sur un compact est bornée et atteint ses bornes.

Démonstration. Soient 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 :

$$A = f(K) = \{f(x) \text{ tel que } x \in K\}. \tag{7.98}$$

Nous considérons le supremum $M = \sup A = \sup_{x\in K} f(x)$ avec la convention suivante : si A n'est pas borné supérieurement, nous posons $M = \infty$ (voir définition 1.393).

Nous allons maintenant construire une suite (x_n) de deux façons différentes selon que $M = \infty$ ou non.

(1) Si $M = \infty$, nous choisissons, pour chaque $n \in \mathbb{N}$, un $x_n \in K$ tel que $f(x_n) > n$. C'est certainement possible parce que si A n'est pas borné, nous pouvons y trouver des nombres aussi grands que nous voulons.

(2) Si $M \neq \infty$, 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-Weierstrass 7.124. Afin d'alléger la notation, nous allons noter (x_n) la sous-suite convergente. Nous avons donc

$$x_n \to x \in K. \tag{7.99}$$

Par la proposition 7.117, nous savons que f prend en x la valeur

$$f(x) = \lim_{n\to\infty} f(x_n). \tag{7.100}$$

Donc $f(x) < \infty$. Évidemment, 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

$$M - \frac{1}{n} < f(x_n) \leqslant M. \tag{7.101}$$

En passant à la limite $n \to \infty$, ces inégalités deviennent

$$M \leqslant f(x) \leqslant M, \tag{7.102}$$

et donc $f(x) = M$, ce qui prouve que f atteint sa borne M au point $x \in K$. □

Lemme 7.127 ([1]).
Soient des compacts A, B et une fonction continue $f\colon A \times B \to \mathbb{R}$. Alors

$$\sup_{(x,y)\in A\times B} |f(x,y)| = \sup_{x\in A}\big(\sup_{y\in B} |f(x,y)|\big). \tag{7.103}$$

Démonstration. Pour chaque $x \in A$, la fonction $f_x\colon B \to \mathbb{R}$ donnée par $f_x(y) = |f(x,y)|$ est continue et atteint donc sa borne[50] en $y_M(x)$. Notons que cela ne définit pas de façon univoque $y_M(x)$ parce que f_x peut atteindre son maximum en plusieurs points. L'important est que pour tout x, le nombre $|f\big(x, y_M(x)\big)|$ ne dépend pas du choix de $y_M(x)$ parmi les y qui réalisent le maximum.

Notons (x_0, y_0) un point de $A \times B$ sur lequel $|f|$ réalise son maximum[51] :

$$\sup_{(x,y)\in A\times B} |f(x,y)| = |f(x_0, y_0)|. \tag{7.104}$$

50. Théorème 7.126.
51. Encore une fois, ce point n'est pas déterminé de façon unique par cette propriété.

Nous avons d'une part

$$\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big)=\sup_{x\in A}|f\big(x,y_M(x)\big)|\leqslant|f(x_0,y_0)| \tag{7.105}$$

Et d'autre part, quelques calculs avec justifications en-dessous :

$$\begin{aligned}\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big)&\leqslant\sup_{x\in A}\sup_{y\in B}|f(x_0,y_0)| &(7.106a)\\ &=|f(x_0,y_0)| &(7.106b)\\ &\leqslant|f\big(x_0,y_M(x_0)\big)| &(7.106c)\\ &\leqslant\sup_{x\in A}|f\big(x,y_M(x)\big)| &(7.106d)\\ &\leqslant\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big). &(7.106e)\end{aligned}$$

Justifications.

— Pour (7.106a). Le point (x_0,y_0) est un maximum de $|f|$.

— Pour (7.106c). y_M est définie pour maximiser, en fonction de x, la quantité $|f(x,y_M(x))|$.

— Pour (7.106d). Au lieu de conserver la valeur x_0 fixé, nous prenons le maximum sur tous les x possibles.

Vu que les premiers et derniers termes des inégalités (7.106) sont égaux, toutes les inégalités sont en réalité des égalités. En particulier, en reprenant (7.104),

$$\sup_{(x,y)\in A\times B}|f(x,y)|=|f(x_0,y_0)|=\sup_{x\in A}\big(\sup_{y\in B}|f(x,y)|\big). \tag{7.107}$$

□

7.7.6 Distance à un ensemble

Définition 7.128.
Si A est une partie de l'espace métrique (X,d), et si $b\in X$, nous définissons

$$d(b,A)=\inf_{y\in A}d(b,y). \tag{7.108}$$

Lemme 7.129 ([1])**.**
Si A est fermé dans (X,d), et si $b\in X$ vérifie $d(b,A)=0$, alors $b\in A$.

Démonstration. Puisque A est fermé, le complémentaire A^c est ouvert (c'est la définition 7.3). Supposons que $b\in A^c$. Alors il existe $r>0$ tel que $B(b,r)\subset A^c$. Si $a\in A$ nous avons alors $d(b,A)\geqslant r$ et donc $d(b,A)\geqslant r>0$. Cela contredit l'hypothèse $d(b,A)=0$.

Nous en déduisons que b n'est pas dans A^c et qu'il est donc dans A. □

Exemple 7.130 (Pas avec un ouvert)**.**
En prenant l'ouvert $A=]0,1[$ dans $\mathbb{R}$ nous avons $d(0,A)=0$, alors que 0 n'est pas dans A. △

Lemme 7.131 ([1])**.**
Soient un espace métrique (X,d) ainsi qu'une partie $A\subset X$. Soit $r>0$. La partie

$$\mathcal{O}=\{x\in X\ \text{tel que}\ d(x,A)<r\} \tag{7.109}$$

est ouverte.

Démonstration. Soit $y\in\mathcal{O}$; nous avons $d(y,A)<r$. Autrement dit,

$$\inf_{a\in A}d(y,a)<r \tag{7.110}$$

et donc il existe $a \in A$ tel que $d(y,a) < r$. Soit $\delta = d(y,a) < r$. Nous montrons à présent que $B(y, r-\delta)$ est dans $\mathcal{O}$. En effet si $z \in B(y, r-\delta)$, alors

$$d(z,a) \leqslant d(z,y) + d(y,a) < r - \delta + \delta = r. \tag{7.111}$$

□

Lemme 7.132 ([1]).
Soit un fermé F de l'espace métrique (X,d). Si $a \in X$ vérifie $d(a,F) = 0$, alors $a \in F$.

Démonstration. Supposons que $d(a,F) = 0$, c'est à dire que $\inf_{x\in F} d(a,x) = 0$. Il existe donc une suite (x_k) dans F telle que $d(a,x_k) \to 0$.

Cela signifie que $x_k \xrightarrow{(X,d)} a$. La proposition 7.47 nous dit alors que $a \in F$. □

Lemme 7.133 ([1]).
Si A est une partie de (X,d), alors la fonction

$$\begin{aligned} f\colon \Omega &\to [0,\infty[\\ x &\mapsto d(x,A) \end{aligned} \tag{7.112}$$

est continue.

7.7.7 Convexité

Définition 7.134 (Partie convexe).
*Une partie A d'un espace vectoriel est **convexe** si pour tout $a,b \in A$ et pour tout $t \in [0,1]$, le point $ta + (1-t)b$ est dans A.*

Autrement dit, une partie est convexe lorsqu'elle contient tous les segments joignant ses points.

Proposition 7.135 ([181]).
Toute intersection de convexes est convexe.

Démonstration. Soit un espace vectoriel E ainsi que des parties convexes $\{C_i\}_{i\in I}$ indexées par un ensemble quelconque I. Nous prouvons que $C = \bigcap_{i\in I} C_i$ est convexe.

Soient $x,y \in C$, ainsi que $i \in I$. Nous avons $x,y \in C_i$ et donc $\{tx + (1-t)y\}_{t\in[0,1]} \subset C_i$. Vu que cela est vrai pour tout i, nous avons

$$\{tx + (1-t)y\}_{t\in[0,1]} \subset \bigcap_{i\in I} C_i, \tag{7.113}$$

et donc le résultat attendu. □

7.7.8 Norme

Définition 7.136 ([182], thème 24).
*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 $N\colon E \to [0,\infty[$ telle que*

(1) $N(x) \geqslant 0$ pour tout x.

(2) $N(x) = 0$ si et seulement si $x = 0$;

(3) $N(\lambda x) = |\lambda| N(x)$ pour tout $\lambda \in \mathbb{R}$ et $x \in E$;

(4) $N(x+y) \leqslant N(x) + N(y)$

pour tout $x,y \in E$ et pour tout $\lambda \in \mathbb{K}$.

*La propriété (4) est appelée **inégalité triangulaire**.*

*Un espace vectoriel muni d'une norme est un **espace vectoriel normé**.*

En prenant $\lambda = -1$ dans la propriété (3), nous trouvons immédiatement que $N(-x) = N(x)$.

Proposition 7.137.
Toute norme N sur l'espace vectoriel E vérifie l'inégalité

$$\big|N(x) - N(y)\big| \leqslant N(x-y) \tag{7.114}$$

pour tout $x, y \in E$.

Démonstration. Nous avons, en utilisant le point (4) de la définition 7.136,

$$N(x) = N(x-y+y) \leqslant N(x-y) + N(y), \tag{7.115a}$$
$$N(y) = N(y-x+x) \leqslant N(y-x) + N(x). \tag{7.115b}$$

Supposons d'abord que $N(x) \geqslant N(y)$. Dans ce cas, en utilisant (7.115a),

$$\big|N(x) - N(y)\big| = N(x) - N(y) \leqslant N(x-y) + N(y) - N(y) = N(x-y). \tag{7.116}$$

Si par contre $N(x) \leqslant N(y)$, alors nous utilisons (7.115b) et nous trouvons

$$\big|N(x) - N(y)\big| = N(y) - N(x) \leqslant N(y-x) + N(x) - N(x) = N(y-x) = N(x-y). \tag{7.117}$$

Dans les deux cas, nous avons retrouvé l'inégalité annoncée. □

Cette proposition signifie aussi que

$$-N(x-y) \leqslant N(x) - N(y) \leqslant N(x-y). \tag{7.118}$$

Le lemme suivant dit que nous pouvons remplacer l'inégalité triangulaire par la convexité de la boule unité dans la définition de norme.

Lemme 7.138 ([183]).
Soit une application $N\colon E \to [0, \infty[$ telle que

(1) $N(x) \geqslant 0$ pour tout $x \in E$,

(2) $N(x) = 0$ si et seulement si $x = 0$,

(3) $N(\lambda x) = |\lambda| N(x)$.

Alors N est une norme si et seulement si la partie

$$B = \{x \in E \text{ tel que } N(x) \leqslant 1\} \tag{7.119}$$

est convexe[52].

Démonstration. Dans les deux sens.

(i) $\Rrightarrow$ Nous supposons que N est une norme et nous prouvons que la boule B est convexe. Soient $x, y \in B$ et $\lambda \in [0,1]$. Nous avons

$$N\big(\lambda x + (1-\lambda)y\big) \leqslant N(\lambda x) + N\big((1-\lambda)y\big) \tag{7.120a}$$
$$= \lambda N(x) + (1-\lambda)N(y) \tag{7.120b}$$
$$\leqslant \lambda + (1-\lambda) \tag{7.120c}$$
$$= 1. \tag{7.120d}$$

Nous avons utilisé diverses propriétés de la norme, ainsi que la majoration $N(x), N(y) \leqslant 1$.

(ii) $\Lleftarrow$ Nous supposons que B est convexe, et nous prouvons que N vérifie l'inégalité triangulaire. Soient $x, y \in E$ que nous choisissons tous deux non nuls (sinon c'est trop facile). Nous posons

$$z = \frac{x+y}{N(x) + N(y)} \tag{7.121}$$

52. Définition 7.134.

et la subtilité sera d'écrire z de telle sorte à être une somme de deux éléments de B. L'astuce est de poser

$$\lambda = \frac{N(x)}{N(x) + N(y)}. \tag{7.122}$$

Une simple vérification montre qu'alors

$$z = \lambda \frac{x}{N(x)} + (1 - \lambda) \frac{y}{N(y)}. \tag{7.123}$$

Nous avons évidemment $x/N(x) \in B$ (et de même avec y). Puisque B est convexe, nous avons $z \in B$. Exprimons le fait que $z \in B$ à partir de la définition (7.121) :

$$\frac{N(x + y)}{N(x) + N(y)} \leqslant 1. \tag{7.124}$$

Cela signifie exactement $N(x + y) \leqslant N(x) + N(y)$.

□

7.139.
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 7.137 s'énoncera donc

$$\big|\|x\| - \|y\|\big| \leqslant \|x - y\|. \tag{7.125}$$

Un espace vectoriel E muni d'une norme est, on l'a déjà dit, un **espace vectoriel normé** ; on le notera $(E, \|.\|)$ pour distinguer la norme fixée.

Une autre inégalité utile de temps en temps.

Corolaire 7.140.
Si a et b sont dans un espace vectoriel normé, alors

$$\big|\|a - b\| - \|b\|\big| \leqslant \|a\|. \tag{7.126}$$

Démonstration. Il s'agit seulement de la proposition 7.137 avec $x = a - b$ et $y = b$. □

Lemme-Définition 7.141 (Distance induite par une norme)**.**
Soit un espace vectoriel normé $(E, \|.\|)$. Nous posons

$$d(x, y) = \|x - y\|. \tag{7.127}$$

Alors

(1) d est invariante par translations : $d(a, b) = d(a + u, b + u)$

(2) d est une distance[53] *sur E.*

*C'est la **distance induite** par la norme.*

Démonstration. Le fait que la formule (7.127) 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. Mais, pour tous $x, y, z \in E$, on a

$$d(x, y) = \|x - y\| = \|x - z + z - y\| \leqslant \|x - z\| + \|z - y\| = d(x, z) + d(z, y). \tag{7.128}$$

□

Définition 7.142.
La topologie associée à une norme est celle associée à la distance donnée en 7.141 par le théorème 7.98.

53. Définition 7.97.

Corolaire 7.143.
Un espace vectoriel normé est un espace vectoriel topologique : en d'autres mots, l'addition et la multiplication par un élément du corps sont continues.

Proposition 7.144.
La norme est une application continue sur un espace vectoriel normé.

Plus précisément, si $(E, \|.\|)$ est un espace vectoriel normé, alors l'application

$$\begin{aligned} f\colon E &\to \mathbb{R} \\ x &\mapsto \|x\| \end{aligned} \tag{7.129}$$

est continue.

Proposition 7.145.
La norme sur un espace vectoriel normé est une fonction de classe C^∞.

Lemme 7.146 ([1])**.**
Soient un espace vectoriel normé E ainsi qu'une partie libre $\{a_i\}$ dans E. Si nous avons $\|\sum_i \lambda_i a_i\| < M$, alors nous avons

$$|\lambda_i|\|a_i\| < M \tag{7.130}$$

pour chaque i.

Lemme 7.147.
Soient un espace vectoriel normé $(V, \|.\|)$ ainsi qu'une isométrie $f\colon V \to V$. Si A est une partie de V telle que $f(A) \subset A$, alors

$$\bar{A} = f(\bar{A}). \tag{7.131}$$

Nous étudierons plus en détail les espaces vectoriels topologiques à partir de la définition 7.148.

7.8 Espaces vectoriels topologiques

Définition 7.148.
Un espace vectoriel V sur le corps valué[54] *$\mathbb{K}$ muni d'une topologie est un* ***espace vectoriel topologique*** *si*

(1) la somme de deux vecteurs est une application continue $V \times V \to V$; et

(2) la multiplication par un scalaire est une application continue $\mathbb{K} \times V \to V$.

Ici, sur $V \times V$ et sur $\mathbb{K} \times V$ nous avons la topologie produit.

Dans toute la suite, nous supposons que $\mathbb{K}$ est un corps avec une topologie métrique.

On le redit quand même : le corps[55] lui-même doit avoir sa topologie. Dans la grande majorité des cas, ce corps est $\mathbb{R}$ ou $\mathbb{C}$ muni de la topologie usuelle.

Mine de rien, le fait que les deux opérations usuelles soient continues a de belles conséquences sur la topologie de l'espace...

Proposition 7.149 ([184])**.**
Soit un espace vectoriel topologique V. Pour $x \in V$ et $\lambda \in \mathbb{K}$, $\lambda \neq 0$ fixés, les fonctions T_x et M_λ définies par :

$$\begin{aligned} T_x :& V \to V \\ & y \mapsto x + y \end{aligned} \quad \text{et} \quad \begin{aligned} M_\lambda :& V \to V \\ & y \mapsto \lambda y \end{aligned} \tag{7.132, 7.133}$$

sont des automorphismes[56] *de l'espace topologique V.*

54. Définition 1.405.
55. Définition 1.237
56. Définition 7.44.

Démonstration. Ce sont des bijections continues, dont les inverses sont respectivement T_{-x} et $M_{1/\lambda}$. □

Corolaire 7.150 (Invariance de la topologie[184]).
Toute base de voisinage de 0 *se transporte en tout point de l'espace vectoriel topologique.*

Plus précisément, si $\{A_i\}_{i\in I}$ *est une base de voisinage de* 0*, alors* $\{A_i + a\}_{i\in I}$ *est une base de voisinage de* a*.*

7.8.0.1 Somme directe topologique

Proposition-Définition 7.151 ([185, 186, 1]).
Soit V *un espace vectoriel topologique et une décomposition en somme directe*[57] $V = V_1 \oplus V_2$*. Alors les trois conditions suivantes sont équivalentes.*

(1) La bijection

$$\begin{aligned} \psi\colon V_1 \times V_2 &\to V \\ (x_1, x_2) &\mapsto x_1 + x_2 \end{aligned} \tag{7.134}$$

est un homéomorphisme[58]*.*

(2) Les parties V_1 *et* V_2 *sont fermées dans* V *et la projection* $s\colon V_2 \to V/V_1$ *est un homéomorphisme.*

(3) Les parties V_1 *et* V_2 *sont fermées dans* V *et la projection* $\pi_2\colon V \to V_2$ *est continue.*

Lorsqu'une décomposition en somme directe vérifie ces conditions, nous disons que la décomposition est ***topologique****.*

Démonstration. Avant de commencer, nous rappelons les topologies.

— Sur V_1 et V_2 nous avons la topologie induite, définition 7.33.

— La topologie produit sur $V_1 \times V_2$ est la définition 7.14.

— La topologie quotient sur V/V_1 est la définition 7.20.

Voici quelque points qui ne dépendent pas des hypothèses (1), (2) ou (3).

(i) **La projection $p\colon V \to V/V_1$ est continue** La projection est toujours continue pour la topologie quotient ; c'est la définition même, voir 7.20.

(ii) **ψ est une bijection continue** Le fait que ψ soit continue fait partie de la définition 7.148 d'un espace vectoriel topologique. Pour que ce soit une bijection, c'est le lemme 4.131.

(iii) **s est injective** Soient $v, w \in V_2$ tels que $s(v) = s(w)$. Alors $\{v + x_1\}_{x_1\in V_1} = \{w + y_1\}_{y_1\in V_1}$. En particulier $v \in \{w+y_1\}_{y_1\in V_1}$. Il existe donc $y_1 \in V_1$ tel que $v = w+y_1$. Donc $v-w \in V_1$. Mais comme V_2 est un espace vectoriel, nous avons aussi $v - w \in V_2$. Donc $v - w \in V_1 \cap V_2 = \{0\}$.

(iv) **s est surjective** Soit $x \in V$. Nous devons trouver $v \in V_2$ tel que $s(v) = [x]$. Nous savons qu'il existe $x_1 \in V_1$ et $x_2 \in V_2$ tels que $x = x_1 + x_2$. Nous avons alors $[x] = s(x_2)$.

(v) **s est continue** Soit un ouvert $\mathcal{O}$ de V/V_1. La partie $p^{-1}(\mathcal{O})$ est ouverte dans V (proposition 7.21), et donc

$$s^{-1}(\mathcal{O}) = p^{-1}(\mathcal{O}) \cap V_2 \tag{7.135}$$

est ouvert dans V_2 parce que la topologie sur V_2 est celle induite[59] de la topologie de V.

Et maintenant on prouve les équivalences.

(i) **Si (1) alors V_1 est fermé** La partie $V_1 \times \{0\}$ est fermée dans $V_1 \times V_2$. Vu que ψ^{-1} est continue, nous en déduisons que $\psi\big(V_1 \times \{0\}\big)$ est fermée dans V par le lemme 7.45. Mais comme $\psi\big(V_1 \times \{0\}\big) = V_1$, nous avons que V_1 est fermé dans V.

57. Définition 4.130.
58. C'est à dire isomorphisme d'espaces vectoriels : bijection continue de réciproque continue. En ce qui concerne la topologie sur $V_1 \times V_2$, elle est donné par la définition 7.14.
59. Topologie induite, définition 7.33.

(ii) **(1) ⇒ (3)** Nous supposons que $\psi\colon V_1 \times V_2 \to V$ est un homéomorphisme. Nous considérons l'application

$$\begin{aligned} \sigma\colon V_1 \times V_2 &\to V_1 \times V_2 \\ (x,y) &\mapsto (0,y). \end{aligned} \tag{7.136}$$

Cette application est continue et permet d'écrire π_2 sous la forme $\pi_2 = \psi \circ \sigma$. En tant que composée d'applications continues, l'application π_2 est continue.

(iii) **(3) ⇒ (2)** L'application $\pi\colon V \to V_2$ est constante sur les classes (modulo V_1). Donc elle descend aux classes[60] en l'application

$$\begin{aligned} \tilde{\pi}\colon V/V_1 &\to V_2 \\ [x] &\mapsto \pi(x). \end{aligned} \tag{7.137}$$

Cette application est continue parce que π l'est et parce que le lemme 7.24 le dit. Le point à remarquer est que $s^{-1} = \tilde{\pi}$ parce que pour tout $x \in V_2$ nous avons

$$s\Big(\tilde{\pi}\big([x]\big)\Big) = s\big(\pi(x)\big) = s(x) = [x] \tag{7.138}$$

parce que $\pi(x) = x$ du fait que $x \in V_2$. Vu que $\tilde{\pi}$ est continue, l'application s^{-1} est également continue.

(iv) **(2) ⇒ (1)** Nous pouvons écrire la projection $\pi_2\colon V \to V_2$ comme composée $\pi_2 = s^{-1} \circ p$. En effet pour $v_1 \in V_1$ et $v_2 \in V_2$ nous avons $(s^{-1} \circ p)(v_1 + v_2) = s^{-1}\big([v_2]\big) = v_2 = \pi_2(v_1 + v_2)$. Nous savon que p est continue (construction de la topologie et tout ça), et que s^{-1} est également continue par hypothèse. Donc π_2 est continue. Étant donné que $\pi_1 + \pi_2 = \mathrm{Id}$ et que l'identité est continue, nous déduisons que π_1 est également continue.

□

7.152 ([187]).
Si V est normé, il existe une façon plus directe (mais pas spécialement plus simple) de prouver l'implication (1) ⇒ (2), et en particulier la continuité de s^{-1}. Voyez 11.66.

Si V est de Banach, la continuité de s^{-1} peut venir du théorème d'isomorphisme de Banach 27.1 parce que s est une bijection continue entre espaces de Banach. Attention toutefois à vérifier que V/V_1 est de Banach[61].

7.8.0.2 Limite dans un espace vectoriel topologique

Lemme 7.153 (Changement de variables).
Soient un espace vectoriel topologique[62] *X ainsi qu'un espace séparé Y et une application $f\colon X \to Y$. Nous supposons que $\lim_{x\to a} f(x) = \ell$. Alors $\lim_{x\to b} f(x+a-b)$ existe et vaut ℓ.*

Démonstration. Soit un voisinage V de ℓ dans Y. Il existe un voisinage U de a tel que $f(U) \subset V$. Nous posons $U' = U - a + b$. C'est un voisinage de b. En posant $g(x) = f(x+a-b)$ nous avons

$$g(U') = f(U - a + b + a - b) = f(U) \subset V. \tag{7.139}$$

Donc $\lim_{x\to b} g(x) = \ell$. C'est cette égalité qui signifie $\lim_{x\to b} f(x+a-b)$. □

Proposition 7.154 (Limite de fonction composée[188]).
Soient des fonctions $f, g\colon \mathbb{R} \to \mathbb{R}$ telles que

$$\lim_{y\to l} f(y) = z \tag{7.140a}$$

$$\lim_{x\to a} g(x) = l. \tag{7.140b}$$

60. Voir la définition 7.23.
61. Je ne l'ai pas fait, et au doigt mouillé je dirais que ça m'étonnerais que ce soit vrai pour tout choix de sous-espace vectoriel V_1. À vos risques et périls. Écrivez-moi si vous avez une idée.
62. Définition 7.148.

Nous supposons qu'il existe un intervalle ouvert I contenant l tel que $g(x) \neq l$ sur $I \backslash \{a\}$.

Alors

$$\lim_{x \to a} (f \circ g)(x) = \lim_{y \to l} f(y) = z. \tag{7.141}$$

Proposition 7.155 (Limite de fonction composée[188])**.**
Soient des fonctions $f, g \colon \mathbb{R} \to \mathbb{R}$ telles que

$$\lim_{y \to l} f(y) = z \tag{7.142a}$$

$$\lim_{x \to a} g(x) = l. \tag{7.142b}$$

Nous supposons que f est continue en l.

Alors

$$\lim_{x \to a} (f \circ g)(x) = \lim_{y \to l} f(y) = z. \tag{7.143}$$

Proposition 7.156.
Toute application linéaire entre espaces vectoriels topologiques de dimension finie est continue.

7.8.1 Corps topologique

Définition 7.157 (Anneau topologique[189])**.**
*Un **anneau topologique** est un anneau*[63] *muni d'une topologie dans laquelle l'addition et la multiplication sont continues*[64]*.*

Proposition-Définition 7.158.
Si $(\mathbb{K}, |.|)$ est un corps valué[65]*, alors l'application*

$$\begin{aligned} d \colon \mathbb{K} \times \mathbb{K} &\to \mathbb{R}^+ \\ (x, y) &\mapsto |x - y| \end{aligned} \tag{7.144}$$

est une distance[66]*.*

Un corps valué muni de sa topologie métrique[67] *est un corps topologique*[68]*.*

Lemme 7.159.
Les corps $\mathbb{R}$ et $\mathbb{C}$ sont des corps valués. Leur topologie métrique (en tant que corps valués) est leur topologie usuelle.

7.8.2 Voisinage symétrique et équilibré

Définition 7.160 (Partie symétrique[184])**.**
*Une partie U d'un espace vectoriel topologique est **symétrique** si $x \in U$ implique $-x \in U$.*

Définition 7.161 (Partie équilibrée[184])**.**
*Une partie U d'un espace vectoriel topologique V est **équilibrée** si pour tout $|\alpha| < 1$ dans $\mathbb{K}$, $\alpha U \subset U$.*

Lemme 7.162 ([184, 1])**.**
Soit un espace vectoriel topologique.

(1) Soit un ouvert A autour de 0 dans l'espace vectoriel topologique V. Il existe $\delta > 0$ dans le corps $\mathbb{K}$ et un voisinage ouvert W de 0 tel que $\lambda W \subset A$ pour tout $|\lambda| < \delta$.

63. Définition 1.38.
64. Définition 7.41(2).
65. Définition 1.405
66. Définition 7.97.
67. Définition 7.98.
68. Définition 7.157.

(2) Tout voisinage de 0 *contient un ouvert équilibré.*

(3) Tout voisinage de 0 *contient un ouvert équilibré et symétrique.*

Démonstration. En plusieurs parties.

(i) **Pour (1)** Nous savons que $0 \cdot 0 = 0$ et que l'application

$$\begin{aligned} f\colon \mathbb{K} \times V &\to V \\ \lambda, x &\mapsto \lambda x \end{aligned} \tag{7.145}$$

est continue. La partie $f^{-1}(A)$ contient $(0,0)$. Il existe donc un voisinage ouvert S de 0 dans $\mathbb{K}$ et un voisinage ouvert W de 0 dans V tel que $S \times W \subset f^{-1}(A)$.

Puisque $\mathbb{K}$ est un corps dont la topologie est métrique [69], il existe une boule $S' = B(0,\delta) \subset S$. Donc nous avons $f(S' \times W) \subset A$ et pour tout $|\lambda| < \delta$, $\lambda W \subset A$.

(ii) **Pour (2)** Soit un voisinage ouvert $\mathcal{O}$ de 0 dans V. Par le point (1), nous considérons un voisinage W de 0 et un $\delta > 0$ tel que $\lambda W \subset \mathcal{O}$ pour tout $|\lambda| < \delta$.

Nous posons

$$U = \{\lambda w < \text{ tel que } |\lambda| < \delta, w \in W\}. \tag{7.146}$$

Nous avons $U \subset A$ par définition de W. De plus U est équilibré parce que si $|\mu| < 1$, et si $x \in U$, il existe $|\lambda| < \delta$ et $w \in W$ tels que $x = \lambda w$. Alors $\mu x = \mu\lambda w$. Nous avons $|\mu\lambda| < \delta$ et donc $\mu\lambda w \in U$.

(iii) **Pour (3)** Nous considérons U équilibré comme dans (2). Ensuite nous posons $U' = U \cap (-U)$. La partie U' est symétrique, elle est ouverte (intersection d'ouverts). Et elle est équilibrée parce que si $x \in U'$ et $|\lambda| < 1$ alors :

— $x \in U$ et U est équilibré, donc $\lambda x \in U$.

— $x \in -U$ et U est équilibré, donc il existe $y \in U$ tel que $x = -y$. Pour ce y nous avons $\lambda y \in U$ et donc $\lambda x = -\lambda y \in -U$. Donc $\lambda x \in -U$.

— Au final, $\lambda x \in U \cap (-U) = U'$ et U' est équilibré.

□

Lemme 7.163 ([190]).
Soit un espace vectoriel topologique V *ainsi qu'un voisinage ouvert* $\mathcal{O}$ *de* 0 *dans* V*. Il existe des voisinages ouverts* U_1 *et* U_2 *de* 0 *dans* V *tels que*

$$U_1 + U_2 \subset \mathcal{O}. \tag{7.147}$$

Démonstration. Par définition d'un espace vectoriel topologique, l'application

$$\begin{aligned} f\colon V \times V &\to V \\ x, y &\mapsto x + y \end{aligned} \tag{7.148}$$

est continue. Donc la partie $f^{-1}(\mathcal{O})$ est un ouvert de $V \times V$ (c'est la définition 7.41(2) de la continuité). La définition 7.14 de la topologie produit, appliquée au point $(0,0) \in V \times V$ implique qu'il existe des voisinages U_1 et U_2 de 0 dans V tels que

$$U_1 \times U_2 \subset f^{-1}(\mathcal{O}). \tag{7.149}$$

Donc $f(U_1 \times U_2) \subset \mathcal{O}$ et en particulier $U_1 + U_2 \subset \mathcal{O}$. □

Proposition 7.164 ([184, 1]).
Soit V *un espace vectoriel topologique, et* $\mathcal{O}$ *un voisinage ouvert de* 0*. Il existe un voisinage ouvert* U *de* 0 *tel que*

69. Le corps $\mathbb{K}$ est un corps valué, et donc métrique par la définition 7.158.

(1) U est symétrique,

(2) U est équilibré

(3) U vérifie $U + U \subset \mathcal{O}$.

(4) U vérifie $U + U + U + U \subset \mathcal{O}$.

Démonstration. En plusieurs petits pas.

(i) **Le point de départ** Le lemme 7.163 donne des voisinages ouverts U_1 et U_2 de 0 dans V tels que $U_1 + U_2 \subset \mathcal{O}$.

(ii) **Symétrique** En posant $U' = U_1 \cap U_2 \cap (-U_1) \cap (-U_2)$, on a un sous-ensemble symétrique de U_1 et U_2 qui vérifie $U' + U' \subset W$. De plus U' est encore un voisinage ouvert de 0 dans V.

(iii) **équilibré** C'est le moment d'utiliser le lemme 7.162. La partie U' contient un voisinage ouvert U'' de 0 qui est symétrique et équilibré. Ce U'' vérifie encore $U'' + U'' \subset \mathcal{O}$.

(iv) **En 4 parties** Maintenant nous ré-appliquons tout ce que nous venons de faire à U'' pour obtenir un voisinage symétrique et équilibré de 0 tel que $U + U \subset U'$. Nous avons alors $U + U + U + U \subset \mathcal{O}$.

Notons que ce U vérifie à fortiori $U + U \subset \mathcal{O}$.

□

Lemme 7.165 ([1]).
Soit un espace vectoriel topologique V sur le corps $\mathbb{K}$. Si $\mathcal{O}$ est un ouvert autour de 0 dans V et si $\lambda \neq 0 \in \mathbb{K}$, il existe un ouvert U autour de 0 tel que $\lambda U \subset \mathcal{O}$.

Démonstration. La réponse est $U = \lambda^{-1}\mathcal{O}$. En effet par définition d'un espace vectoriel topologique, la fonction donnée par $f(x) = \lambda x$ est continue ; donc $U = f^{-1}(\mathcal{O})$ est un ouvert. De plus $\lambda U = \mathcal{O}$. □

7.8.3 Limite de suites

Si (x_n) est une suite dans un espace vectoriel topologique, rien ne garantit qu'elle ait une limite, ni qu'elle soit unique. Donc lorsque nous écrivons

$$x_n \xrightarrow{V} x, \tag{7.150}$$

nous sous-entendons seulement que x est une limite.

De même, dans la proposition 7.167, nous montrerons que $x_n + y_n \xrightarrow{V} x + y$ et $\lambda x_n \xrightarrow{V} \lambda x$. Cela signifie que si x et y sont des limites de (x_n) et (y_n), alors $x + y$ est une limite de $(x_n + y_n)$ et que λx est une limite de (λx_n).

Si V est un espace vectoriel topologique dans lequel il n'y a pas unicité de la limite [70], nous ne pouvons pas exactement dire que le processus de limite est une opération linéaire sur l'ensemble des suites convergentes.

Lemme 7.166.
Soient un espace vectoriel topologique V ainsi qu'une suite (x_n) dans V. Nous avons

$$x_n \xrightarrow{V} x \tag{7.151}$$

si et seulement si

$$x_n - x \xrightarrow{V} 0. \tag{7.152}$$

Proposition 7.167 ([1]).
Soit V, un espace vectoriel topologique. Soient deux suites convergentes $x_n \xrightarrow{V} x$ et $y_n \xrightarrow{V} y$ ainsi que $\lambda \in \mathbb{K}$. Alors

70. La proposition 7.53 dit qu'il y a unicité de la limite dans les espaces topologiques séparés.

(1)

$$x_n + y_n \xrightarrow{V} x + y. \tag{7.153}$$

(2)

$$\lambda x_n \xrightarrow{V} \lambda x. \tag{7.154}$$

Démonstration. En deux parties.

(i) **(1)** Nous allons montrer que $x_n + y_n - (x + y) \xrightarrow{V} 0$; ce sera suffisant par le lemme 7.166.
Soit un ouvert $\mathcal{O}$ autour de 0. Soient des ouverts U_1 et U_2 autour de 0 tels que $U_1 + U_2 \subset \mathcal{O}$ (lemme 7.163).
Vues les convergences de (x_n) et de (y_n), il existe un N tel que $n \geqslant N$ implique $x_n - x \in U_1$ et $y_n - y \in U_2$. Dans ce cas, $x_n + y_n - (x + y) \in U_1 + U_2 \subset \mathcal{O}$.
Donc pour $n \geqslant N$ nous avons bien $x_n - y_n - (x + y) \in \mathcal{O}$, ce qui signifie que $x_n + y_n \xrightarrow{V} x + y$.

(ii) **(2)** En plusieurs étapes.

(i) $x_n - x \xrightarrow{V} 0$ C'est le lemme 7.166.

(ii) $\lambda x_n - x \xrightarrow{V} 0$ Soit un ouvert $\mathcal{O}$ autour de 0. Par le lemme 7.165, il existe un ouvert U autour de 0 tel que $\lambda U \subset \mathcal{O}$. Comme $x_n - x \xrightarrow{V} 0$, il existe N tel que $n \geqslant N$ implique $x_n - x \in U$.
Pour ces N et n nous avons aussi $\lambda(x_n - x) \in \lambda U \subset \mathcal{O}$. Nous avons donc démontré que $\lambda x_n - \lambda x \xrightarrow{V} 0$.

(iii) **Conclusion** Encore le lemme 7.166 nous permet de déduire que $\lambda x_n \xrightarrow{V} \lambda x$.

□

7.9 Continuité de fonctions

7.9.1 Continuité

La définition de la continuité d'une fonction est donnée en 7.41.

7.168.
Lorsque nous écrivons $f\colon X \to Y$, nous entendons que f est définie sur tout X, mais pas qu'elle soit surjective sur Y. En particulier, pour que f soit continue en a, il faut que a soit dans le domaine de définition de f.

Dans le cas de fonctions $\mathbb{R} \to \mathbb{R}$, l'espace X sera la partie de $\mathbb{R}$ sur laquelle f sera définie, et la topologie sera la topologie induite de $\mathbb{R}$.

Proposition 7.169 ([191]).
Soient deux espaces topologiques X et Y. Une application $f\colon X \to Y$ est continue[71] *si et seulement si pour tout $x \in X$ et pour tout voisinage*[72] *V de $f(x)$, la partie $f^{-1}(V)$ est un voisinage de x dans X.*

Démonstration. En deux parties.

(i) $\Rightarrow$ Soient $x \in X$ et un voisinage V de $f(x)$ dans Y. Il existe alors un ouvert $\mathcal{O}$ de Y tel que $f(x) \in \mathcal{O} \subset V$.
La partie $f^{-1}(\mathcal{O})$ vérifie :
— $f^{-1}(\mathcal{O})$ est un ouvert de X parce que f est continue.
— $x \in f^{-1}(\mathcal{O})$

71. Définition 7.41.
72. Définition 7.4

— $f^{-1}(\mathcal{O}) \subset f^{-1}(V)$.

Donc $f^{-1}(V)$ contient un ouvert contenant x. Donc $f^{-1}(V)$ est un voisinage de x dans X.

(ii) $\Leftarrow$ Soit un ouvert $\mathcal{O}$ de Y. Nous devons prouver que $f^{-1}(\mathcal{O})$ est un ouvert de X. Pour cela nous prouvons que $f^{-1}(\mathcal{O})$ contient un ouvert autour de chacun de ses éléments et utilisons le théorème 7.7.

Soit donc $x \in f^{-1}(\mathcal{O})$. La partie $\mathcal{O}$ est un voisinage de $f(x)$. Donc $f^{-1}(\mathcal{O})$ est un voisinage de x. Il existe donc un ouvert V de X tel que

$$x \in V \subset f^{-1}(\mathcal{O}). \tag{7.155}$$

Nous en déduisons que $f^{-1}(\mathcal{O})$ contient bien un ouvert autour de chacun de ses points.

□

La proposition 7.258 donnera des détails sur ce qu'il se passe lorsque l'espace est métrique.

Théorème 7.170.
Une fonction $f\colon X \to Y$ est une fonction continue si et seulement si elle est continue en chacun des points de X.

Démonstration. En deux parties.

(i) $\Rightarrow$ Nous supposons que f est une fonction continue. Soient $a \in X$ et W un voisinage de $f(a)$. Nous considérons $\mathcal{O}$, un voisinage ouvert de $f(a)$ contenu dans W ; 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 W.

(ii) $\Leftarrow$ 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 7.7 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 choisi dans $f^{-1}(\mathcal{O})$. Donc la continuité de f en a 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 7.171.
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 l'exemple 7.50 et la proposition 7.95.

Lemme 7.172 ([1])**.**
Soient une fonction $f\colon X \to Y$, et un point d'accumulation $a \in X$ [73]. La fonction f est continue en a si et seulement si $f(a)$ est une limite de f en a.

Démonstration. En deux parties.

(i) $\Rightarrow$ Nous supposons que f est continue en $a \in X$. Soit un ouvert V de Y contenant $f(a)$. Par continuité de f au point [74] a, il existe un voisinage U de a tel que $f(U) \subset V$. À fortiori, $f\big(U\backslash a\big) \subset W$ comme le demande la définition de la limite.

(ii) $\Leftarrow$ Nous supposons que $f(a)$ est une limite de $f(x)$ lorsque x tend vers a. Si W est un ouvert de Y contenant $f(a)$, il existe un voisinage V de a dans X tel que $f\big(V\backslash a\big) \subset W$. Mais puisque $f(a) \in W$, nous avons $f(V) \subset W$.

□

73. Un point d'accumulation de X n'est pas spécialement dans X, si X est un sous-espace d'un autre. Par exemple 0 est un point d'accumulation de $]0,1[$ dans $\mathbb{R}$. Ici nous supposons que $a \in X$, sinon il n'y a de toute façon pas de continuité en a.

74. Continuité en un point, définition 7.41(1).

7.9.1.1 Continuité séquentielle

Définition 7.173.
*Si X et Y sont deux espaces topologiques, une fonction $f\colon X \to \mathbb{R}$ est **séquentiellement continue** en un point a si pour toute suite convergente $x_n \to a$ dans X nous avons $f(x_n) \to f(a)$ dans Y.*

7.174.
Nous allons maintenant voir deux résultats disant que si une fonction est continue, alors elle peut être permutée avec une limite de suite. Dans le cas des espaces métriques, la proposition 7.220 montrera la réciproque : si pour toute suite $x_n \to a$, nous avons $\lim_{n\to\infty} f(x_n) = y$, alors f a une limite en a qui vaut y.

Proposition 7.175 (Permuter limite et fonction continue[1])**.**
Soient deux espaces topologiques X et Y ainsi qu'une fonction $f\colon X \to Y$. Soit $a \in X$ et $\ell \in Y$. Si

$$\lim_{x\to a} f(x) = \ell, \tag{7.156}$$

alors, pour toute suite (x_k) telle que $x_k \to a$, on a

$$\lim f(x_k) = \ell. \tag{7.157}$$

Démonstration. Nous considérons une suite (x_k) qui converge vers a dans X. Soient V un voisinage de ℓ et W un voisinage de a tels que $f(W) \subset V$ (définition 7.92 de la continuité en un point). Par la convergence $a_k \to a$, il existe N tel que pour tout $k \geqslant N$, $a_k \in W$, et donc tel que $f(a_k) \in V$, ce qui donne la continuité séquentielle de f. □

7.9.1.2 Application réciproque

Définition 7.176 (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 atteints, 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 domaine sur laquelle nous la considérons.

Définition 7.177.
*Soit $f\colon A \to B$ une bijection. L'**application réciproque** de f est la fonction*

$$\begin{aligned} f^{-1}\colon B &\to A \\ y &\mapsto \text{le } x \in A \text{ tel que } f(x) = y. \end{aligned} \tag{7.158}$$

Plus généralement si $f\colon X \to Y$ est une application quelconque et si $S \subset Y$ nous notons

$$f^{-1}(S) = \{x \in X \text{ tel que } f(x) \in S\}, \tag{7.159}$$

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 7.178.
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 7.179.
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 $f(I)$.

7.9.2 Continuité et topologie induite

Proposition 7.180 ([1])**.**
Soit une fonction $f\colon X \to Y$, continue sur l'ouvert A de X au sens où elle est continue en chaque point de A. Alors la fonction restriction $\tilde{f}\colon A \to Y$ est également continue pour la topologie sur A, induite[75] *de X.*

Démonstration. Soit $a \in A$, et montrons que $\tilde{f}$ est continue en a, c'est-à-dire que $\tilde{f}(a) = f(a)$ soit une limite de $\tilde{f}$ en a. Soit un voisinage V de $\tilde{f}(a)$ dans Y. Par la continuité de f, nous avons un ouvert W de X tel que

$$f\big(W\backslash\{a\}\big) \subset V. \tag{7.160}$$

La partie $W \cap A$ est un voisinage de a pour la topologie de A, et vérifie

$$f\big(W \cap A\backslash\{a\}\big) \subset V. \tag{7.161}$$

donc $f(a)$ est une limite de $\tilde{f}$ pour $x \to a$. La fonction $\tilde{f}\colon A \to Y$ est continue en chaque point de A. □

Au niveau de la notion de continuité, il n'y a pas trop de changements en passant de $\mathbb{R}$ à $\mathbb{Q}$ muni de la topologie induite.

Exemple 7.181.
Que signifie d'être continue pour une fonction $f\colon \mathbb{Q} \to \mathbb{R}$? D'après le théorème 7.170, il s'agit d'être continue en chaque point de $\mathbb{Q}$. Il s'agit donc, par la définition 7.41 que pour tout $q \in \mathbb{Q}$, le nombre $f(q)$ soit une limite de f pour $x \to q$.

L'espace d'arrivée étant $\mathbb{R}$, un voisinage de $f(q)$ est pris comme une boule de taille ϵ. La continuité de f exige qu'il y ait un voisinage W de q dans $\mathbb{Q}$ tel que pour tout $q' \in W$ (différent de q), $|f(q) - f(q')| < \epsilon$.

Qu'est-ce qu'un ouvert dans $\mathbb{Q}$? D'après la définition 7.33 de la topologie induite, ce sont les ensembles $\mathbb{Q} \cap \mathcal{O}$ avec $\mathcal{O}$ ouvert dans $\mathbb{R}$. Tout cela pour dire que pour tout $\epsilon > 0$, il doit exister $\delta > 0$ tel que pour tout $q' \in \mathbb{Q}$ tel que $0 < |q - q'| < \delta$, nous ayons $|f(q) - f(q')| < \epsilon$.

Bref, c'est exactement le mécanisme usuel de la continuité sur $\mathbb{R}$, sauf qu'il ne faut considérer que les rationnels. △

Lemme 7.182 (Application partielle[1])**.**
Soient trois espaces topologiques X_1, X_2 et Y. Nous considérons une fonction continue $f\colon X_1 \times X_2 \to Y$ ainsi que $x_1 \in X_1$. Alors l'application

$$\begin{aligned} g\colon X_2 &\to Y \\ x_2 &\mapsto f(x_1, x_2) \end{aligned} \tag{7.162}$$

est continue.

Démonstration. Soit un ouvert $\mathcal{O}$ de Y ; par hypothèse sur f, la partie $f^{-1}(\mathcal{O})$ est ouverte dans $X_1 \times X_2$. Notre but est de prouver que $g^{-1}(\mathcal{O})$ est un ouvert de X_2. Nous avons :

$$g^{-1}(\mathcal{O}) = \{x_2 \in X_2 \text{ tel que } (x_1, x_2) \in f^{-1}(\mathcal{O})\}. \tag{7.163}$$

Nous considérons $x_2 \in g^{-1}(\mathcal{O})$ et nous prouvons qu'il existe dans X_2 un voisinage de x_2 entièrement contenu dans $g^{-1}(\mathcal{O})$.

75. Définition 7.33.

Étant donné que (x_1, x_2) est dans $f^{-1}(\mathcal{O})$ qui est ouvert, la définition 7.14 de la topologie sur $X_1 \times X_2$ nous donne des ouverts A_1 dans X_1 et A_2 dans X_2 tels que

$$(x_1, x_2) \in A_1 \times A_2 \subset f^{-1}(\mathcal{O}). \tag{7.164}$$

Nous montrons à présent que $A_2 \subset g^{-1}(\mathcal{O})$. Soit $y_2 \in A_2$. Par construction $(x_1, y_2) \in A_1 \times A_2 \subset f^{-1}(\mathcal{O})$, donc

$$g(y_2) = f(x_1, y_2) \in \mathcal{O}. \tag{7.165}$$

Cela termine la démonstration. □

7.9.3 Continuité et connexité

Proposition 7.183.
Un espace topologique X est connexe si et seulement si toute application continue $X \to \mathbb{Z}$ est constante.

Proposition 7.184.
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 des valeurs intermédiaires 10.82.

Exemple 7.185.
Les espaces topologiques $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes. △

Démonstration. Supposons par l'absurde que $f\colon \mathbb{R} \to \mathbb{R}^2$ soit un homéomorphisme. Nous posons $E = f\big(\mathbb{R}\backslash\{0\}\big)$ et $z_0 = f(0)$. Puisque f est bijective nous avons

$$E = \mathbb{R}^2\backslash\{z_0\}, \tag{7.166}$$

qui est connexe.

Comme E est connexe et que f^{-1} est continue, la proposition 7.184 nous dit que $f^{-1}(E)$ est connexe. Mais par définition, $f^{-1}(E) = \mathbb{R}\backslash\{0\}$ qui n'est pas connexe. □

7.9.4 Continuité et compacité

Théorème 7.186.
L'image d'un compact[76] *par une fonction continue est un compact.*

Dans le cadre des espaces vectoriels normés, ce théorème est démontré en la proposition 7.254.

Démonstration. Soit $K \subset X$, un ensemble compact, et étudions $f(K)$; en particulier, nous considérons Ω, un recouvrement de $f(K)$ par des ouverts. Nous avons

$$f(K) \subseteq \bigcup_{\mathcal{O}\in\Omega} \mathcal{O}. \tag{7.167}$$

76. Définition 7.66.

Par construction, nous avons aussi

$$K \subseteq \bigcup_{\mathcal{O}\in\Omega} f^{-1}(\mathcal{O}), \tag{7.168}$$

en effet, si $x \in K$, alors $f(x)$ est dans un des ouverts de Ω, disons $f(x) \in \mathcal{O}$, 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 $\{f^{-1}(\mathcal{O}_1), \dots, f^{-1}(\mathcal{O}_n)\}$ tels que

$$K \subseteq \bigcup_{i=1}^{n} f^{-1}(\mathcal{O}_i). \tag{7.169}$$

Dans ce cas, nous avons

$$f(K) \subseteq \bigcup_{i=1}^{n} \mathcal{O}_i, \tag{7.170}$$

ce qui prouve la compacité de $f(K)$. □

Lemme 7.187.
Soit un espace compact K et un espace topologique séparé X. Si $f\colon K \to X$ est une bijection continue, alors f est un isomorphisme d'espaces topologiques.

7.9.5 Topologie et matrices

Lemme-Définition 7.188 (Topologie sur les matrices)**.**
Si $\mathbb{K}$ est un corps valué[77]*, alors l'opération*

$$\begin{aligned} \|.\|_{\mathbb{M}}\colon \mathbb{M}(n\times m,\mathbb{K}) &\to \mathbb{R}^+ \\ M &\mapsto \max_{kl} \|M_{kl}\|_{\mathbb{K}}. \end{aligned} \tag{7.171}$$

est une norme[78]*.*

*Cette norme est appelée **norme maximum** et nous considérons sur $\mathbb{M}(n\times m,\mathbb{K})$ la topologie associée à cette norme*[79]

Proposition 7.189.
La multiplication matricielle est une opération continue.

7.10 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.

7.10.1 Distance et norme produit

Proposition-Définition 7.190 (Distance produit)**.**
Si (E_1, d_1),..., (E_n, d_n) sont des espaces métriques alors la formule

$$d(x,y) = \max_{i=1,\dots,n} d_i(x_i, y_i) \tag{7.172}$$

*définit une distance sur le produit cartésien $E = E_1 \times \dots \times E_n$. Elle est la **distance produit**.*

La définition de la norme sur un produit d'espaces vectoriels normés découle immédiatement de la définition de la distance 7.190 :

77. Définition 1.405.
78. Définition 7.136.
79. Définition 7.142.

Lemme-Définition 7.191 ([192]).
Soient V et W deux espaces vectoriels sur $\mathbb{K}$. Si (v_1, w_1) et (v_2, w_2) sont des éléments de $V \times W$ et si λ est un élément de $\mathbb{K}$, alors les opérations suivantes donnent une structure d'espace vectoriel au produit $V \times W$:

— $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$
— $\lambda(v_1, w_1) = (\lambda v_1, \lambda w_1)$.

Démonstration. Il faut seulement faire les vérifications d'usage. □

Lemme-Définition 7.192.
Soient deux espaces vectoriels normés V et W.

(1) L'opération
$$\|(v, w)\|_{V\times W} = \max\{\|v\|_V, \|w\|_W\}. \tag{7.173}$$
est une norme[80] *sur $V \times W$.*

(2) La topologie de cette norme[81] *est la même que la topologie produit*[82] *sur $V \times W$.*

Démonstration. En plusieurs parties.

(i) **Norme** On doit vérifier les trois conditions de la définition 7.136.

— 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 $\|a(v, w)\|_{V\times W}$ se calcule de la façon suivante :
$$\|a(v, w)\|_{V\times W} = \max\{\|av\|_V, \|aw\|_W\} = |a| \max\{\|v\|_V, \|w\|_W\} = |a|\|(v, w)\|_{V\times W}. \tag{7.174}$$

— Soient (v_1, w_1) et (v_2, w_2) dans $V \times W$.
$$\begin{aligned}\|(v_1, w_1) + (v_2, w_2)\|_{V\times W} &= \max\{\|v_1 + v_2\|_V, \|w_1 + w_2\|_W\}\\ &\leqslant \max\{\|v_1\|_V + \|v_2\|_V, \|w_1\|_W + \|w_2\|_W\}\\ &\leqslant \max\{\|v_1\|_V, \|w_1\|_W\} + \max\{\|v_2\|_V, \|w_2\|_W\}\\ &= \|(v_1, w_1)\|_{V\times W} + \|(v_2, w_2)\|_{V\times W}.\end{aligned} \tag{7.175}$$

Dans cette preuve, nous considérons la « topologie de $V \times W$ » comme étant la topologie produit et « la topologie métrique de $V \times W$ » la topologie de la norme produit.

(ii) **Équivalence** (i) **Dans un sens** La définition 7.14 de la topologie produit dit qu'une prébase de $V \times W$ est donnée par
$$\big\{B(v, r) \times B(w, s) \text{ tel que } v \in V; w \in W; r, s > 0\big\}. \tag{7.176}$$

Nous prouvons maintenant que la partie $S = B(v_0, r) \times B(w_0, s)$ est un ouvert de l'espace $\big(V \times W, \|.\|_{V\times W}\big)$. Pour cela nous prouvons que tout élément de S contient un voisinage métrique contenu dans S.
Soit $(v_1, w_1) \in S$. Nous posons
$$d\big((v_1, w_1), (v_0, w_0)\big) = \delta < \max\{r, s\}. \tag{7.177}$$

Nous considérons $\epsilon > 0$ et nous montrons que si ϵ est assez petit, $B\big((v_1, w_1), \epsilon\big) \subset S$. Pour cela nous considérons $(v, w) \in B\big((v_1, w_1), \epsilon\big)$ et nous calculons un tout petit peu :
$$d\big((v, w), (v_0, w_0)\big) \leqslant d\big((v, w), (v_1, w_1)\big) + d\big((v_1, w_1), (v_0, w_0)\big) \tag{7.178a}$$
$$< \epsilon + \delta. \tag{7.178b}$$

80. Définition 7.136.
81. Topologie associée à une norme : c'est la topologie associée à la distance correspondante, définition 7.142.
82. Topologie produit, définition 7.14.

Si ϵ est assez petit, le tout reste plus petit que $\max\{r, s\}$.
Donc S est bien un ouvert métrique par le théorème 7.7. Vu que la topologie métrique contient une prébase de la topologie produit, tout ouvert de la topologie produit est un ouvert de la topologie métrique.

(ii) **Dans l'autre sens** Soient un ouvert métrique $\mathcal{O}$ ainsi que $(v_0, w_0) \in \mathcal{O}$; il existe $r > 0$ tel que

$$B\big((v_0, w_0), r\big) \subset \mathcal{O}. \tag{7.179}$$

Nous affirmons que $B(v_0, r) \times B(w_0, r)$ est contenu dans $\mathcal{O}$, de telle sorte que $\mathcal{O}$ soit un ouvert de la topologie produit. Pour $(v_1, w_1) \in B(v_0, r) \times B(w_0, r)$ nous avons

$$d\big((v_1, w_1), (v_0, w_0)\big) = \max\{\|v_1 - v_0\|, \|w_1 - w_0\|\} < r \tag{7.180}$$

parce que $v_1 \in B(v_0, r)$ et $w_1 \in B(w_0, r)$.
Donc tout élément de $\mathcal{O}$ admet un voisinage « produit » contenu dans $\mathcal{O}$; donc $\mathcal{O}$ est ouvert pour le topologie produit.

□

7.193.
En particulier, pour la topologie de la norme maximum, la convergence d'une suite implique la convergence « composante par composante » par la proposition 7.56.

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

$$\begin{aligned} \mathtt{proj}_V \colon V \times W &\to V \\ (v, w) &\mapsto v \end{aligned} \tag{7.181}$$

et

$$\begin{aligned} \mathtt{proj}_W \colon V \times W &\to W \\ (v, w) &\mapsto w. \end{aligned} \tag{7.182}$$

Les inégalités suivantes sont évidentes

$$\begin{aligned} \|\mathtt{proj}_V(v, w)\|_V &\leqslant \|(v, w)\|_{V \times W} \\ \|\mathtt{proj}_W(v, w)\|_W &\leqslant \|(v, w)\|_{V \times W}. \end{aligned} \tag{7.183}$$

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

(1) Les projections sont des **applications ouvertes**. Cela veut dire que l'image par $\mathtt{proj}_V$ (respectivement $\mathtt{proj}_W$) de toute partie ouverte de $V \times W$ est une partie ouverte de V (respectivement W).

(2) Pour toute partie A de V et B de W, nous avons $\operatorname{Int}(A \times B) = \operatorname{Int} A \times \operatorname{Int} B$.

Une propriété moins facile à 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 11.67.

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 7.194.
Si V et W sont deux espaces vectoriels, nous pouvons considérer le produit $E = V \times W$. Les projections $\mathtt{proj}_V$ et $\mathtt{proj}_W$, définies dans la section 7.10, sont des applications linéaires.

En effet, la projection $\mathtt{proj}_V \colon V \times W \to V$ est donnée par $\mathtt{proj}_V(v, w) = v$. Alors,

$$\begin{aligned} \mathtt{proj}_V\big((v,w)+(v',w')\big) &= \mathtt{proj}_V\big((v+v'),(w+w')\big) \\ &= v+v' \\ &= \mathtt{proj}_V(v,w)+\mathtt{proj}_V(v',w'), \end{aligned} \tag{7.184}$$

et

$$\mathtt{proj}_V\big(\lambda(v,w)\big) = \mathtt{proj}_V\big((\lambda v,\lambda w)\big) = \lambda v = \lambda\,\mathtt{proj}_V(v,w). \tag{7.185}$$

Nous laissons à la lectrice le soin d'adapter ces calculs pour montrer que $\mathtt{proj}_W$ est également une projection[83]. △

Proposition 7.195.
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. Puisque $\mathcal{O}$ est un voisinage, il contient un ouvert et donc une boule

$$B\big((a,b),r\big) = \{(v,w) \in V \times W \text{ tel que } \max\{\|v-a\|,\|w-b\|\} < r\}. \tag{7.186}$$

Évidemment l'ensemble $B(a, r) \times B(b, r)$ est dedans. □

7.11 Topologie réelle en dimension n

Nous considérons sur $\mathbb{R}$ la topologie donnée par la valeur absolue, et sur $\mathbb{R}^n$ celle de la topologie produit ou du maximum, qui sont identiques par le lemme 7.192.

En particulier, nous n'avons pas encore la norme donnée par $\|x\| = \sqrt{\sum_i x_i^2}$, parce qu'elle demande la racine carré, définie en 10.85.

7.11.1 Ouverts et fermés

Proposition 7.196.
Une partie A de $\mathbb{R}^n$ est ouverte si et seulement si pour tout $a \in A$ il existe $r > 0$ tel que $B(a, r) \subset A$.

Cette proposition est évidemment à mettre en rapport avec le théorème 7.7.

Lemme 7.197.
Pour tout $x \in \mathbb{R}^n$ et tout $r > 0$ la boule[84] *$B(x, r)$ 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,

$$d(x,p') \leqslant d(x,p)+d(p,p') \leqslant d(x,p)+r-d(p,x) = r. \tag{7.187}$$

□

83. Écrivez moi si ça pose un problème.
84. Définition 7.98.

7.11.2 Point d'accumulation, point isolé

Les définitions de point d'accumulation et de point isolé sont 7.39 et 7.40. Nous voyons maintenant ce que ces définitions donnent dans le cas de l'espace topologique $\mathbb{R}$.

Lemme 7.198.
Soit $D \subset \mathbb{R}$. Un point $a \in D$ est isolé dans D si et seulement si il existe $\varepsilon > 0$ tel que

$$[a - \varepsilon, a + \varepsilon] \cap D = \{a\}. \tag{7.188}$$

Autrement dit, il existe un intervalle autour de a dans lequel a est le seul élément de D.

Lemme 7.199.
Un point $a \in \mathbb{R}$ est un point d'accumulation de D si pour tout $\varepsilon > 0$,

$$\Big([a - \varepsilon, a + \varepsilon]\backslash\{a\}\Big) \cap D \neq \emptyset. \tag{7.189}$$

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 7.200.
Prenons $D = [0, 1[\cup]2, 3]$. Cet ensemble n'a pas de point isolé, 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 7.201.
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. △

Exemple 7.202.
Soit $D =]1, 2[\cup \{12\}$. Le point 12 est adhérence, mais pas d'accumulation parce que le voisinage $]9, 14[$ n'intersectionne pas $D\backslash\{12\}$. △

7.11.3 Limite de suite

Définition 7.203 (Limite d'une suite dans $\mathbb{R}^m$)**.**
*Une suite de points (x_n) dans $\mathbb{R}^m$ est dite **convergente** si il existe un élément $\ell \in \mathbb{R}^m$ tel que*

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que } \forall n \geqslant N, \|x_n - \ell\| < \varepsilon. \tag{7.190}$$

*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 10.23.

Remarque 7.204.
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 7.205 (Unicité de la limite)**.**
Il ne peut pas y avoir deux nombres différents qui satisfont à la condition (7.190). 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

$$\|x_n - \ell\| < \varepsilon \tag{7.191}$$

pour tout $n \geqslant N$, et $N' > 0$ tel que

$$\|x_n - \ell'\| < \epsilon \tag{7.192}$$

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

$$\|\ell - \ell'\| = \|\ell - x_n + x_n - \ell'\| \leqslant \|\ell - x_n\| + \|x_n - \ell'\| < 2\varepsilon. \tag{7.193}$$

Cela prouve que $\|\ell - \ell'\| = 0$. □

Le théorème de Bolzano-Weierstrass 7.124 dit que dans le cas métrique, la compacité séquentielle est équivalente à la compacité.

7.12 Topologie et distance

Lemme 7.206.
Soient (X_1, d_1) et (X_2, d_2) des espaces métriques séparables. Alors $X_1 \times X_2$ admet une base dénombrable de topologie constituée de produits de boules 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

$$\big\{B(y_1, r_1) \times B(y_2, r_2)\big\}_{\substack{y_i \in A_i \\ r_i \in \mathbb{Q}^+}} \tag{7.194}$$

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[85], 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

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}). \tag{7.195}$$

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

$$d_i(z_i, x_i) \leqslant d(z_i, y_i) + d(y_i, x_i) \leqslant \frac{r_i}{2} + \frac{r_i}{2} = r_i. \tag{7.196}$$

Au final,

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}) \subset \mathcal{O}. \tag{7.197}$$

□

Définition 7.207.
Si (X, d_X) et (Y, d_Y) sont des espaces métriques, une ***isométrie*** *est une application bijective $f\colon X \to Y$ telle que pour tout $x, y \in X$ nous ayons*

$$d_Y\big(f(x), f(y)\big) = d_X(x, y). \tag{7.198}$$

Remarque 7.208.
Une application vérifiant (7.198) est automatiquement injective. En pratique, il ne faut donc vérifier que la surjectivité.

Exemple 7.209 (Manque de surjectivité)**.**
Si $X = [0, \infty[$ et $f(x) = x + 1$ alors f vérifie (7.198) pour la distance $d(x, y) = |x - y|$, mais n'est pas surjective. △

Proposition-Définition 7.210 (Groupe des isométries)**.**
Si (X, d) est un espace métrique,

85. Définition 7.14.

(1) l'ensemble des isométries de X, noté $\mathrm{Isom}(X)$ *est un groupe pour la composition.*

(2) Ce groupe agit fidèlement[86] *sur X.*

Proposition 7.211.
Une isométrie entre deux espaces métriques est continue.

Démonstration. Soient $f\colon X \to Y$ 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. □

Exemple 7.212.
Si X est un ensemble, nous pouvons écrire la **distance discrète** :

$$d(x,y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y. \end{cases} \tag{7.199}$$

La topologie résultante est la topologie discrète, côtoyée dans l'exemple 7.9[87].

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. △

7.12.0.1 Distance point-ensemble

Définition 7.213.
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*

$$d(x,A) = \inf_{a \in A} d(x,a). \tag{7.200}$$

FIGURE 7.1 – 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 $d(x,p)$.

7.12.1 Suites et espaces métriques

Proposition 7.214 (Caractérisation séquentielle de la limite[1])**.**
Soient deux espaces métriques X et Y ainsi qu'une fonction $f\colon X \to Y$. Soit $a \in X$ et $\ell \in Y$. On a

$$\lim_{x \to a} f(x) = \ell, \tag{7.201}$$

si et seulement si, pour toute suite (x_k) telle que $x_k \to a$, on a

$$\lim f(x_k) = \ell. \tag{7.202}$$

Par ailleurs, l'une des deux limites existe si et seulement si l'autre existe.

86. Si vous ne savez pas ce que c'est, alors vous avez zappé la définition 2.30.
87. Vérifiez-le tout de même !

Démonstration. Le sens direct est la proposition 7.175. Pour la réciproque, nous passons par la contraposée. C'est-à-dire que nous supposons que ℓ n'est pas une limite de f pour $x \to a$. Il existe un ϵ tel que pour tout δ, il existe un x vérifiant $d_X(x;a) < \delta$ et $d_Y(f(x);\ell) > \epsilon$.

Nous construisons à présent une suite de la manière suivante. Pour $\delta = \frac{1}{n}$ nous considérons x_n tel que $d_X(x_n;a) < \delta$ et $d_Y(f(x_n);\ell) > \epsilon$. Cette suite converge vers a, mais la suite $f(x_n)$ ne converge manifestement pas vers ℓ : elle ne rentre jamais dans la boule $B(\ell,\epsilon)$. □

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.

Corolaire 7.215 (Caractérisation séquentielle de la continuité en un point[1])**.**
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 en ce point.

Démonstration. Paraphrasons la preuve précédente. Nous supposons que X et Y sont 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. □

Définition 7.216 (Fermeture séquentielle[193])**.**
*Une partie F d'un espace topologique X est dit **séquentiellement fermé** si la convergence d'une suite (x_n) de F vers x implique que x appartient à F.*

Les espaces métriques ont une propriété importante que la fermeture séquentielle est équivalente à la fermeture.

Proposition 7.217 (Caractérisation séquentielle d'un fermé)**.**
Soient 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. Une suite contenue dans un fermé ne peut converger que vers un élément de ce fermé : c'était la proposition 7.47. Le point le plus important est donc l'autre sens : si toute suite d'éléments de F converge dans F alors F est fermé.

Par contraposée, 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, convergeant vers x qui n'est pas dans F. □

Lemme 7.218.
Soit X un espace métrique, et soit (x_n) une suite convergente contenue dans un ensemble $A \subset X$. Alors la limite x_n appartient à $\bar{A}$.

Ce lemme est précisément la version « espace métrique » du corolaire 7.48 ; mais, donnons-en une preuve tout de même.

Démonstration. Supposons que nous ayons une partie A de X, 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[88] $B(\ell, r) \cap A = \varnothing$. Si tous les éléments x_n de la suite sont dans A, il n'y en a donc aucun tel que $d(x_n, \ell) < r$. Cela contredit la notion de convergence $x_n \to \ell$. □

Corolaire 7.219.
Soit X un espace métrique, $A \subset X$ et $a \in \bar{A}$. Alors il existe une suite d'éléments dans A qui converge vers a.

88. Une autre manière de dire la même chose : si $\ell \notin \bar{A}$, alors $d(\ell, A) > 0$.

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.). □

En termes savants, ce corolaire signifie que la fermeture $\bar{A}$ est composé de A plus de toutes les limites de toutes les suites contenues dans A.

Proposition 7.220 (Caractérisation séquentielle de la continuité[1])**.**
Soient X et Y deux espaces topologiques séparés. Nous supposons que X est métrisable. Une application $f\colon X \to Y$ est continue sur X si et seulement si elle est séquentiellement continue sur X.

Démonstration. Le sens direct est déjà prouvé dans la proposition 7.175. Nous nous concentrons donc sur la réciproque.

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 $x_k \in X\backslash f^{-1}(\mathcal{O})$ pour tout k. Nous allons montrer que $x \in X\backslash f^{-1}(\mathcal{O})$ et la caractérisation séquentielle[89] de la fermeture conclura que $X\backslash f^{-1}(\mathcal{O})$ est fermé.

Pour tout k, nous avons $f(x_k) \in X\backslash\mathcal{O}$, et $f(x_k) \xrightarrow{Y} f(x)$ parce que f est séquentiellement continue. Vu que $f(x_k)$ est une suite dans le fermé $Y\backslash\mathcal{O}$, la limite est également dans $Y\backslash\mathcal{O}$. Nous en déduisons que $f(x) \in Y\backslash\mathcal{O}$, de telle sorte que $x \in X\backslash f^{-1}(\mathcal{O})$. □

Proposition 7.221.
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 7.117, ou proposition 7.215). Soient 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

$$f(a_n) \to f(x) \tag{7.203a}$$
$$g(a_n) \to g(x), \tag{7.203b}$$

mais pour tout n, $f(a_n) = g(a_n)$. Par unicité de la limite[90] dans Y, $f(x) = g(x)$. □

7.12.2 Espace métrisable

Définition 7.222 (Espace vectoriel topologique métrisable[194])**.**
*Un espace topologique est **métrisable** si il existe une distance compatible avec la topologie.*

Proposition 7.223 ([195])**.**
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

$$V_n = \{x \in X \text{ tel que } d(x, A) < \frac{1}{n}\}. \tag{7.204}$$

Et juste pour faire simple nous notons $V_0 = X$.

89. Proposition 7.217, valable parce que la topologie de X provient d'une métrique.
90. Proposition 7.95.

(i) **Les parties V_n sont ouvertes** Soit $x \in V_n$. Trouvons un voisinage de x contenu dans V_n afin de pouvoir encore invoquer le théorème 7.7. D'abord, vu que $x \in V_n$, il existe $a \in A$ tel que $d(x,a) < \frac{1}{n}$ (ici les inégalités strictes sont importantes).
Soient $\epsilon > 0$ que nous fixerons plus bas, et $y \in B(x,\epsilon)$. L'inégalité triangulaire donne

$$d(y,a) \leqslant d(y,x) + d(x,a) < \epsilon + \frac{1}{n}. \tag{7.205}$$

Nous pouvons donc choisir ϵ de telle sorte que $d(y,a) < 1/n$. Avec ce ϵ, nous avons, pour tout $y \in B(x,\epsilon)$:

$$d(y,A) \leqslant d(y,a) < \frac{1}{n} \tag{7.206}$$

et donc $y \in V_n$.

(ii) **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 $d(a,A) = 0$, et donc $a \in A$ par le lemme 7.129.

Ceci démontre le point (1).

En ce qui concerne la seconde partie, nous appliquons la première partie au complémentaire. Si $\mathcal{O}$ est ouvert, $\mathcal{O}^c$ est fermé et

$$\mathcal{O}^c = \bigcap_{n\in\mathbb{N}} V_n, \tag{7.207}$$

ce qui donne immédiatement

$$\mathcal{O} = \bigcup_{n\in\mathbb{N}} V_n^c \tag{7.208}$$

où les V_n^c sont fermés. □

Corolaire 7.224.
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 7.223. □

7.13 Suites de Cauchy, métrique et espaces complets

7.13.1 Généralités

Définition 7.225 (Suite de τ-Cauchy, espace vectoriel topologique[196, 184])**.**
*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 7.226 (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*[91] *vers un élément de A.*

Définition 7.227 (Suite de Cauchy, espace métrique)**.**
*Une suite (a_k) dans un espace métrique (V,d) est **de Cauchy** si pour tout $\epsilon \in \mathbb{R}$, il existe N tel que si $n,m \geqslant N$ alors $d(a_n,a_m) < \epsilon$.*

Notons qu'ici, même si l'espace V n'a rien à voir avec $\mathbb{R}$, nous prenons ϵ dans $\mathbb{R}$ et la distance à valeurs dans $\mathbb{R}$. Cela semble une évidence, mais il faut se rendre compte que $\mathbb{R}$ commence à prendre une place centrale dans nos constructions. Ce n'était pas le cas du temps où nous parlions de suites de Cauchy et de complétude dans des corps totalement ordonnés (définitions 1.318). Dans ce contexte, le ϵ était pris dans le corps lui-même.

91. Définition 7.12.

Définition 7.228 (Métrique complète)**.**
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.*

7.229.
Ces définitions méritent quelques 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 7.231 que le même espace topologique peut accepter plusieurs distances différentes (donnant la même topologie) donnant lieu à des suites de 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 à 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.

Définition 7.230 (espace de Banach, algèbre de Banach[197])**.**
Un ***espace de Banach*** *est un espace vectoriel normé complet*[92] *pour la topologie de la norme.*

Une ***algèbre de Banach*** *est une algèbre commutative et associative qui est un espace vectoriel normé complet.*

Exemple 7.231 (La complétude n'est pas une propriété topologique[198])**.**
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 $\mathbb{N}$:

$$d_1(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right|. \tag{7.209}$$

Pour vérifier que cette formule définit bien une distance (définition 7.97), le seul point non immédiat est l'inégalité triangulaire :

$$d_1(x, y) = \left|\frac{1}{x} - \frac{1}{y}\right| \leqslant \left|\frac{1}{x} - \frac{1}{z}\right| + \left|\frac{1}{z} - \frac{1}{y}\right| = d_1(x, z) + d_1(z, y). \tag{7.210}$$

Au niveau de la topologie induite par cette distance, c'est la topologie discrète. En effet, soit $x \in \mathbb{N}$ et $\epsilon > 0$; nous voulons déterminer la boule $B(x, \epsilon)$ en résolvant l'équation

$$\left|\frac{1}{x} - \frac{1}{y}\right| < \epsilon \tag{7.211}$$

pour $y \in \mathbb{N}$. Nous trouvons que $\frac{1}{y} > \frac{1}{x} - \epsilon$ et $\frac{1}{y} < \frac{1}{x} + \epsilon$, soit

$$\begin{cases} y > \dfrac{1}{\frac{1}{x} + \epsilon} & \text{(7.212a)} \\ y < \dfrac{1}{\frac{1}{x} - \epsilon}. & \text{(7.212b)} \end{cases}$$

Si ϵ est assez petit, la seule solution entière est $y = x$. Les ouverts sont donc toutes les parties parce que tous les singletons sont ouverts.

92. Définition 7.228.

L'espace topologique associé à $(\mathbb{N}, d_1)$ est donc la topologie discrète[93].

Si nous considérons par contre la distance usuelle sur $\mathbb{N}$, à savoir $d(x,y) = |x-y|$, nous obtenons encore la topologie discrète. Nous avons donc un isomorphisme d'espaces topologiques

$$(\mathbb{N}, d) \simeq (\mathbb{N}, d_1). \tag{7.213}$$

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 1.375) nous avons, pour $n, m > N$:

$$\Big|\frac{1}{n} - \frac{1}{m}\Big| < \frac{1}{n} < \frac{1}{N} < \epsilon. \tag{7.214}$$

Or cette suite ne converge pas. Soit en effet un candidat limite k. Calculons

$$d_1(x_n, k) = \Big|\frac{1}{n} - \frac{1}{k}\Big| \to \frac{1}{k} \neq 0. \tag{7.215}$$

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.

△

Proposition 7.232 ([199]).
Le dual[94] *d'un espace de Banach*[95] *est de Banach.*

7.13.2 Espace topologique métrique

Dans les espaces vectoriels topologiques métriques, il n'y a pas d'ambiguïté.

Proposition 7.233 (Caractérisations avec la distance d).
Soit (E, d) un espace vectoriel topologique métrique.

(1) Une suite (x_n) dans E est convergente[96] *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[97] *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 :

(i) **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 $d(x, x_k) \leqslant \epsilon$.

93. Celle dont toutes les parties sont des ouverts.
94. Définition 4.119.
95. Définition 7.230.
96. Définition 7.12.
97. Définition 7.225.

(ii) **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 7.98 et en particulier la liste des ouverts est donnée par (7.92). Il existe donc une boule $B(x, \epsilon)$ incluse 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 :

(i) **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 7.141(1)) nous avons

$$d(u, 0) = d(x_p - x_q, 0) = d(x_p, x_q). \tag{7.216}$$

Par conséquent $d(x_p, x_q) \leqslant \epsilon$.

(ii) **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

$$d(x_p, x_q) = d(x_p - x_q, 0), \tag{7.217}$$

et comme cela est plus petit que ϵ, nous avons $x_p - x_q \in B(0, \epsilon) \subset \mathcal{O}$.

□

Proposition 7.234 ([200]).
Toute suite convergente dans un espace métrique est de Cauchy.

Démonstration. Nous utilisons les caractérisations de la proposition 7.233 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 7.233(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

$$d(x_n, x_m) \leqslant d(x_m, \ell) + d(l, x_m) \leqslant 2\epsilon. \tag{7.218}$$

Cela prouve que (x_n) est une suite de Cauchy. □

7.13.3 Compacts, fermés

Proposition 7.235 (Séparation [184, 201]).
Soit V, un espace vectoriel topologique. Soient des parties K et F de V telles que :

(1) K est compact,

(2) F est fermé,

(3) $K \cap F = \varnothing$.

Alors il existe un voisinage U de 0 tel que

$$(K + U) \cap (F + U) = \varnothing. \tag{7.219}$$

Autrement dit, la topologie d'un espace vectoriel topologique sépare les fermés des compacts.

Démonstration. Soit un élément $x \in K$. La partie F^c est un ouvert qui contient x ; donc $F^c - x$ est un voisinage ouvert de 0.

(i) **Quelque chose sans intersection avec F** La proposition 7.164, il existe un ouvert symétrique U_x autour de 0 tel que

$$U_x + U_x + U_x + U_x \subset F^c + x. \tag{7.220}$$

En enlevant un des éléments de la somme, nous gardons la même inclusion :

$$U_x + U_x + U_x \subset F^c + x. \tag{7.221}$$

Et en passant le x de l'autre côté :

$$x + U_x + U_x + U_x \subset F^c. \tag{7.222}$$

Cela pour dire que nous avons, pour chaque $x \in K$ un voisinage ouvert U_x de 0 tel que

$$(x + U_x + U_x + U_x) \cap F = \varnothing. \tag{7.223}$$

(ii) $\underline{(x + U_x + U_x) \cap (U_x + F) = \varnothing}$ Supposons $z \in (x + U_x + U_x) \cap (U_x + F)$, et déduisons une contradiction. En particulier $z \in U_x + F$ et donc il existe $u \in U_x$, $f \in F$ tel que $z = v + f$. Nous avons alors (nous utilisons le fait que U_x soit symétrique)

$$f = z - v \in z + U_x \tag{7.224}$$

Mais z est aussi dans $x + U_x + U_x$. Donc

$$f \in z + U_x \subset x + U_x + U_x + U_x. \tag{7.225}$$

Cela prouve que $x + U_x + U_x + U_x \cap F \neq \varnothing$, en contradiction avec (7.223). Nous en concluons que

$$(x + U_x + U_x) \cap (U_x + F) = \varnothing. \tag{7.226}$$

(iii) **Un sous-recouvrement fini** Fini de rigoler. Si nous avons un compact dans les hypothèses, il fallait que ça arrive. Pour chaque $x \in K$ nous avons un voisinage ouvert U_x de 0 et donc un voisinage ouvert $x + U_x$ de x. Du coup $\{U_x + x\}_{x \in K}$ forme un recouvrement de K par des ouverts. Et hop, sous-recouvrement fini (c'est la définition 7.66 d'un compact) : nous avons $\{x_i\}_{i=1,\dots,n}$ tels que

$$K \subset \bigcup_{i=1}^{n} (x_i + U_{x_i}). \tag{7.227}$$

Nous posons $U = \bigcap_{i=1}^{n} U_{n_i}$. Même si c'est évident, remarquez que U est un ouvert autour de 0 et que U est plus petit que chacun des U_{x_i}. Récrivons (7.226) pour chacun des x_i :

$$(x + U_{x_i} + U_{x_i}) \cap (F + U_{x_i}) = \varnothing. \tag{7.228}$$

Lorsqu'on enlève des éléments dans une intersection vide, l'intersection reste vide. Donc en remplaçant des U_{x_i} par des U dans (7.228), nous conservons une intersection vide :

$$(x + U_{x_i} + U) \cap (F + U) = \varnothing. \tag{7.229}$$

(iv) **Conclusion** Aucun ensemble de la forme $x + U_{x_i} + U$ n'intersecte $F + U$. L'union ne l'intersecte pas non plus :

$$\bigcup_{i=1}^{n} (x + U_{x_i}) + U \cap (F + U) = \varnothing. \tag{7.230}$$

Et donc

$$(K + U) \cap (F + U) = \varnothing, \tag{7.231}$$

comme nous le voulions.

□

Définition 7.236.
Une distance d sur un espace vectoriel topologique V est dite ***compatible*** *avec la topologie si la topologie induite*[98] *de d est celle de V.*

98. Définition 7.98.

Définition 7.237.
Une distance d sur un espace vectoriel V est dite ***invariante*** *si pour tout $x, y, a \in V$ nous avons*

$$d(x+a, y+a) = d(x, y). \tag{7.232}$$

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.

Lemme 7.238 ([1])**.**
Soit un espace vectoriel muni d'une distance invariante. Alors

$$B(a, r) + x = B(a + x, r). \tag{7.233}$$

Démonstration. Un élément de $B(a, r) + x$ est de la forme $y + x$ avec $d(a, y) < r$. Nous avons alors $d(y + x, a + x) = d(y, a) < r$, de telle sorte que $y + x \in B(a + x, r)$.

Dans le sens inverse, si $y \in V$ vérifie $d(a + x, y) < r$, alors je prétend que $y - x \in B(a, r)$. En effet $d(y - x, a) = d(y - x + x, a + x) = d(y, a + x) < r$. □

<++>

Une version plus complète du lemme suivant sera dans la proposition 11.129 et ce qui suit.

Lemme 7.239.
Nous introduisons l'ensemble $\mathbb{D}_2$ des suites finies dans $\{0, 1\}$ [99].

L'application

$$\begin{aligned} \varphi\colon \mathbb{D}_2 &\to [0, 1[\\ x &\mapsto \sum_{i=1}^{l(x)} \frac{x_i}{2^i}. \end{aligned} \tag{7.234}$$

où $l(x)$ est la longueur de la suite finie x est injective.

Démonstration. Le lemme 1.417 donne la somme partielle de la série géométrique. Dans notre cas, $q = 1/2$. Si $x \in D_n$, alors en majorant chacun des c_i par 1,

$$x \leqslant \sum_{i=1}^{n} \left(\frac{1}{2}\right)^i = \frac{1 - (1/2)^{n+1}}{1/2} - 1 = 1 - (1/2)^{n+1} < 1. \tag{7.235}$$

Donc c'est bon pour dire que φ prend ses valeurs dans $[0, 1[$.

Il reste à voir l'injectivité. Nous supposons que $\varphi(x) = \varphi(y)$. Quitte à alonger x ou y par des zéros, nous supposons qu'elles ont même longueur N. Enfin nous définissons n_0 le plus petit indice pour lequel $x_i \neq y_i$. Nous avons :

$$0 = \frac{1}{2^{n_0}} + \sum_{i=n_0+1}^{N} \frac{x_i - y_i}{2^i}. \tag{7.236}$$

99. C'est en gros ce qui se fera dans 11.127 pour les développements de nombres dans une base donnée, sauf qu'ici nous n'avons pas besoin de subtilités sur les queues de suites.

Le deuxième terme est majoré de la façon suivante :

$$|\sum_{i=n_0+1}^{N} \frac{x_i - y_i}{2^i}| \leqslant \sum_{i=n_0+1}^{N} \frac{|x_i - y_i|}{2^i} \tag{7.237a}$$

$$\leqslant \sum_{i=n_0+1}^{N} \frac{1}{2^i} \tag{7.237b}$$

$$= \sum_{i=1}^{N} \left(\frac{1}{2}\right)^i - \sum_{i=1}^{n_0} \left(\frac{1}{2}\right)^i \tag{7.237c}$$

$$= \frac{1-(1/2)^{N+1}}{1/2} - \frac{1-(1/2)^{n_0+1}}{1/2} \tag{7.237d}$$

$$= \frac{1}{2^{n_0}} - \frac{1}{2^N} \tag{7.237e}$$

$$< \frac{1}{2^{n_0}}. \tag{7.237f}$$

L'égalité (7.236) n'est donc pas possible. Nous déduisons que n_0 n'existe en fait pas et que $x = y$. D'où l'injectivité de φ. □

Toujours avec l'application φ du lemme 7.239 nous avons ceci.

Lemme 7.240.
Soient $r, s \in \varphi(\mathbb{D}_2)$ tels que $r + s < 1$. Alors $r + s \in \varphi(\mathbb{D}_2)$.

Lemme 7.241.
Soient $r, s \in \varphi(\mathbb{D}_2)$ tels que $r + s < 1$. Nous posons $u = \varphi^{-1}(r)$, $v = \varphi^{-1}(s)$ et $w = \varphi^{-1}(r + s)$. Si k est le plus petit entier tel que $w_k \neq u_k + v_k$, alors $u_k = v_k = 0$ et $w_k = 1$.

Théorème 7.242 (Espace topologique métrisable[184, 1])**.**
Si V est un espace vectoriel topologique. Nous supposons :

(1) Tout point admet une base dénombrable de topologie[100]*.*

(2) Le singleton $\{0\}$ est fermé[101]*.*

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.

Démonstration. Vu que V admet une base dénombrable de topologie, nous en considérons une autour de 0 dans V : $\{A_i\}_{i\in\mathbb{N}}$.

(i) **Nouvelle base de topologie** Nous construisons une nouvelle base de topologie $\{U_i\}_{i\in\mathbb{N}}$ autour de 0 de la façon suivante. La proposition 7.164 donne un ouvert symétrique équilibré dans A_0. Nous le nommons U_0 :

$$U_0 \subset A_0. \tag{7.238}$$

Pour les suivants, U_{k+1} est un ovuert symétrique équilibré dans $U_k \cap A_k$ vérifiant

$$U_{k+1} + U_{k+1} + U_{k+1} + U_{k+1} \subset A_k \cap U_k. \tag{7.239}$$

Le fait est que $A_k \cap U_k$ est un ouvert autour de 0 ; donc la proposition 7.164 permet de construire un ouvert symétrique et équilibré U_{k+1} autour de 0 ayant la propriété demandée.

(ii) **C'est bien une base** Notons d'abord que $U_k \subset A_k$. Si $\mathcal{O}$ est un ouvert autour de 0, il existe k tel que $A_k \subset \mathcal{O}$: c'est la définition 7.2 d'une base de topologie. Nous avons donc $U_k \subset A_k \subset \mathcal{O}$, et les $\{U_k\}$ forment une base dénombrable de la topologie autour de 0.

100. Définition 7.2.

101. Attention : dans [184], cela fait partie de la définition d'un espace vectoriel topologique et n'est donc pas listé dans les hypothèses de ce théorème.

(iii) **Quelques inclusions** Notons au passage quelques inclusions. Pour tout $n \in \mathbb{N}$ nous avons

$$U_{n+1} + U_{n+1} + U_{n+1} + U_{n+1} \subset U_n. \tag{7.240}$$

À fortiori nous avons

$$U_{n+1} \subset U_n \tag{7.241}$$

et pour tout $n, k \in \mathbb{N}$, nous obtenons alors

$$U_{n+1} + U_{n+2} + \cdots + U_{n+(k-1)} + U_{n+k} \subset U_{n+1} + U_{n+1} \subset U_n. \tag{7.242}$$

(iv) **L'ensemble D** Nous considérons l'ensemble $\mathbb{D}_2$ et l'application φ du lemme 7.239. Nous notons $D = \varphi(\mathbb{D}_2) \subset [0, 1[$. L'application $\varphi\colon \mathbb{D}_2 \to D$ est injective, et nous pouvons donc parler de $\varphi^{-1}(r)$ pour tout $r \in D$.

(v) **La fonction ϕ** Nous définissons maintenant une fonction à valeurs dans la topologie τ_V de V :

$$\begin{aligned} \phi\colon D \cup [1, +\infty[&\to \tau_V \\ r &\mapsto \begin{cases} V & \text{si } r \geqslant 1; \\ \sum_i \varphi^{-1}(r)_i U_i & \text{si } r \in D. \end{cases} \end{aligned} \tag{7.243}$$

La dernière somme est toujours une somme finie et est un ouvert parce que la multiplication par un scalaire et l'addition sont des ouverts.

Notez aussi que plus r est petit, plus ce sont les grands i qui tendront à avoir $\varphi^{-1}(r) \neq 0$. En effet, pour $r = 1/8$, nous avons $\varphi^{-1}(1/8) = (0, 0, 0, 1)$, et plus généralement pour $r < 1/2^N$, les N premiers termes de $\varphi^{-1}(r)$ seront nuls.

Quelques remarques sur cette fonction.

(i) **Une égalité facile** Si je ne me trompe pas d'un off-by-one, nous avons

$$\phi(1/2^N) = U_N. \tag{7.244}$$

(ii) **$0 \in \phi(r)$ pour tout r** En effet, $\phi(r)$ n'est jamais vide, c'est toujours un voisinage de 0.

(iii) **$\phi(r) + \phi(s) \subset \phi(r + s)$** Si $r + s \geqslant 1$, c'est clair. Sinon, nous posons $u = \varphi^{-1}(r)$, $v = \varphi^{-1}(s)$ et $w = \varphi^{-1}(r + s)$. Dans la suite, nous prolongeons u et v pour qu'elles aient le même nombre d'éléments que nous notons N.

Deux cas se produisent : soit $w_i = u_i + v_i$ pour tout i soit non.

(i) **Premier cas** Si $w_i = u_i + v_i$ pour tout i, alors en particulier les u_i et v_i ne peuvent pas être 1 en même temps et nous pouvons séparer clairement les termes de la somme :

$$\phi(r + s) = \sum_i w_i U_i = \sum_i u_i U_i + \sum_i v_i U_i = \phi(r) + \phi(s); \tag{7.245}$$

(ii) **Second cas** Sinon, posons k le plus petit entier tel que $w_k \neq u_k + v_k$. Alors le lemme 7.241 dit que $u_k = v_k = 0$ et $w_k = 1$. Nous avons d'abord

$$\phi(r) = \sum_{i=1}^{k-1} u_i U_i + \sum_{i=k+1}^{n} u_i U_i \subset \sum_{i=1}^{k-1} u_i U_i + U_{k+1} + U_{k+1} \tag{7.246}$$

où nous avons utilisé (7.241) pour tous les derniers termes. De même nous avons :

$$\phi(s) = \sum_{i=1}^{k-1} v_i U_i + \sum_{i=k+1}^{n} v_i U_i \subset \sum_{i=1}^{k-1} v_i U_i + U_{k+1} + U_{k+1}. \tag{7.247}$$

En combinant, et en utilisant (7.240),

$$\phi(r)+\phi(s)=\sum_{i=1}^{k-1}u_iU_i+\sum_{i=1}^{k-1}v_iU_i+\underbrace{U_{k+1}+U_{k+1}+U_{k+1}+U_{k+1}}_{\subset U_k} \tag{7.248a}$$

$$\subset\sum_{i=1}^{k-1}w_iU_i+U_k \tag{7.248b}$$

$$=\sum_{i=1}^{k}w_iU_i \tag{7.248c}$$

$$\subset\phi(r+s). \tag{7.248d}$$

Pour (7.248c) nous avons utilisé le fait que $w_k=1$. Au final nous avons bien $\phi(r)+\phi(s)\subset\phi(r+s)$.

(iv) **ϕ est croissante** Dans notre contexte « croissante » signifie que si $r<s$ alors $\phi(r)\subset\phi(s)$. Il suffit d'écrire

$$\phi(r)\subset\phi(r)+\phi(s-r)\subset\phi(s). \tag{7.249}$$

(iv) **Définition de la distance (enfin !)** Nous définissons l'application

$$\begin{aligned}f\colon V&\to\mathbb{R}\\ x&\mapsto\inf\{r\text{ tel que }x\in\phi(r)\},\end{aligned} \tag{7.250}$$

et ensuite ce qui va être notre distance :

$$d(x,y)=f(y-x). \tag{7.251}$$

Notons que cet infimum est un réel, et qu'il n'est pas spécialement dans $\varphi(\mathbb{D}_2)$. Nous devons prouver que ce d est une distance invariante par translation et compatible avec la topologie.

(v) **d est invariante par translation** Il s'agit seulement d'écrire $d(x+a,y+a)$ et de remarquer que $(y+a)-(x+a)=y-x$.

(v) $d(x,x)=0$ Oui, car 0 est dans $\phi(r)$, pour tout r, puisque les U_i sont des voisinages de 0.

(vi) **$d(x,y)=0$ implique $x=y$** Nous montrons que $f(x)\neq 0$ dès que $x\neq 0$. Rappellez-vous que nous avons posé l'hypothèse que $\{0\}$ est un fermé dans V. Vu que $\{x\}$ est un compact [102], la proposition 7.235 nous indique qu'il existe un voisinage A de 0 qui ne contient pas x.

Vu que les $\{U_k\}_{k\in\mathbb{N}}$ forment une base de la topologie autour de 0, il existe un n_0 tel que $x\notin U_{n_0}$. Et comme $U_{k+1}\subset U_k$ nous avons

$$x\notin U_k \tag{7.252}$$

pour tout $k\geqslant n_0$. Nous savons aussi que ϕ est croissante et que $\phi(1/2^N)=U_N$. Donc pour tout $\epsilon<\frac{1}{2^k}$, nous avons

$$\phi(\epsilon)\subset U_k \tag{7.253}$$

et donc $x\notin\phi(\epsilon)$. Donc

$$f(x)\geqslant\frac{1}{2^k}>0 \tag{7.254}$$

la dernière inégalité est stricte et c'est important.

En ce qui concerne la distance, si $d(x,y)=0$, alors $f(y-x)=0$. Cela signifie que $y-x=0$ et donc que $x=y$. Ok.

(vii) $d(x,y)=d(y,x)$ Oui, car tous les voisinages considérés sont symétriques. Donc $x-y\in\phi(r)$ si et seulement si $y-x\in\phi(r)$.

102. Les singletons sont toujours des compacts dans les espaces topologiques.

(viii) $d(x,z) \leqslant d(x,y) + d(y,z)$ Nous supposons que x, y et z sont trois points distincts, de telle sorte que les trois distances soient strictement positives.

Soit $\epsilon > 0$. Par définition des distances comme infimums, et grâce au corolaire 1.379, il existe r et s dans $\varphi(\mathbb{D}_2)$ tels que :

$$d(x,y) < r < d(x,y) + \frac{\epsilon}{2} \quad \text{et} \quad d(y,z) < s < d(y,z) + \frac{\epsilon}{2}. \tag{7.255}$$

En sommant :

$$d(x,y) + d(y,z) < r + s < d(x,y) + d(y,z) + \epsilon. \tag{7.256}$$

Vu que ϕ est croissante, on a $y - x \in \phi(r)$ et $z - y \in \phi(s)$; donc

$$z - x = (y - x) + (z - y) \in \phi(r) + \phi(s) \subset \phi(r + s) \tag{7.257}$$

Ainsi, pour tout $\epsilon > 0$, on a

$$d(x,z) \leqslant r + s < d(x,y) + d(y,z) + \epsilon. \tag{7.258}$$

(ix) **Compatibilité avec la topologie** Nous devons montrer que les d-ouverts sont les mêmes que les ouverts de la topologie de V. Nous allons utiliser la proposition 7.2.

(i) **Les ouverts sont des d-ouverts** Soit un ouvert $\mathcal{O}$ contenant $x \in V$. Alors $\mathcal{O} - x$ est un ouvert contenant 0, et il existe un n tel que

$$U_n \subset \mathcal{O} - x. \tag{7.259}$$

Nous avons alors

$$B(0, \frac{1}{2^n}) \subset U_n \subset \mathcal{O} - x. \tag{7.260}$$

Nous utilisant le lemme 7.238 pour additionner x dans toutes les inclusions :

$$B(x, \frac{1}{2^n}) = B(0, \frac{1}{2^n}) + x \subset \mathcal{O}. \tag{7.261}$$

Les boules forment donc une base de la topologie de V. Donc tous les ouverts de V sont des unions de boules et sont donc des d-ouverts.

(ii) **Les d-ouverts sont des ouverts** Considérons une boule $B(0,r)$. Il existe un n tel que $B(0, 1/2^n) \subset B(0,r)$ et donc tel que

$$U_n \subset B(0,r). \tag{7.262}$$

En procédant par translations et tout ça, on en déduit que les $\{U_n + x\}_{n\in\mathbb{N}, x\in V}$ forment une base de la d-topologie. Donc les d-ouverts sont des ouverts.

□

7.13.4 Équivalence entre Cauchy et $\tau-$Cauchy

Lemme 7.243.
Soit un espace vectoriel topologique[103] *V et une distance $d\colon V \times V \to \mathbb{R}^+$ compatible*[104] *avec la topologie de V. Si d est invariante*[105]*, alors les suites de Cauchy pour d et les suites τ-Cauchy sont les mêmes.*

Démonstration. Nous avons deux implications à prouver.

103. Définition 7.148.
104. Définition 7.236.
105. Définition 7.237.

(i) **Cauchy pour d implique τ-Cauchy** Soit (x_n), une suite de Cauchy dans V pour d, et un voisinage U de 0. Vu que d est compatible avec la topologie de V, il existe une boule ouverte $B(0,\epsilon)$ incluse 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

$$d(0, x_m - x_n) < \epsilon, \tag{7.263}$$

c'est-à-dire $x_m - x_n \in B(0,\epsilon) \subset U$. La suite (x_n) est donc τ-Cauchy.

(ii) **τ-Cauchy implique Cauchy pour d** 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 $d(x_n, x_m) < \epsilon$.

□

Tout ceci nous mène à donner une large classe d'espaces vectoriels topologiques sur lesquelles les notions de suites de Cauchy pour une distance et τ-Cauchy coïncident.

Théorème-Définition 7.244.
Soit V un espace vectoriel topologique métrisable[106]*, alors il admet une métrique d compatible avec la topologie telle que une suite dans V est de Cauchy pour d si et seulement si elle est τ-Cauchy.*

*Une **suite de Cauchy** dans un espace vectoriel métrique (E, d) est une suite τ-Cauchy ou de Cauchy pour d.*

Démonstration. Soit d une métrique sur V satisfaisant au théorème 7.242. Vu qu'elle est invariante par translation, les suites d-Cauchy sont exactement les suites τ-Cauchy par le lemme 7.243. □

Remarque 7.245.
Même si V est métrisable, si on choisit la métrique n'importe comment, on ne peut rien espérer.

7.246.
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.

Le théorème suivant donne la complétude de $\mathbb{R}$ et le critère de Cauchy pour les définitions métriques et topologiques usuelles. Lorsqu'on dit que $\mathbb{R}$ est complet, le plus souvent nous parlons de ce théorème, et non de 1.389 qui en est un lemme indispensable mais qui parle de notions différentes, bien que très liées.

Théorème 7.247 (Complétude de $\mathbb{R}$, critère de Cauchy[15])**.**
Nous avons :

(1) L'espace métrique $(\mathbb{R}, d)$ est complet (définition 7.228).

(2) Une suite dans $\mathbb{R}$ est convergente (définition 7.12) si et seulement si elle est de Cauchy (définition 7.244).

Démonstration. Tout ce théorème se base sur le fait que la définition de suite de Cauchy dans $(\mathbb{R}, d)$ et de suite convergente dans $(\mathbb{R}, d)$ coïncident avec les définitions correspondantes dans $\mathbb{R}$ vu comme simple corps ordonné (définitions 1.318).

Donc si (x_n) est de Cauchy dans $(\mathbb{R}, d)$, elle est de Cauchy dans le corps ordonné $(\mathbb{R}, \leqslant)$. Donc le théorème 1.389 nous dit que (x_n) est convergente dans $(\mathbb{R}, \leqslant)$. Et donc convergente dans $(\mathbb{R}, d)$.

Toutes les autres affirmations se prouvent de la même manière. □

106. Voir la proposition 7.242 pour une condition suffisante.

Si vous n'êtes pas sûr ou si vous ne voulez pas étudier les notations de convergence et de suites de Cauchy dans les corps, vous pouvez simplement recopier la démonstration du théorème 1.389 en remplaçant partout $\mathbb{Q}$ par $\mathbb{R}$, et aussi en remplaçant les $|x-y|$ par $d(x,y)$.

7.248.
Nous pouvons également mettre une structure d'espace métrique sur $\mathbb{C}$ en posant

$$d(z,z') = |z-z'|. \tag{7.264}$$

Proposition 7.249.
L'espace métrique $(\mathbb{C},d)$ est complet.

Démonstration. Commençons par nous rendre compte que pour tout $z \in \mathbb{C}$ nous avons $|\operatorname{Re}(z)| \leqslant |z|$. C'est bon ? Vous vous en êtes rendu compte ? Ok. Continuons.

Soit une suite de Cauchy (z_k) dans $\mathbb{C}$ et $\epsilon > 0$. Si $x_k = \operatorname{Re}(z_j)$, nous avons

$$|x_k - x_l| = |\operatorname{Re}(z_k - z_l)| \leqslant |z_k - z_l|. \tag{7.265}$$

Vu que (z_k) est de Cauchy, il existe un N tel que si $k,l \geqslant N$,

$$|x_k - x_l| \leqslant |z_k - z_l| \leqslant \epsilon. \tag{7.266}$$

Donc la suite des parties réelles converge par la complétude de $(\mathbb{R},d)$ du théorème 7.247. Notez que le d ici n'est pas tout à fait le même, et que la démonstration fonctionne parce que la distance prise sur $\mathbb{R}$ est la restriction à $\mathbb{R}$ de la distance prise sur $\mathbb{C}$. Notons x la limite de (x_k).

De la même manière la suite des parties imaginaires $y_k = \operatorname{Im}(z_k)$ converge vers un réel que nous notons y. Avec tout cela, la suite z_k converge dans $\mathbb{C}$ vers $x+iy$. En effet pour ϵ donné et pour un k suffisament grand,

$$|z_k - (x+iy)| = \big|\operatorname{Re}(z_k) - x + i(\operatorname{Im}(z_k) - y)\big| \leqslant |x_k - x| + |y_k - y| \leqslant \epsilon. \tag{7.267}$$

□

7.14 Norme, espace vectoriel normé

La valeur absolue est essentielle pour introduire les notions de limite et de continuité pour les fonctions d'une variable. Par exemple nous verrons dans la proposition 10.73 que la fonction $f\colon \mathbb{R} \to \mathbb{R}$ est continue en a si et seulement si pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$|x-a| \leqslant \delta \Rightarrow |f(x) - f(a)| \leqslant \varepsilon. \tag{7.268}$$

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 éloignés 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 notions 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 première 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 :

$$\text{taille de } (a,b) = \sqrt{a^2+b^2}. \tag{7.269}$$

Nous pouvons introduire une notion de distance entre les éléments de $\mathbb{R}^2$ de façon similaire :

$$d\big((a_x,a_y),(b_x,b_y)\big) = \sqrt{(a_x-b_x)^2+(a_y-b_y)^2}. \tag{7.270}$$

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 définition possibles de « taille » et distance en $\mathbb{R}^2$?

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.

7.14.0.1 Critère de Cauchy

Théorème 7.250 (Bolzano-Weierstrass, thème 31)**.**
Toute suite contenue dans un compact admet une sous-suite convergente.

Démonstration. Nous faisons la preuve par l'absurde en supposant que (x_k) n'admette pas de sous-suite convergente. Soit $a \in K$; aucune sous-suite de (x_k) ne converge vers a. En particulier, il existe un voisinage ouvert $\mathcal{O}_a$ de a et une partie finie I_a de $\mathbb{N}$ tel que $x_k \in \mathcal{O}_a$ seulement pour $k \in I_a$.

Les ouverts $\mathcal{O}_a$ recouvrent K ; nous pouvons en extraire un sous-recouvrement fini (c'est la définition 7.66 de la compactié). Nous avons donc des points $a_1,\dots,a_n$ tels que

$$K \subset \bigcup_{i=1}^{n} \mathcal{O}_{a_i} \tag{7.271}$$

et tels que pour chaque $\mathcal{O}_{a_i}$, nous avons $x_k \in \mathcal{O}_{a_i}$ seulement pour $k \in I_{a_i}$. Bien entendu, toute la suite est dans K et donc dans l'union.

En conclusion, nous avons $\mathbb{N} = \bigcup_{i=1}^{n} I_{a_i}$, ce qui prouve que $\mathbb{N}$ est un ensemble fini. Contradiction avec la proposition 1.115 qui dit que $\mathbb{N}$ est infini. □

Corolaire 7.251 ([1])**.**
Une suite dans un compact dont toutes les sous-suites convergentes convergent vers la même limite est convergente.

Démonstration. Soient un espace topologique X, un compact K et une suite (x_k) dans K. Nous supposons que toutes les sous-suites convergentes de (x_k) convergent vers $a \in X$. Nous devons montrer que $x_k \xrightarrow{X} a$.

Supposons que ce ne soit pas le cas. Alors il existe un voisinage V de a tel que pour tout $N \in \mathbb{N}$, il existe k tel que $x_k \notin V$. Cela produit une sous-suite hors de V. Cette sous-suite est encore dans K et possède donc une sous(-sous)-suite convergente (théorème 7.250). Par hypothèse, cette sous-sous-suite doit converger vers a, ce qui est impossible.

Contradiction et le corolaire est prouvé. □

Lemme 7.252.
Une suite de Cauchy[107] *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

107. Définition 7.227.

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

$$\|a_k - \ell\| \leqslant \|a_k - a_n\| + \|a_n - \ell\| \leqslant 2\epsilon. \tag{7.272}$$

□

Dans le cas des espaces de dimension finie, le fait d'être complet est automatique, comme le montre la proposition suivante.

Proposition 7.253.
Soit $\big(E, \|.\|\big)$ un espace vectoriel normé de dimension finie sur un corps $\mathbb{K}$ qui est complet[108]*. Alors E est complet*[109]*.*

Pour rappel, la complétude de l'espace métrique $\mathbb{R}$ est la proposition 1.341.

Démonstration. Nous considérons une suite de Cauchy (f_n) dans E et si $\{e_\alpha\}$ est une base orthonormée de E nous définissons les coefficients $f_n = \sum_\alpha a_{n\alpha} e_\alpha$. La somme sur α est finie par hypothèse sur la dimension de E.

Nous avons

$$\|f_n - f_m\| = \|\sum_\alpha (a_{n\alpha} - a_{m\alpha}) e_\alpha\| = \sum_\alpha |a_{n\alpha} - a_{m\alpha}|^2. \tag{7.273}$$

Pour tout ϵ, il existe N tel que si $m, n > N$ alors $|a_{n\alpha} - a_{m\alpha}| < \sqrt{\epsilon}$. Autrement dit, pour chaque α, la suite $(a_{n\alpha})_{\alpha \in \mathbb{N}}$ est de Cauchy dans $\mathbb{K}$ et converge donc dans $\mathbb{K}$. Soit a_α la limite et définissons $f = \sum_\alpha a_\alpha e_\alpha$. Nous avons alors

$$\|f_n - f\| = \|\sum_\alpha (a_{n\alpha} - a_\alpha) e_\alpha\|, \tag{7.274}$$

dont la limite $n \to \infty$ est bien zéro. Donc la suite (f_n) converge vers $f \in E$. L'espace E est alors complet. □

Proposition 7.254.
Soient V et W deux espaces vectoriels normés. Soient 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 7.186.

Démonstration. Nous allons prouver que $f(K)$ est fermée et bornée.

(i) **$f(K)$ 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 7.250). 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

$$f(a) = \lim f(x'_n) = \lim y_n. \tag{7.275}$$

Cela prouve que la limite de (y_n) est dans $f(K)$ et par conséquent que $f(K)$ est fermé.

108. La définition est 1.318, mais si vous n'avez pas envie de vous embarquer trop loin, dites juste « toutes les suites de Cauchy convergent ». Typiquement c'est $\mathbb{R}$ ou $\mathbb{C}$.
109. Définition 7.228.

(ii) **$f(K)$ est borné** Si $f(K)$ n'est pas borné, nous pouvons trouver une suite (x_n) dans K telle que

$$\|f(x_n)\|_W > n \tag{7.276}$$

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

$$|f(a)| = \lim |f(x'_n)|. \tag{7.277}$$

La suite $f(x'_n)$ est alors une suite bornée, ce qui n'est pas possible au vu de la condition (7.276) imposée à la suite de départ (x_n).

□

Corolaire 7.255.
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 7.254 montre que $f(K)$ est compact et donc borné. □

7.15 Espaces métriques

7.15.1 Espaces métrisables

Définition 7.256.
*Un espace topologique est **métrisable** si il est homéomorphe à un espace métrique.*

Proposition 7.257.
Une fonction séquentiellement continue sur un espace métrisable et à valeurs dans un espace métrique est continue.

Démonstration. Soient E un espace métrique et $\phi\colon X \to (E, d)$ un homéomorphisme. 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

$$\begin{aligned} \tilde{f}\colon E &\to Y \\ a &\mapsto f\big(\phi^{-1}(a)\big). \end{aligned} \tag{7.278}$$

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

$$\tilde{f}(a_k) = f\big(\phi^{-1}(a_k)\big), \tag{7.279}$$

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

$$\lim_{k\to\infty} \tilde{f}(a_k) = \lim_{k\to\infty} f\big(\phi^{-1}(a_k)\big) = f\big(\phi^{-1}(a)\big) = \tilde{f}(a). \tag{7.280}$$

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 7.220. L'application $f = \tilde{f} \circ \phi$ est donc continue en tant que composée d'applications continues. □

7.15.2 Fonctions continues

La propriété suivante donne des caractérisations importantes de la continuité dans le cas des espaces métriques.

Proposition 7.258 (Continuité, ouverts et voisinages et limite[202]).
Soient $f\colon E \to F$ 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 7.92,

(6) Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $\|x - a\| < \delta$ implique $\|f(x) - f(a)\| < \epsilon$.

La proposition 7.292 nous montrera que ces équivalences tiennent encore lorsque l'espace a une topologie de seminormes.

Démonstration. L'équivalence (1) $\Leftrightarrow$ (2) est la définition 7.41. L'équivalence (3) $\Leftrightarrow$ (4) est une simple paraphrase.

Montrons (2) $\Rightarrow$ (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 [110], il en contient un autour de a : $V' = B(a, \delta) \subset V$. A fortiori nous avons $f(V') \subset W$.

Montrons (3) $\Rightarrow$ (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) $\Leftrightarrow$ (5) est la définition 7.41 de la continuité en un point couplée à l'unicité de la limite due à la proposition 7.95 parce qu'un espace métrique est séparé.

Prouvons (5) $\Rightarrow$ (6). Soient $\epsilon > 0$ 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) $\Rightarrow$ (5) est une réécriture de la définition de la limite en un point. □

Voici un théorème qui parle de fermés emboîtés dans un espace métrique. Le corolaire 7.79 parle du cas $\cap_i A_i = \emptyset$ dans un compact.

Théorème 7.259 (Théorème des fermés emboîtés[203]).
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. En deux parties.

(i) **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.

(ii) **Condition nécessaire** Soit (x_n) un suite de Cauchy. Nous considérons les ensembles

$$F_n = \overline{\{x_i \text{ tel que } i \geqslant n\}}. \tag{7.281}$$

Le fait que la suite soit de Cauchy implique que $\operatorname{diam}(F_n) \to 0$. Par hypothèse, nous avons alors

$$\bigcap_{n\in\mathbb{N}} F_n = \{a\}. \tag{7.282}$$

110. Cela est le théorème-définition 7.98 des ouverts dans un espace métrique, à ne pas confondre avec le théorème 7.7.

Pour s'assurer que a est bien la limite de (x_n), il suffit de remarquer que

$$d(x_n, a) \leqslant \operatorname{diam} F_n \to 0. \tag{7.283}$$

□

Proposition 7.260.
Soient (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 7.233(1).

Par la caractérisation séquentielle des fermés (un fermé contient les limites de toutes ses suites, proposition 7.217), la suite (x_k) étant dans F, la limite est dans F. Donc $x \in F$. □

Lemme 7.261.
Soit A_n une suite décroissante de fermés dans un espace métrique[111] *compact K. Alors*

$$C = \bigcap_{n\in\mathbb{N}} A_n \tag{7.284}$$

est non vide.

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 7.82), 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 7.124. Une queue de la suite y_n est dans A_2 et nous considérons donc une sous-suite convergente dans A_2 donnée par

$$z_n = y_{\sigma_2(n)} = x_{\sigma_1\sigma_2(n)}. \tag{7.285}$$

En continuant ainsi nous construisons une suite convergente dans A_k. Nous considérons enfin la suite

$$y_n = x_{\sigma_1\ldots\sigma_n(n)}. \tag{7.286}$$

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 7.262.
Cette propriété est fausse pour les ouverts. Par exemple

$$\bigcap_{n>1}]0, \frac{1}{n}[= \varnothing. \tag{7.287}$$

Lemme 7.263.
Si K est un compact dans un espace métrique et F un fermé disjoint de K, alors $d(K, F) > 0$.

Démonstration. La fonction

$$\begin{aligned} K &\to \mathbb{R} \\ x &\mapsto d(x, F) \end{aligned} \tag{7.288}$$

est une fonction continue sur K, et donc atteint son minimum par le théorème de Weierstrass 7.126. 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. □

111. L'hypothèse métrique provient de l'utilisation de Bolzano-Weierstrass, lequel est vrai pour les espaces séquentiellement compacts, dont les espaces métriques.

Proposition 7.264 ([180]).
Une isométrie d'un espace métrique compact sur lui-même est une bijection.

Démonstration. Soient 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 7.263 appliqué au fermé $\{x\}$ et au compact $f(K)$ donne un $r > 0$ tel que

$$d\big(x, f(K)\big) > r. \tag{7.289}$$

Soit la suite $u_n = f^n(x)$; c'est une suite dans K et possède donc une sous-suite convergente (Bolzano-Weierstrass7.124) que l'on nomme (y_n). Vu que f est une isométrie,

$$d(y_n, y_{n+1}) = d(x, y_m) > r \tag{7.290}$$

pour un certain $m \leqslant n+1$. Cela 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 7.265.
Soient (X, d) un espace métrique compact et (u_n) une suite de X telle que

$$\lim_{n\to\infty} d(u_n, u_{n+1}) = 0. \tag{7.291}$$

Alors l'ensemble des points d'accumulation[112] *de (u_n) est connexe.*

Démonstration. Nous notons Γ l'ensemble des points d'accumulation de la suite.

(i) **Γ est compact** Nous notons $A_p = \{u_n \text{ tel que } n \geqslant p\}$ et nous avons

$$\Gamma = \bigcap_{p\in\mathbb{N}} \overline{A_p} \tag{7.292}$$

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 7.6). En tant que fermé dans un compact, Γ est compact (lemme 7.82).

(ii) **Recouvrement par deux compacts** Supposons que Γ ne soit[113] pas connexe. Nous pouvons alors considérer S et O, deux ouverts disjoints recouvrant Γ et intersectant tous deux Γ. Nous posons alors

$$A = S \cap \Gamma \tag{7.293a}$$
$$B = O \cap \Gamma, \tag{7.293b}$$

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[114].

112. Définition 7.39.
113. est-ce qu'il faut vraiment un subjonctif ici ?
114. Lemme 7.82.

(iii) **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 7.263. Nous notons

$$A'=\{x\in X \text{ tel que } d(x,A)<\frac{\alpha}{3}\} \tag{7.294a}$$

$$B'=\{x\in X \text{ tel que } d(x,B)<\frac{\alpha}{3}\} \tag{7.294b}$$

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

$$K=X\backslash(A'\cup B') \tag{7.295}$$

qui est fermé en tant que complémentaire d'ouvert, et donc compact. Étant donné que $A\subset A'$ et $B\subset B'$, nous avons $K\cap\Gamma=\emptyset$.

L'idée est maintenant de montrer que K contient un point d'accumulation de (u_n).

(iv) **Sous-suites de (u_n)** L'hypothèse sur la suite (u_n) nous indique qu'il existe un N_0 tel que $\forall n\geqslant N_0$,

$$d(u_n,u_{n+1})<\frac{\alpha}{3}. \tag{7.296}$$

Soient $N>N_0$ 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 $B'\cap A'=\emptyset$, 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 (7.296) est important. Nous avons donc $N_0<n_1<n_0<n_2$.

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

$$\begin{aligned}d(u_{n_0},B)&\geqslant d(u_{n_0-1},B)-d(u_{n_0-1},u_{n0})\\&\geqslant d(A,B)-d(u_{n_0-1},A)-d(u_{n_0-1},u_{n_0})\\&\geqslant\alpha-\frac{\alpha}{3}-\frac{\alpha}{3}\\&=\frac{\alpha}{3}.\end{aligned} \tag{7.297}$$

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 7.266.
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 7.267.
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) $K_n \subset \mathrm{Int}(K_{n+1})$.

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 dans $\mathrm{Int}(K_n)$ pour un certain n.

*Une telle suite de compacts est une **suite exhaustive** de compacts pour Ω.*

Démonstration. Nous considérons les ensembles

$$V_n = \{z \in E \text{ tel que } |z|\} \cup \bigcup_{a \notin \Omega} B(a, \frac{1}{n}), \tag{7.298}$$

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
$$\frac{1}{n+1} < \delta < \frac{1}{n} \tag{7.299}$$
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 $K_n \subset \mathrm{Int}(K_{n+1})$, 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 $\Omega = \bigcup_{n=0}^{\infty} \mathrm{Int}(K_n)$. En effet nous avons
$$\Omega = \bigcup_{n=0}^{\infty} K_n \subset \bigcup_{n=0}^{\infty} \mathrm{Int}(K_{n+1}) \subset \bigcup_{n=0}^{\infty} \mathrm{Int}(K_n). \tag{7.300}$$
L'inclusion dans l'autre sens est facile.
Soit K compact dans Ω. Vu que Ω est l'union des $\mathrm{Int}(K_n)$, nous avons
$$K \subset \bigcup_{n=0}^{\infty} \mathrm{Int}(K_n). \tag{7.301}$$
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 $K \subset \mathrm{Int}(K_m)$.

□

Notons qu'avec la suite de K_n telle que construite, le dernier point est réglé en prenant

$$\frac{1}{n+1} < \delta_n < \frac{1}{n}. \tag{7.302}$$

Lemme 7.268 ([204]).
Soient un compact $K \subset \mathbb{R}^d$ ainsi que des ouverts $\{\Omega_i\}_{i=1,\dots,n}$ tels que $K \subset \bigcup_{i=1}^n \Omega_i$.

Il existe des compacts $\{K_i\}_{i=1,\dots,n}$ tels que

— $K_i \subset \Omega_i$,

— $K \subset \bigcup_{i=1}^n K_i$.

Démonstration. Soit $x \in K$. Vu que les Ω_i recouvrent K, il existe un $k(x) \in \{1,\dots,n\}$ tel que $x \in \Omega_{k(x)}$. De plus, vu que $\Omega_{k(x)}$ est ouvert, il existe un voisinage de x contenu dans $\Omega_{k(x)}$ (théorème 7.7). Autrement dit, il existe $r(x) > 0$ tel que

$$\overline{B\big(x, r(x)\big)} \subset \Omega_{k(x)}. \tag{7.303}$$

Vu que l'ensemble $\{B\big(x, r(x)\big)\}_{x\in k}$ est un recouvrement de K par des ouverts, nous pouvons en extraire un sous-recouvrent fini [115]. Soient donc $x_1,\dots,x_m \in K$ tels que

$$K \subset \bigcup_{i=1}^m B\big(x_i, r(x_i)\big). \tag{7.304}$$

Pour chaque $j = 1,\dots,n$, nous posons

$$A_j = \big\{l \in \{1,\dots,m\} \text{ tel que } k(x_l) = j\big\}. \tag{7.305}$$

Et enfin nous définisssons, pour $j = 1,\dots,n$ les parties

$$K_j = \bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)} \tag{7.306}$$

et il nous reste à prouver que ces ensembles répondent bien à la question.

(i) $\underline{\bigcup_{j=1}^n A_j = \{1,\dots,m\}}$ **est une union disjointe** Un élément l de $A_i \cap A_j$ devrait vérifier $i = k(x_l) = j$. Si $s \in \{1,\dots,m\}$, alors $s \in A_{k(x_s)}$. Donc oui, l'union des A_j est tout $\{1,\dots,m\}$.

(ii) $\underline{K_j \subset \Omega_j}$ Nous avons

$$K_j = \bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)} \subset \bigcup_{l\in A_j} \Omega_{k(x_l)} = \bigcup_{l\in A_l} \Omega_j = \Omega_j. \tag{7.307}$$

(iii) $\underline{K \subset \bigcup_{i=1}^n K_i}$. Par (7.304), et vu que $\{1,\dots,m\} = \bigcup_{j=1}^n A_j$,

$$\begin{aligned}
K &\subset \bigcup_{i=1}^m B\big(x_i, r(x_i)\big) && \text{(7.308a)}\\
&= \bigcup_{j=1}^n \bigcup_{l\in A_j} B\big(x_l, r(x_l)\big) && \text{(7.308b)}\\
&= \bigcup_{j=1}^n \underbrace{\bigcup_{l\in A_j} \overline{B\big(x_l, r(x_l)\big)}}_{K_j} && \text{(7.308c)}\\
&= \bigcup_{j=1}^n K_j. && \text{(7.308d)}
\end{aligned}$$

115. C'est la définition 7.66 d'un compact.

□

Théorème 7.269 (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 7.275.

7.15.3 Ensembles enchainés

Soit (X, d) un espace métrique.

Définition 7.270.
Une ***ϵ-chaine*** *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 enchainée*** *si pour tout $\epsilon > 0$ et pour tout $a, b \in A$, il existe une ϵ-chaine joignant a et b dans A.*

Lemme 7.271.
Les rationnels dans $\mathbb{R}$ sont bien enchainés.

Démonstration. Soient p et q des rationnels avec $p < q$, ainsi que $\epsilon > 0$. Le lemme 1.375 nous permet de considérer un rationnel δ vérifiant $0 < \delta < \epsilon$. Et nous définissons les rationnels

$$r_k = p + k\delta. \tag{7.309}$$

Vu que $\mathbb{Q}$ est archimédien[116], il existe K tel que $r_K > q$. D'autre part, $r_0 = p < q$. Donc il existe $N = \max\{k \in \mathbb{N} \text{ tel que } r_k < q\}$.

Nous considérons la chaine $(r_0, \dots, r_N, q)$. Elle débute à $r_0 = p$ et termine à q ; pas de problèmes avec ça. À part pour le dernier pas, nous avons

$$|r_n - r_{n-1}| = \delta < \epsilon, \tag{7.310}$$

donc c'est bien une ϵ-chaine. Il reste à voir $|q - r_N|$. Nous avons $r_N \leqslant q \leqslant r_{N+1}$, et donc

$$0 \leqslant q - r_N \leqslant r_{N+1} - r_N = \delta \leqslant \epsilon. \tag{7.311}$$

Donc ok aussi pour ce dernier pas. □

Proposition 7.272 ([205, 1])**.**
Un espace métrique connexe[117] *est bien enchainé.*

Démonstration. Soit un espace métrique X et $\epsilon > 0$. La relation $x \sim y$ si et seulement si x et y peuvent être reliés par une ϵ-chaine est une relation d'équivalence.

Soit $x \in X$. Nous prouvons que la classe $[x]$ est ouverte. En effet soit $y \in [x]$, si $z \in B(y, \epsilon)$ nous avons $z \in [y]$, et donc $z \in [x]$. Nous en déduisons que $B(y, \epsilon) \subset [x]$, et donc que $[x]$ est ouvert par le théorème 7.7.

Donc les classes sont des ouverts.

Supposons que X n'est pas bien enchainé. Alors il existe ϵ pour lequel X possède plus qu'une classe d'équivalence. Soit $\{[x_k]\}_{k \in I}$ l'ensemble des classes d'équivalences.

Nous considérons un $i_0 \in I$ quelconque, et nous définissons les ouverts $A = [x_{i_0}]$ et

$$B = \bigcup_{k \in I \setminus \{i_0\}} [x_k]. \tag{7.312}$$

Ce sont deux ouverts disjoints qui recouvrent X qui n'est donc pas connexe. □

116. Proposition 1.335.
117. Définition 7.60.

Proposition 7.273.
La fermeture d'un ensemble bien enchainé dans un espace métrique compact (X, d) est connexe.

Démonstration. Soit $A \subset X$ un ensemble bien enchainé, 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 $\frac{1}{n}$-chaine $\{v_i^{(n)}\}_{i \in I_n}$ 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 7.265 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

$$\bigcup_{x \in \bar{A}} C_{a,x} = \bar{A}. \tag{7.313}$$

Vu que le membre de gauche est une union de connexes, c'est un connexe par la proposition 7.64. □

Corolaire 7.274.
Un espace métrique compact est connexe si et seulement si il est bien enchainé.

Démonstration. Dans le sens direct, c'est la proposition 7.272. Dans l'autre sens, si X est compact, alors X est fermé par le lemme 7.82(2). Et vu qu'il est fermé et bien enchainé, la proposition 7.273 implique qu'il est connexe. □

7.15.4 Produit fini d'espaces métriques

Pour rappel, la distance sur un espace produit est donnée par la définition 7.190.

Théorème 7.275 ([1]).
Un produit fini d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Démonstration. Soient $K_1, \ldots, K_n$ des compacts et $K = K_1 \times \ldots \times K_n$ le produit muni de sa métrique usuelle de la définition (7.190) (attention : chacun des K_i peut être de dimension infinie) :

$$d(\alpha, \beta) = \max\{d_i(\alpha_i, \beta_i)\} \tag{7.314}$$

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 7.124). 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 $k = 1, \ldots, n$, il existe $N_k > 0$ tel que si $p > N_k$ alors

$$d\big((\alpha_p)_k - \ell_p\big) \leqslant \epsilon. \tag{7.315}$$

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

$$d\big(\alpha_n, (\ell_1, \ldots, \ell_n)\big) \leqslant \epsilon. \tag{7.316}$$

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

$$\mathcal{P}_i = \mathcal{O}_i \times K_2 \times \ldots \times K_n, \tag{7.317}$$

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 7.276 ([206]).
Soient (E_n, d_n) des espaces métriques. Sur l'ensemble produit $E = \prod_{i=1}^{\infty} E_i$ nous définissons la métrique

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} d'_k(x_i, y_i) \tag{7.318}$$

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 7.277 (Tykhonov dénombrable[206]).
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 7.269.

7.15.5 Équicontinuité

Définition 7.278 ([1, 207, 208]).
*Soient un espace topologique X et un espace vectoriel topologique Y. Une famille H d'applications $X \to Y$ est **équicontinue** en $a \in X$ si pour tout voisinage V de 0 dans Y, il existe un voisinage U de a dans X tel que*

$$h(U) \subset h(a) + V \tag{7.319}$$

pour tout $h \in H$.

Nous disons que H est équicontinue si elle est équicontinue en tout point.

Lemme 7.279 ([1]).
Soient un espace métrique X, un espace vectoriel normé Y ainsi qu'une famille H d'isométries linéaires $X \to Y$. Alors H est équicontinue.

Démonstration. Nous suivons la définition 7.278 de l'équicontinuité. Soient $a \in X$ et un voisinage V de 0 dans Y. Nous considérons $r > 0$ tel que $B(0, r) \subset V$, et nous posons $U = B(a, r)$.

Si $x \in U$ et $h \in H$, nous avons

$$\|h(x) - h(a)\| = \|h(x - a)\| = d(x, a) < r, \tag{7.320}$$

de telle sorte que $h(x) \in h(a) + B(0, r)$.

Donc H est continue en a. Vu que a est arbitraire, H est équicontinue en tout point et donc équicontinue sur X. □

Lemme 7.280 ([208]).
Soit une famille de fonctions $f_i \colon X \to E$ indexée par un ensemble I où X est un espace topologique et E un espace métrique. Cette famille est équicontinue[118] en $x \in X$ si pour tout $\epsilon > 0$, il existe un voisinage V de x tel que

$$\|f_i(x) - f_i(y)\| < \epsilon \tag{7.321}$$

118. Définition 7.278.

pour tout i dès que $x, y \in V$.

La proposition suivante permet de montrer que certaines fonctions définies par une limite sont continues. Ce sera par exemple le cas de la fonction puissance, proposition 12.411.

Proposition 7.281 ([1, 208])**.**
Soit une suite équicontinue (f_i) de fonctions qui converge simplement vers f, alors f est continue.

Démonstration. Soit une suite équicontinue $f_i \colon X \to E$ convergeant simplement vers f. Soit $a \in X$. Nous prouvons que f est continue en a.. Pour cela nous considérons $\epsilon > 0$ et, conformément à l'hypothèse équicontinuité un voisinage V de a tel que $|f_i(a) - f_i(x)| < \epsilon$ pour tout $x \in V$.

Nous avons la majoration

$$|f(x) - f(a)| \leqslant |f(x) - f_i(x)| + |f_i(x) - f_i(a)| + |f_i(a) - f(a)|. \tag{7.322a}$$

Plusieurs majorations.

— Vu que $f_i \to f$, il existe N_1 tel que $|f(x) - f_i(x)| < \epsilon$ pour tout $i > N_1$.
— De plus, par définition de V, nous avons aussi $|f_i(x) - f_i(a)| \leqslant \epsilon$.
— Vu que $f_i \to f$, il existe N_2 tel que $|f_i(a) - f(a)| < \epsilon$ pour tout $i > N_2$.

Donc en prenant $x \in V$ et $i > \max\{N_1, N_2\}$ nous avons

$$|f(x) - f(a)| \leqslant 3\epsilon. \tag{7.323}$$

□

7.15.6 Continuité uniforme

Définition 7.282 ([209])**.**
*Soient deux espaces métriques (E, d) et (E', d'). Une application $f \colon E \to E'$ est **uniformément continue** si pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $d(x, y) \leqslant \delta$ implique $d'\big(f(x), f(y)\big) \leqslant \epsilon$.*

Dans l'uniforme continuité, le α qui fait fonctionner ϵ doit le faire fonctionner pour tous les $x, y \in E$. C'est la différence avec la continuité simple dans laquelle nous pouvons choisir, pour un même ϵ, un δ différent en chaque point.

Nous parlons plus d'uniforme continuité dans la section 12.8.

7.16 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 métriques, voire plus. Voir aussi la section 7.18 sur les espaces de Baire.

Définition 7.283.
*Un ensemble est dit **nulle part dense** si il n'est dense dans aucun intervalle.*

*Un ensemble dans $\mathbb{R}$ est de **première catégorie** ou **maigre** si il est une union dénombrable d'ensembles nulle part dense (c'est-à-dire d'ensembles denses sur aucun intervalle).*

Théorème 7.284 (Baire[210])**.**
Une réunion dénombrable d'ensembles nulle part denses est d'intérieur vide.

Démonstration. Soient $a \in S$ 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

$$B(x_1, r_1) \cap A_1 = \varnothing. \tag{7.324}$$

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) $B(x_n, r_n) \cap A_j = \emptyset$ 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 emboîtés de rayon décroissant vers zéro), elle converge [119] 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. □

7.17 Topologie des seminormes

Les principaux espaces topologiques construit avec des seminormes seront les espaces de fonctions de la définition 30.13. Nous verrons également la topologie $*$-faible sur $\mathscr{D}'(\Omega)$ en la définition 30.21.

Définition 7.285.
*Si E est un espace vectoriel sur le corps $\mathbb{K} = \mathbb{R}, \mathbb{C}$, une **seminorme** 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) $p(x + y) \leqslant p(x) + p(y)$

pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$.

La seule différence avec une norme est qu'une seminorme peut s'annuler en des éléments non-nuls de l'espace.

Lemme 7.286 ([211]).
Si p est une seminorme [120] nous avons

$$|p(x) - p(y)| \leqslant p(x - y). \tag{7.325}$$

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,

$$-p(h) \leqslant p(x + h) - p(x) \leqslant p(h) \tag{7.326}$$

ou encore

$$|p(x + h) - p(x)| \leqslant p(h) \tag{7.327}$$

qui donne le résultat demandé en posant $h = y - x$. □

Soit $(p_i)_{i \in I}$ une famille de seminormes sur E. Nous construisons alors une topologie sur E de la façon suivante.

Proposition-Définition 7.287 (Topologie et seminormes[212, 213]).
*Soient des seminormes $\{p_i\}_{i \in I}$. Pour tout J fini dans I nous définissons les **boules ouvertes***

$$B_J(x, r) = \{y \in E \text{ tel que } p_j(y - x) < r \, \forall j \in J\}. \tag{7.328}$$

*La **topologie** sur E donnée par la famille de seminorme 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 $\mathcal{O}$.*

Cela définit une topologie.

Proposition 7.288.
Soit un ensemble E muni de la topologie des seminormes $\{p_i\}_{i \in I}$. Une suite (x_n) dans E converge vers x si et seulement si pour tout $i \in I$,

$$p_i(x - x_n) \to 0. \tag{7.329}$$

119. Par la proposition 1.341
120. Définition 7.285.

Démonstration. Si la suite (x_n) converge[121] 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 $x_n \in B_J(x, r)$. □

¡¡ Avertissement/question au lecteur !! 7.289
Je n'ai pas vérifié si la proposition 7.290 est correcte. D'ailleurs je même pas trouvé l'énoncé ; et j'avoue n'avoir pas trop cherché.

La preuve serait sans doute similaire à ce qu'on a pour le lemme 7.192.

Proposition 7.290 ([1]).
Soient des espaces vectoriels munis de seminormes $\big(E, \{p_i\}_{i\in I}\big)$ et $\big(F, \{q_j\}_{j\in J}\big)$. Nous posons

$$\begin{aligned} r_{ij} \colon E \times F &\to \mathbb{R} \\ (x, y) &\mapsto \max\big\{p_i(x), q_j(y)\big\}. \end{aligned} \tag{7.330}$$

Alors :

(1) Les r_{ij} sont des seminormes.

(2) La topologie induite sur $E \times F$ par ces seminormes est la topologie produit.

Proposition 7.291 ([214, 1]).
Un espace vectoriel muni de seminormes sur un corps valué est un espace vectoriel topologique[122].

Démonstration. Soit un espace vectoriel E muni des seminormes $\{p_i\}_{i\in I}$. Sa topologie est donnée par la définition 7.287. Sur le corps $\mathbb{K}$, nous avons la topologie métrique 7.158.

(i) **Somme** Nous commençons par prouver que

$$\begin{aligned} f \colon E \times E &\to R \\ (x, y) &\mapsto x + y \end{aligned} \tag{7.331}$$

est continue. Soit un ouvert $\mathcal{O}$ de E ; nous allons prouver que $f^{-1}(\mathcal{O})$ est ouvert en prouvant qu'il contient une boule ouverte autour de chacun de ses points (théorème 7.7). Notez que $f^{-1}(\mathcal{O}) \subset E \times E$; la topologie sur cet ensemble est celle est seminormes r_{ij} données en (7.330). Nous allons en particulier utiliser la seminorme $q_i = r_{ii}$ donnée par

$$\begin{aligned} q_i \colon E \times E &\to \mathbb{R} \\ (x, y) &\mapsto \max\big\{p_i(x), p_i(y)\big\}. \end{aligned} \tag{7.332}$$

Soit $(a, b) \in f^{-1}(\mathcal{O})$. Vu que $a + b \in \mathcal{O}$ et que $\mathcal{O}$ est ouvert, la partie $\mathcal{O}$ contient une boule ouverte autour de $a + b$ (définition 7.287). Soit $i \in I$ et $r > 0$ tels que

$$B_i(a + b, r) \subset \mathcal{O}. \tag{7.333}$$

Nous allons prouver qu'il existe un $s > 0$ tel que $B_i\big((a, b), s\big) \subset f^{-1}(\mathcal{O})$, et plus précisément que

$$f\Big(B_i\big((a, b), s\big)\Big) \subset B_i(a + b, r). \tag{7.334}$$

À gauche, B_i est la boule dans $E \times E$ pour la seminorme (7.332). Soit $(x, y) \in B_i\big((a, b), s\big)$, c'est à dire

$$q_i\big((a, b) - (x, y)\big) \leqslant s. \tag{7.335}$$

121. Définition 7.12.
122. Définition 7.148.

Pour savoir si $f(x,y) \in B_i(a+b,r)$, nous posons $x = a+h$ et $y = b+k$ et nous calculons

$$\begin{aligned}
p_i\big(f(x,y)-(a+b)\big) &= p_i(x+y-a-b) && (7.336a)\\
&= p_i(h+k) && (7.336b)\\
&\leqslant p_i(h)+p_i(k) && (7.336c)\\
&\leqslant 2\max\{p_i(h),p_i(k)\} && (7.336d)\\
&= 2q_i(h,k) && (7.336e)\\
&\leqslant 2s \qquad \text{par (7.335).} && (7.336f)
\end{aligned}$$

En posant $s = r/2$, nous avons bien $f(x,y) \in \mathcal{O}$, et donc $f^{-1}(\mathcal{O})$ est un ouvert ; f est alors continue.

(ii) **Produit** Nous nommons $\mathbb{K}$ le corps de l'espace vectoriel E. Nous devons voir que l'application

$$\begin{aligned}
f\colon \mathbb{K}\times E &\to E\\
(\lambda,x) &\mapsto \lambda x
\end{aligned} \tag{7.337}$$

est continue.

La topologie sur $\mathbb{K}$ est sa topologie métrique, c'est à dire la topologie de son unique seminorme $\lambda \mapsto |\lambda|$. La topologie sur $\mathbb{K}\times E$ est donc celle des seminormes

$$\begin{aligned}
q_i\colon \mathbb{K}\times E &\to \mathbb{R}\\
(\lambda,x) &\mapsto \max\{|\lambda|,p_i(x)\}.
\end{aligned} \tag{7.338}$$

Nous pouvons donc reprendre le même cheminement que celui que nous avons pris pour la somme. Soit un ouvert $\mathcal{O}$ dans E ; nous considérons $(\lambda,a) \in f^{-1}(\mathcal{O})$. Vu que $f(\lambda,a) \in \mathcal{O}$, et que $\mathcal{O}$ est ouvert pour la topologie des $\{p_i\}_{i\in I}$, il existe $i \in I$ et $r > 0$ tel que $B_i\big(f(\lambda,a),r\big) \subset \mathcal{O}$.

Nous allons prouver qu'il existe $s > 0$ tel que

$$f\Big(B_i\big((\lambda,a),s\big)\Big) \subset B_i(\lambda a,r). \tag{7.339}$$

Ici encore, à gauche B_i est la boule pour la seminorme q_i donnée en (7.338). Soit $(\mu,x) \in B_i\big((\lambda,a),s\big)$, c'est à dire

$$q_i\big((\mu,x)-(\lambda,a)\big) = \max\{|\lambda-\mu|,p_i(a-x)\} < s. \tag{7.340}$$

En particulier nous avons les deux inégalités

$$\begin{cases}
|\lambda-\mu| < s, & (7.341a)\\
p_i(a-x) < s. & (7.341b)
\end{cases}$$

Nous avons le calcul suivant :

$$\begin{aligned}
p_i\big(f(\mu,x),\lambda a\big) &= p_i(\mu x-\lambda a) && (7.342a)\\
&= p_i(\mu x-\lambda x+\lambda x-\lambda a) && (7.342b)\\
&\leqslant p_i\big((\mu-\lambda)x\big)+|\lambda|p_i(x-a) && (7.342c)\\
&= |\mu-\lambda|p_i(x)+|\lambda|p_i(x-a) && (7.342d)\\
&\leqslant sp_i(x)+|\lambda|s. && (7.342e)
\end{aligned}$$

C'est le moment de chercher une majoration pour $p_i(x)$:

$$p_i(x) = p_i\big(a+(x-a)\big) \leqslant p_i(a)+p_i(x-a) \leqslant p_i(a)+s. \tag{7.343}$$

Nous pouvons continuer la majoration (7.342) tout en ne nous posant pas de questions sur le sens de l'inégalité parce que nous cherchons $s > 0$:

$$p_i\big(f(\mu,x),\lambda a\big) \leqslant sp_i(x) + |\lambda|s \tag{7.344a}$$
$$\leqslant s\big(p_i(a)+s\big) + |\lambda|s \tag{7.344b}$$
$$= s^2 + \big(|\lambda| + p_i(a)\big)s. \tag{7.344c}$$

Nous devons prouver l'existence d'un $s > 0$ tel que $s^2 + \big(|\lambda| + p_i(a)\big)s < r$; autrement dit nous devons résoudre l'inéquation

$$s^2 + \big(|\lambda| + p_i(a)\big)s - r < 0. \tag{7.345}$$

Nous sommes en présence d'un polynôme du second degré en s qui vaut $-r < 0$ en $s = 0$. Par continuité, il existe un voisinage de $s = 0$ dans $\mathbb{R}$ sur lequel le polynôme reste strictement négatif. Il suffit de prendre un s positif dans ce voisinage.

□

La proposition suivante est pratiquement une copie de la proposition 7.258.

Proposition 7.292.
Soit $f\colon \mathbb{R} \to (E, p_i)_{i\in I}$ une application. Nous avons équivalence entre

(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

$$f\big(B(t_0,\delta)\big) \subset B_i\big(f(t_0),\epsilon\big). \tag{7.346}$$

Démonstration. L'équivalence (1) ⇔ (2) est la définition 7.41.

Prouvons (2) ⇒ (3). Soient $i \in I$ et $\epsilon > 0$. 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

$$f\big(B(t_0,\delta)\big) \subset f(V) \subset B_i\big(f(t_0),\epsilon\big). \tag{7.347}$$

Prouvons (3) ⇒ (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

$$f\big(B(t_0,\delta)\big) \subset B_i\big(f(t_0),\epsilon\big) \subset W. \tag{7.348}$$

□

Lorsqu'on a un espace E muni d'une quantité dénombrable de seminormes $\{p_k\}_{k\in I}$ nous définissons l'écart [123]

$$d(x,y) = \sup_{k\geqslant 1} \min\big\{\frac{1}{k}, p_k(x-y)\big\}. \tag{7.349}$$

Notons que cet écart est invariant par translation au sens où pour tout x, y, h dans E nous avons

$$d(x+h, y+h) = \sup_{k\geqslant 1} \min\big\{\frac{1}{k}, p_k(x-y)\big\} = d(x,y). \tag{7.350}$$

Proposition 7.293.
Si X est un espace topologique dont la topologie est donnée par une famille dénombrable de seminormes, alors il est métrisable.

123. Dans le cas de $E = \mathscr{D}(K)$, la première seminorme est numérotée à zéro, donc il faudra poser $d(\varphi_1,\varphi_2)$ avec p_{k-1} au lieu de p_k.

Proposition 7.294 ([211]).
La topologie donnée par les boules

$$B_k(a,r) = \{x \in E \text{ tel que } \forall\, k \leqslant \frac{1}{r}, p_k(x-a) < r\} \tag{7.351}$$

est la même que celle « usuelle » donnée par les seminormes. 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 si il est ouvert pour l'autre.

Démonstration. Pour cette démonstration nous allons préfixer par d les notions topologiques issues des boules (7.351) et par P celle des seminormes : P-continue, d-ouvert, etc.

D'abord nous avons

$$B(a,r) = \bigcap_{k \leqslant \frac{1}{k}} B_k(a,r). \tag{7.352}$$

Si $\mathcal{O}$ est un d-ouvert, il contient une d-boule autour de chacun de ses points. Or d'après la formule (7.352), 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 seminormes 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

$$\begin{aligned}
|p_k(y) - p_k(x)| &\leqslant p_k(x-y) && (7.353a)\\
&= \min\{\frac{1}{k}, p_k(x-y)\} && (7.353b)\\
&\leqslant d(x,y) && (7.353c)\\
&\leqslant \epsilon. && (7.353d)
\end{aligned}$$

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 $x \in B_k(a,r)$. 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

$$\big|p_k(a-x) - p_k(a-y)\big| \leqslant d(x,y) \leqslant \epsilon. \tag{7.354}$$

Par conséquent le nombre $p_k(a-y)$ est dans l'intervalle

$$p_k(a-x) \pm \epsilon \tag{7.355}$$

et il suffit de prendre $\epsilon < \frac{r - p_k(a-x)}{2}$. □

7.17.0.1 Norme induite sur la topologie quotient

Proposition-Définition 7.295 (Norme quotient[215]).
Soient un espace vectoriel topologique normé E, et un sous-espace M. Pour $\alpha \in E/M$ nous posons

$$\|\alpha\|_{E/M} = d(\alpha, M) \tag{7.356}$$

où $d(\alpha, M)$ est la distance entre la partie α et la partie M.

Nous avons :

(1) La formule $\|\alpha\| = \inf_{u \in \alpha} \|u\|_E$.

(2) L'opération $\|.\|$ est une seminorme[124] *sur l'espace vectoriel E/M.*

(3) C'est une norme si et seulement si M est fermé.

*Nous parlons de **(semi)norme quotient**.*

124. Définition 7.285.

Démonstration. Point par point.

(i) **Pour (1)** La définition de $\|\alpha\|$ est $\|\alpha\| = d(\alpha, M) = \inf_{\substack{u\in\alpha\\ v\in M}} \|u - v\|$. Mais

$$\{u - v \text{ tel que } u \in \alpha, v \in M\} = \alpha, \tag{7.357}$$

donc

$$d(\alpha, M) = \inf_{u\in\alpha} \|u\|. \tag{7.358}$$

(ii) **Pour (2)** Nous devons vérifier les propriétés de la définition 7.285. D'abord en tant que distance, nous avons $\|\alpha\| \geqslant 0$ pour tout α.

Si $\lambda \in \mathbb{R}$ nous avons aussi

$$\|\lambda\alpha\| = \inf_{u\in\lambda\alpha} \|u\| = \inf_{u\in\alpha} \|\lambda u\| = |\lambda| \inf_{u\in\alpha} \|u\| = |\lambda|\|\alpha\|. \tag{7.359}$$

Enfin si $\alpha, \beta \in E/M$, nous avons

$$\inf_{u\in\alpha+\beta} \|u\| \leqslant \inf_{\substack{u\in\alpha\\ v\in\beta}} \|u + v\| \leqslant \inf_{\substack{u\in\alpha\\ v\in\beta}} \big(\|u\| + \|v\|\big) = \inf_{u\in\alpha} \|u\| + \inf_{v\in\beta} \|v\| = \|\alpha\| + \|b\|. \tag{7.360}$$

(iii) **Pour (3) en supposant que M est femré** Nous supposons que M est fermé et nous montrons que $\|.\|$ est une norme. Nous supposons donc que $d(\alpha, M) = 0$ et nous prouvons que $\alpha = 0$. Soit $x \in E$ tel que $\alpha = [x]$. Nous avons donc

$$\alpha = \{x - v \text{ tel que } v \in M\} \tag{7.361}$$

Notez qu'on a écrit $-v$ et non $+v$. De toutes façons M est vectoriel ; ça ne change rien et ça tombera mieux plus bas.

Nous avons donc

$$0 = \|\alpha\| = d(\alpha, M) = \inf_{u\in\alpha} \|u\| = \inf_{v\in M} \|x - v\|. \tag{7.362}$$

Il existe donc une suite (v_n) dans M telle que $\|x - v_n\| \to 0$. La suite est donc convergente : $v_n \to x$. Comme M est fermé, la proposition 7.217 nous indique que la limite doit être dans M. Autrement dit : $x \in M$. Par définition des classes nous avons alors $\alpha = [x] = 0$.

(iv) **Pour (3) en supposant que $\|.\|$ est une norme** Supposons que M n'est pas fermé. Il ne contient donc pas son adhérence. Soit $a \in \mathrm{Adh}(M)\backslash M$. Vu que M est vectoriel, nous supposons que $a \neq 0$.

Étant donné que a est dans l'adhérence de M nous avons $d(a, M) = 0$ et donc $\|[a]\| = 0$.

□

¡¡ Avertissement/question à la lectrice !! 7.296

Je ne suis pas certain de la proposition 7.297. Peut-être qu'il faut ajouter l'hypothèse que M est fermé.

Proposition 7.297.

Soient un espace vectoriel topologique normé E, et un sous-espace M. La topologie quotient[125] *sur E/M est la même que celle de la seminorme induite*[126].

125. Topologie quotient, définition 7.20.
126. Voir la définition 7.295.

7.17.1 Espace dual

Nous parlerons plus en détail d'espace dual d'un espace normé en la section 11.14.

Définition 7.298.
Soient F un espace métrique et E un espace topologique vectoriel. Une topologie possible[127] *sur l'espace des applications linéaires $\mathcal{L}(E,F)$ est la* ***topologie $*$-faible*** *qui est la topologie des seminormes*

$$p_v(T) = \|T(v)\|_F. \tag{7.363}$$

C'est une famille de seminormes 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 7.299.
Soient E un espace muni de la topologie des seminormes $\{p_i\}_{i\in I}$ et F un espace métrique. Soient 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 :

$$T_n \xrightarrow{*} T \tag{7.364a}$$

$$p_v(T_n - T) \to 0 \,\forall v \in E \qquad \text{proposition 7.288} \tag{7.364b}$$

$$\|T_n(v) - T(v)\|_E \to 0 \,\forall v \in E \tag{7.364c}$$

$$T_n(v) \xrightarrow{E} T(v). \tag{7.364d}$$

□

7.17.2 Espace $C^k(\mathbb{R}, E')$

Nous revenons à nos histoires de limites de la définition 7.12.

Proposition 7.300 (Unicité de la limite dans un dual topologique)**.**
Soient E un espace métrique et E' son dual topologique muni de sa topologie de la définition 7.298. 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 $t \to t_0$. Soient $\epsilon > 0$ 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 $\mathbb{R}$) :

$$\lim_{t\to t_0} u_t(x) = T(x). \tag{7.365}$$

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

$$\lim_{t\to t_0} u_t(x) = T'(x). \tag{7.366}$$

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 7.301.
Soit $u\colon \mathbb{R} \to E'$ une fonction continue. Alors

(1) pour tout $x \in E$ la fonction $t \mapsto u_t(x)$ est continue,

127. C'est, dans l'idée, celle qui sera choisie pour les espaces de distributions, voir la définition 30.21.

(2) pour tout $x \in E$ nous avons la limite dans $\mathbb{R}$

$$\lim_{t \to t_0} u_t(x) = u_{t_0}(x), \tag{7.367}$$

(3) nous avons la limite dans E'

$$\lim_{t \to t_0} u_t = u_{t_0}. \tag{7.368}$$

Démonstration. Soient $x \in E$ et $\epsilon > 0$. Par la proposition 7.292 la continuité de u donne un $\delta > 0$ tel que

$$u_{B(t_0,\delta)} \subset B_x(u_{t_0}, \epsilon). \tag{7.369}$$

C'est-à-dire que si $|t - t_0| \leqslant \delta$ nous avons

$$\big|u_{t_0}(x) - u_t(x)\big| < \epsilon, \tag{7.370}$$

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 (7.367).

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

$$u_{B(t_0,\delta)} \subset B_i(u_{t_0}, \epsilon) \subset \mathcal{O}. \tag{7.371}$$

Cela signifie bien que

$$|t - t_0| \leqslant \delta \Rightarrow u_t \in \mathcal{O}, \tag{7.372}$$

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 7.92 de la limite et le résultat d'unicité 7.300. □

Définition 7.302.
*Si nous avons une application $u \colon \mathbb{R} \to E'$ nous considérons sa **dérivée** donnée par la limite*

$$u'_{t_0} = \lim_{t \to t_0} \frac{u_t - u_{t_0}}{t - t_0}. \tag{7.373}$$

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 utilisations que nous ferons de ces espaces seront les espaces de fonctions à valeurs dans les distributions tempérées dont nous parlerons dans la section 30.4.

7.18 Espaces de Baire

Définition 7.303.
*Un **espace de Baire** est un espace topologique dans lequel toute intersection dénombrable d'ouverts denses est dense.*

Lemme 7.304 ([216]).
Un espace topologique est de Baire si et seulement si toute union dénombrable de fermés d'intérieur vides est d'intérieur vide.

Théorème 7.305 (Théorème de Baire[216]).
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. (i) **Espaces topologiques localement compacts**

(ii) **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

$$\bigcap_{n\in\mathbb{N}} U_n \cap V \tag{7.374}$$

contient l'intersection des B_n. Par le théorème 7.259 des fermés emboîtés (que nous utilisons parce que E est métrique et complet), cette intersection est non vide.

(iii) **Ouvert d'un espace de Baire**

□

Parmi les applications du théorème de Baire, nous avons

— Le théorème de Banach-Steinhaus 11.136.

— Le théorème de l'application ouverte 11.145.

Chapitre 8

Espaces affines

8.1 Vecteurs agissant sur un espace

Définition 8.1.
Soit E, un espace vectoriel. Un ***espace affine modelé sur*** *E est un ensemble $\mathcal{E}$ sur lequel le groupe $(E, +)$ agit à droite transitivement et librement*[1].

Étant donné que E est un groupe commutatif, l'action peut être vue indifféremment à gauche ou à droite. Si $M \in \mathcal{E}$ et si $x \in E$ nous notons $M + x$ au lieu de $x \cdot M$ le résultat de l'action de x sur M.

8.2.
Lorsque nous écrivons « $M + x$ », le symbole plus n'est pas une loi de composition interne de $\mathcal{E}$, mais une action.

Proposition-Définition 8.3.
Soient $N, M \in \mathcal{E}$. Il existe un unique $x \in E$ tel que $M + x = N$.

Nous noterons $\overrightarrow{MN}$ ce vecteur.

Démonstration. La transitivité de l'action assure l'existence et la liberté assure l'unicité. □

Lemme 8.4.
Pour tout élément A nous avons $A + 0 = A$.

Démonstration. Soit $B \in \mathcal{E}$. Nous avons :

$$B + 0 = B + (0 + 0) = (B + 0) + 0 \tag{8.1}$$

parce que le $+$ dénote une action.

En appliquant cette égalité à l'élément $B = A - 0$ nous trouvons l'égalité demandée. □

Proposition 8.5.
Soit un espace affine $\mathcal{E}$ modelé sur l'espace vectoriel E. Soient $A, B, C \in \mathcal{E}$. Nous avons les égalités suivantes dans E :

(1) $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ *(relations de Chasles),*

(2) $\overrightarrow{AA} = 0$,

(3) $\overrightarrow{BA} = -\overrightarrow{AB}$.

Démonstration. Point par point.

1. Définition 2.45.

(i) **Pour (1)** Nous avons, par définition 8.3 les égalités

$$\begin{cases} C = A + \overrightarrow{AC} & (8.2a) \\ B = A + \overrightarrow{AB} & (8.2b) \\ C = B + \overrightarrow{BC} & (8.2c) \end{cases}$$

En substituant les deux premières dans la troisième, nous trouvons $A + \overrightarrow{AB} + \overrightarrow{BC} = A + \overrightarrow{AC}$. Par liberté de l'action, nous pouvons « simplifier » par A et trouver la relation de Chasles.

(ii) **Pour (2)** Nous avons $A + \overrightarrow{AA} = A$, mais aussi $A + 0 = A$. Par unicité nous avons $\overrightarrow{AA} = 0$.

(iii) **Pour (3)** Nous avons $B + \overrightarrow{BA} = A$ et $A + \overrightarrow{AB} = B$. En mettant bout à bout,

$$B + \overrightarrow{BA} + \overrightarrow{AB} = B. \tag{8.3}$$

Donc $\overrightarrow{BA} + \overrightarrow{AB} = 0$.

□

8.6.
Si E est un espace vectoriel, le groupe $(E, +)$ agit sur E par l'action $t_y(x) = y + x$. Utilisant cette action nous construisons l'**espace affine canonique** de E. En particulier nous notons $\mathcal{E}_n(\mathbb{K})$ l'espace affine canonique de $\mathbb{K}^n$ vu comme espace vectoriel sur $\mathbb{K}$.

— En tant qu'ensembles, $\mathcal{E}_n(\mathbb{K}) = \mathbb{K}^n$.

— Sur cet espace en particulier, si $M, N \in \mathcal{E}_n(\mathbb{K})$, nous avons $\overrightarrow{MN} = N - M$ où à droite, la différence est la différence vectorielle dans $\mathbb{K}^n$.

Ces deux points se généralisent immédiatement à un espace vectoriel E au lieu de $\mathbb{K}^n$.

8.2 Repères cartésiens affines

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathcal{E}$ un espace affine construit sur E.

Définition 8.7.
*Un multiplet $(A, e_1, \dots, e_n)$ où A est un point de $\mathcal{E}$ et $\{e_i\}$ est une base de E est un **repère cartésien** de $\mathcal{E}$.*

*Nous disons que $\{e_i\}$ est la **base associée** au repère.*

Proposition 8.8.
Si $\mathcal{E}$ est un espace affine modelé sur l'espace vectoriel E de dimension n sur le corps $\mathbb{K}$, et si $\big(A, \{e_i\}_{i=1,\dots,n}\big)$ est un repère cartésien, alors

$$\begin{aligned} \phi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1, \dots, x_n) &\mapsto A + \sum_i x_i e_i. \end{aligned} \tag{8.4}$$

est une bijection.

*Ces nombres x_i sont les **coordonnées** du point $A + \sum_i x_i e_i$ dans le repère (A, e_i).*

Démonstration. L'application φ est surjective parce que l'action de E sur $\mathcal{E}$ est transitive et injective parce que l'action est libre. □

8.3 Classification affine des coniques

Soit une conique $f(x, y) = 0$ avec

$$f(x, y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f \tag{8.5}$$

dans le repère $R = (A, e_i)$.

Lemme 8.9.
La signature de la quadratique

$$q(x,y) = ax^2 + 2bxy + cy^2 \tag{8.6}$$

ne dépend pas de la base choisie et un changement de variables

$$\begin{cases} \tilde{x} = \alpha x + \beta y & (8.7a) \\ \tilde{y} = \gamma x + \delta y & (8.7b) \end{cases}$$

peut nous amener dans trois cas :

$$q(x,y) = \begin{cases} \tilde{x}^2 + \tilde{y}^2 & \textit{genre ellipse} \\ \tilde{x}^2 - \tilde{y}^2 & \textit{genre hyperbole} \\ \tilde{x}^2 & \textit{genre parabole.} \end{cases} \tag{8.8}$$

Dans le troisième cas, la matrice de q est de rang 1.

Nous cherchons maintenant à savoir si un point $I = (x_0, y_0)$ est un centre de symétrie de $f(x,y) = 0$. Pour cela nous choisissons le repère centré en I, c'est-à-dire que nous posons

$$\begin{cases} x = x_0 + \tilde{x} & (8.9a) \\ y = y_0 + \tilde{y}. & (8.9b) \end{cases}$$

Un peu de calcul montre qu'alors la conique s'écrit

$$f(x_0, y_0) + q(\tilde{x}, \tilde{y}) + (2ax_0 + 2by_0 + 2d)\tilde{x} + (2bx_0 + 2cy_0 + 2e)\tilde{y} = 0. \tag{8.10}$$

Lemme 8.10.
Le point I sera un centre de symétrique si les termes linéaires en $\tilde{x}$ et $\tilde{y}$ s'annulent, c'est-à-dire si

$$\begin{cases} ax_0 + by_0 + d = 0 & (8.11a) \\ bx_0 + cy_0 + e = 0. & (8.11b) \end{cases}$$

Nous supposons que $(d,e) \neq (0,0)$, sinon la conique de départ serait déjà centrée. Le déterminant du système (8.11) est

$$\delta = ac - b^2. \tag{8.12}$$

Si ce dernier est différent de zéro, le système possède une unique solution et la conique aura alors un unique centre de symétrie.

Si le déterminant du système est nul, il y a soit aucun centre de symétrie, soit une infinité. Dans le premier cas nous sommes en présence d'une parabole, et dans le second cas de deux droites parallèles.

Exemple 8.11.
Soit

$$f(x,y) = x^2 + 2xy - y^2 - 6x + 2y - 1 = 0 \tag{8.13}$$

donnée dans le repère affine $R = (A, \{e_i\})$. Nous commençons par étudier la signature de $q(x,y) = x^2 + 2xy - y^2$ dont la matrice symétrique est

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{8.14}$$

Son polynôme caractéristique est $\lambda^2 - 2$ dont les racines sont $\pm\sqrt{2}$. La signature est donc $(1,1)$ et nous sommes en présence d'une conique de genre hyperbole. Nous cherchons le centre en suivant le lemme 8.10. Nous posons $x = \tilde{x} + x_0$, $y = \tilde{y} + y_0$, et nous cherchons à résoudre le système

$$\begin{cases} x_0 + y_0 - 3 = 0 & (8.15a) \\ x_0 - y_0 + 1 = 0. & (8.15b) \end{cases}$$

L'unique solution est $(x_0, y_0) = (1, 2)$. Nous considérons le repère centré en (x_0, y_0), c'est-à-dire le repère

$$R' = (I, \{e_i\}) \tag{8.16}$$

avec $I = A + x_0 e_1 + y_0 e_2$ où A est l'origine du repère dans lequel l'équation (8.13) était donnée.

Par construction dans ce repère nous avons la conique

$$f(x_0, y_0) + q(\tilde{x}, \tilde{y}) = 0, \tag{8.17}$$

c'est-à-dire

$$\tilde{x}^2 + 2\tilde{x}\tilde{y} - \tilde{y}^2 - 2 = 0. \tag{8.18}$$

Maintenant, nous avons une quadrique centrée que nous voulons mettre sous une forme plus canonique :

$$\left(\frac{1}{\sqrt{2}}(\tilde{x} + \tilde{y})\right)^2 - \tilde{y}^2 - 1 = 0. \tag{8.19}$$

Nous posons donc

$$\begin{cases} X = \dfrac{1}{\sqrt{2}}(\tilde{x} + \tilde{y}) & (8.20a) \\ Y = \tilde{y}, & (8.20b) \end{cases}$$

pour trouver l'hyperbole

$$X^2 - Y^2 - 1 = 0. \tag{8.21}$$

Cherchons le changement de base correspondant. Pour trouver les coordonnées de e'_1 dans la base (e_1, e_2) nous cherchons pour quelles valeurs de x, y nous avons $e'_1 = xe_1 + ye_2$. Le point e'_1 étant caractérisé par $X = 1$, $Y = 0$ nous avons à résoudre

$$\begin{cases} \dfrac{1}{\sqrt{2}}(x + y) = 1 & (8.22a) \\ y = 0, & (8.22b) \end{cases}$$

ce qui donne $x = \sqrt{(2)}$ et $y = 0$. Donc

$$e'_1 = \sqrt{2}e_1. \tag{8.23}$$

Pour trouver e'_2, c'est le même raisonnement en posant $X = 0$ et $Y = 1$. Le résultat est :

$$e'_2 = -e_1 + e_2. \tag{8.24}$$

Résumons :

$$\begin{cases} e'_1 = \sqrt{2}e_1 & (8.25a) \\ e'_2 = -e_1 + e_2. & (8.25b) \end{cases}$$

Il y a un dicton qui dit que les vecteurs de base se transforment avec la matrice inverse des coefficients. Prenons la matrice M donnée par

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = M \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \tag{8.26}$$

Calculons la matrice inverse.

```
sage: M=matrix([  [1/sqrt(2), 1/sqrt(2)],[0,1]  ])
sage: M
[1/2*sqrt(2) 1/2*sqrt(2)]
[          0            1]
sage: M.inverse()
[sqrt(2)       -1]
[      0        1]
```

tex/sage/sageSnip023.sage

Nous voyons que les colonnes de la matrice M^{-1} donnent les coordonnées des vecteurs e'_1 et e'_2.

△

8.4 Applications affines

Voici la définition d'une application affine entre deux espaces affines. La définition 9.143 donnera le définition d'une application affine entre espaces vectoriels.

Définition 8.12.
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines sur les espaces vectoriels E et E' (sur le même corps $\mathbb{K}$). Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est dite ***affine*** *si pour tout $M \in \mathcal{E}$, il existe une application linéaire*[2] *$u_M\colon E \to E'$ telle que*

$$f(M+x) = f(M) + u_M(x) \tag{8.27}$$

pour tout $x \in E$.

La définition suivante permet de décomposer une application affine en une partie linéaire et une translation. À partir de là, la proposition 8.62 nous donnera une structure de groupe sur $\mathrm{Aff}(\mathbb{R}^n)$.

Lemme-Définition 8.13 (partie linéaire d'une application affine[1])**.**
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines sur les espaces vectoriels E et E' (sur le même corps $\mathbb{K}$). Nous considérons une application affine $f\colon \mathcal{E} \to \mathcal{E}'$.

Il existe une unique application linéaire $u\colon E \to E'$ telle que

$$f(M+x) = f(M) + u(x) \tag{8.28}$$

pour tout $x \in E$ et pour tout $M \in \mathcal{E}$.

Cette application linéaire est appelée ***partie linéaire*** *de f. Pour varier les notations, nous noterons souvent $f = \alpha \circ \tau_v$ pour une application linéaire α et la translation τ_v de vecteur v.*

Démonstration. En plusieurs étapes.

(i) **Unicité** Supposons que u_1 et u_2 vérifient la propriété, alors pour tout $x \in E$ et tout $M \in \mathcal{E}$ nous avons $f(M+x) = f(M) + u_1(x)$ et $f(M+x) = f(M) + u_2(x)$. Cela suffit à nous convaincre que $u_1 = u_2$.

(ii) $u_M = u_N$ Avant de prouver l'existence, nous considérons $M, N \in \mathcal{E}$ et les applications linéaires u_M et u_N vérifiant l'équation (8.27) pour M et N respectivement. Prouvons que $u_M = u_N$.
Posons

$$f(M+x) = f(M) + u_M(x) \tag{8.29a}$$
$$f(N+y) = f(N) + u_N(y). \tag{8.29b}$$

Définissons $a \in E$ par $N = M + a$; nous avons d'une part

$$f(N+y) = f(M+y+a) = f(M) + u_M(y+a), \tag{8.30}$$

et d'autre part

$$f(N+y) = f(M+a) + u_N(y) = f(M) + u_M(a) + u_N(y). \tag{8.31}$$

Par conséquent $u_M(y+a) = u_M(a) + u_N(y)$. Par linéarité $u_N = u_M$.

(iii) **Existence** Soit $M \in \mathcal{E}$. Nous affirmons que u_M fait l'affaire. En effet, soient $N \in \mathcal{E}$ et $x \in E$. Puisque $u_M = u_N$ nous avons

$$f(N+x) = f(M) + u_N(x) = f(M) + u_M(x). \tag{8.32}$$

Donc effectivement u_M peut être utilisé en tout point de $\mathcal{E}$.

□

Ce lemme est important car il permet de démontrer qu'une application est affine en prouvant la linéarité des u_M séparément sans devoir prouver qu'elles sont égales.

2. Définition 4.27.

8.4.1 Autres propriétés

Lemme 8.14 ([1]).
Soient $M \in \mathcal{E}$ et $A, B \in \mathcal{E}$ deux points donnés par $A = M + x_a$, $B = M + x_b$. Soit encore une application affine f sur $\mathcal{E}$. Alors

$$\overrightarrow{AB} = u_f(x_b - x_a). \tag{8.33}$$

Démonstration. En appliquant f à $A = M + x_a$ et $B = M + x_b$,

$$f(A) = f(M) + u_f(x_a) \tag{8.34a}$$
$$f(B) = f(M) + u_f(x_b). \tag{8.34b}$$

Donc $f(B) = f(A) - u_f(x_a) + u_f(x_b)$ ou encore

$$f(B) = f(A) + u_f(x_b - x_a). \tag{8.35}$$

□

Remarque 8.15.
La condition (8.27) pour tout $M \in \mathcal{E}$ est équivalente à demander

$$f \circ t_x = t_{u(x)} \circ f \tag{8.36}$$

pour tout $x \in E$.

Proposition 8.16 ([1]).
Soit une application affine $f \colon \mathcal{E} \to \mathcal{E}'$.

(1) Il existe une unique application linéaire u_f telle que $f(M + x) = f(M) + u_f(x)$ pour tout $M \in \mathcal{E}$ et tout $x \in E$.

(2) L'application u_f est injective si et seulement si f est injective.

(3) L'application u_f est surjective si et seulement si f est surjective.

Démonstration. En plein de parties.

(i) **Pour (1)** La partie (1) est le lemme 8.13.

(ii) **Si u_f est injective** Soient $M, N \in \mathcal{E}$ tels que $f(M) = f(N)$. Nous avons

$$f(M) = f(N) = f\big(M + (N - M)\big) = f(M) + u_f(N - M), \tag{8.37}$$

donc $u_f(N - M) = 0$. Vu que u_f est injective, nous déduisons que $N - M = 0$.

(iii) **Si f est injective** Soient $x, y \in E$ tels que $u_f(x) = u_f(y)$. Soit M quelconque dans $\mathcal{E}$; nous avons

$$f(M + x) = f(M) + u_f(x) = f(M) + u_f(y) = f(M + y). \tag{8.38}$$

L'injectivité de f nous indique alors que $M + x = M + y$ et donc que $x = y$ parce que l'action de E sur $\mathcal{E}$ est libre.

(iv) **Si u_f est surjective** Soit $M \in \mathcal{E}$. Nous allons trouver un élément de $\mathcal{E}$ dont l'image par f est M. Soient $N \in \mathcal{E}$ et $x \in E$ tels que $u_f(x) = M - f(N)$.
Alors nous avons $f(N + x) = f(N) + u_f(x) = M$.

(v) **Si f est surjective** Soit $a \in E$. Nous voulons $x \in E$ tel que $u_f(x) = a$. Soit $M \in E$. Vu que f est surjective, il existe $N \in \mathcal{E}$ tel que $f(N) = f(M) + a$.
Posons $x = N - M$. Nous avons d'une part

$$f(M + x) = f(M) + u_f(x) \tag{8.39}$$

et d'autre part

$$f(M + x) = f\big(M + (N - M)\big) = f(N) = f(M) + a. \tag{8.40}$$

En égalisant nous trouvons $u_f(x) = a$.

□

Proposition 8.17.
Soient des espaces affines $\mathcal{E}$ et $\mathcal{E}'$ de même dimension. Une application affine $f\colon \mathcal{E} \to \mathcal{E}'$ est injective si et seulement si elle est surjective.

Démonstration. Nous allons utiliser les équivalences de la proposition 8.16, ainsi que le corollaire 4.46 pour la partie linéaire. Nous avons les équivalences :

$$f \text{ est injective} \Leftrightarrow u_f \text{ est injective} \Leftrightarrow u_f \text{ est surjective} \Leftrightarrow f \text{ est surjective}. \tag{8.41}$$

□

Exemple 8.18.
L'espace $\mathbb{R}^n$ est très particulier parce qu'il agit sur lui-même ; il est donc un espace affine à lui tout seul : $\mathcal{E} = E = \mathbb{R}^n$.

Dans le cas de $\mathbb{R}^n$, en posant $M = 0$ dans la condition (8.27), si f est une application affine il existe une application linéaire α et un vecteur v tel que $f = \tau_v \circ \alpha$.

Notons que ça n'a pas de sens de poser $M = 0$, et la décomposition $f = \tau_v \circ \alpha$ n'a aucun sens en général. En particulier, nous ne pouvons pas appliquer une application linéaire à un élément d'un espace affine général. △

Proposition 8.19.
Si $f\colon \mathcal{E} \to \mathcal{E}'$ et $g\colon \mathcal{E} \to \mathcal{E}''$ sont des applications affines, alors $g \circ f\colon \mathcal{E} \to \mathcal{E}''$ est affine et $u_{g\circ f} = u_g \circ u_f$.

Démonstration. Si $M \in \mathcal{E}$ et $x \in E$ nous avons

$$\begin{aligned}(g \circ f)(M + x) &= g\big(f(M) + u_f(x)\big)\\ &= g\big(f(M)\big) + u_g\big(u_f(x)\big)\\ &= (g \circ f)(M) + (u_g \circ u_f)(x).\end{aligned} \tag{8.42}$$

□

Théorème 8.20.
Soient $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines de dimensions finies p et q sur $\mathbb{K}$. Soient les repères cartésiens $R = (O, \{e_i\})$ et $R' = (O', \{e'_i\})$. Une application $f\colon \mathcal{E} \to \mathcal{E}'$ est affine si et seulement si il existe une matrice $a \in \mathrm{M}_{p,q}(\mathbb{K})$ et $b \in \mathbb{K}^q$ tels que

$$f(x) = b + ax. \tag{8.43}$$

Remarque 8.21.
L'équation (8.43) est écrite en utilisant un abus de notation entre le vecteur $x \in \mathbb{K}^p$ et le point de $\mathcal{E}$ qui est représenté par x dans le repère $(A, \{e_i\})$.

8.5 Isomorphismes

Définition 8.22.
*Un **isomorphisme** entre les espaces affines $\mathcal{E}$ et $\mathcal{E}'$ est une application affine $f\colon \mathcal{E} \to \mathcal{E}'$ inversible dont l'inverse est affine.*

Proposition 8.23.
Une application affine bijective est un isomorphisme. Si f est un isomorphisme d'espaces affines, alors $u_{f^{-1}} = (u_f)^{-1}$.

Proposition 8.24.
Un espace affine de dimension finie n sur un corps $\mathbb{K}$ est isomorphe à l'espace affine canonique $\mathcal{E}_n(\mathbb{K})$.

Démonstration. Si nous considérons le repère $R = (A, \{e_i\})$ de l'espace affine $\mathcal{E}$ alors l'application

$$\begin{aligned} \varphi\colon \mathbb{K}^n &\to \mathcal{E} \\ (x_1, \dots, x_n) &\mapsto A + \sum_i x_i e_i \end{aligned} \tag{8.44}$$

est un isomorphisme. □

8.6 Sous espaces affines

Définition 8.25.
Soit $\mathcal{E}$ un espace affine sur l'espace vectoriel E. Un ***sous-espace affine*** *de $\mathcal{E}$ est une orbite de l'action d'un sous-espace vectoriel de E.*

Si $\mathcal{F}$ est un sous-ensemble de $\mathcal{E}$, il sera un sous-espace affine de $\mathcal{E}$ si et seulement si l'ensemble

$$F = \{AB \text{ tel que } A, B \in \mathcal{F}\} \tag{8.45}$$

est un sous-espace vectoriel de E. Dans ce cas nous disons que F est la **direction** de $\mathcal{F}$. Si $A \in \mathcal{F}$, alors l'orbite de A sous F est $\mathcal{F}$. La **dimension** de $\mathcal{F}$ est la dimension de sa direction.

Si $\mathcal{F}$ et $\mathcal{G}$ sont des sous-espaces affines de $\mathcal{E}$ de directions F et G, nous disons que $\mathcal{F}$ est **parallèle** à $\mathcal{G}$ si $F \subset G$.

Proposition 8.26.
Soit $\mathcal{F}$ un sous-espace affine de dimension k dans l'espace affine $\mathcal{E}$ de dimension n. Alors il existe une application affine $f\colon \mathcal{E} \to \mathbb{K}^{n-k}$ telle que $\mathcal{F} = f^{-1}(0)$.

Démonstration. Soient F la direction de $\mathcal{F}$ et $A \in \mathcal{F}$. Nous considérons une base $\{e_i\}$ adaptée à F au sens $\{e_1, \dots, e_k\}$ est une base de F. Nous considérons maintenant le repère cartésien $(A, \{e_i\})$ avec $A \in \mathcal{F}$ et nous construisons l'application affine

$$\begin{aligned} f\colon \mathcal{E} &\to \mathbb{K}^{n-k} \\ A + \sum_{i=1}^n x_i e_i &\mapsto \begin{pmatrix} x_{k+1} \\ \vdots \\ x_n \end{pmatrix}. \end{aligned} \tag{8.46}$$

Par construction nous avons $f(M) = 0$ si et seulement si $M \in \mathcal{F}$. □

Proposition 8.27 ([40])**.**
Soit σ une partie de l'espace affine $\mathcal{E}$.

(1) L'intersection de tous les sous-espaces affines contenant σ est un sous-espace affine, noté $\mathcal{F}$.

(2) Si $A \in \sigma$, alors la direction de $\mathcal{F}$ est le sous-espace vectoriel

$$F = \operatorname{Span}\{\overrightarrow{AM} \text{ tel que } M \in \sigma\}. \tag{8.47}$$

Le sous-espace affine donné par la proposition 8.27 est le sous-espace affine **engendré** par la partie σ, et il est noté $\operatorname{eae}(\sigma)$.

Proposition 8.28.
Soit $\mathcal{E}$ un espace affine de dimension n sur $\mathbb{K}$, soit $f\colon \mathcal{E} \to \mathbb{K}^r$ une fonction affine. Pour tout $a = (a_1, \dots, a_r) \in \mathbb{K}^r$, l'ensemble $f^{-1}(a)$ est un sous-espace affine de dimension $\dim \ker(u_f)$.

Démonstration. Nous considérons le repère $(A, \{e_i\})$ de $\mathcal{E}$. Étant donné que f est affine nous avons

$$f\big(A + \sum_i x_i e_i\big) = f(A) + u_f\big(\sum_i x_i e_i\big). \tag{8.48}$$

Nous avons donc $f\big(A+\sum_i x_i e_i\big)=a$ lorsque

$$u_f(\sum_i x_i e_i)=a-f(A). \tag{8.49}$$

Nous avons donc

$$f^{-1}(a)=A+(u_f)^{-1}\big(a-f(A)\big), \tag{8.50}$$

dont la dimension est le rang de $(u_f)^{-1}=u_{f^{-1}}$ (proposition 8.23). Le rang de $(u_f)^{-1}$ est la dimension du noyau de u_f. □

Proposition 8.29.
Soit un espace vectoriel normé[3] $(V,\|.\|)$. *Pour tout* $a\in V$ *et* $r>0$, *la boule* $B(a,r)$ *est convexe*[4]. *La boule fermée* $\overline{B(a,r)}$ *également.*

Démonstration. En deux parties.

(i) **La boule centrée en zéro** Soient $x,y\in B(0,r)$ et $\lambda\in\,]0,1[$. Alors

$$\|\lambda x+(1-\lambda)y\|\leqslant|\lambda|\|x\|+|1-\lambda|\|y\|<(|\lambda|+|1-\lambda|)r\leqslant r \tag{8.51}$$

où nous avons utilisé le fait que $|\lambda|=\lambda$ et $|1-\lambda|=1-\lambda$.
Cela prouve que $\lambda x+(1-\lambda)y\in B(0,r)$. Notez l'inégalité stricte due au fait que $\|x\|<r$ et $\|y\|<r$. Dans le cas de la boule fermée, nous avons une inégalité large.

(ii) **La boule centrée autre part** Soient $x,y\in B(a,r)$. Alors $x-a$ et $y-a$ sont dans $B(0,r)$, de telle sorte que

$$\lambda(x-a)+(1-\lambda)(y-a)\in B(0,r) \tag{8.52}$$

par la première partie. En développant et simplifiant,

$$\lambda x+(1-\lambda)y-a\in B(0,r), \tag{8.53}$$

ce qui signifie que $\lambda x+(1-\lambda)y\in B(a,r)$.

□

Proposition 8.30.
Soit A *un ensemble convexe*[5] *dans un espace vectoriel et* $v_1,\dots,v_n$ *des éléments de* A. *Alors toute combinaison*

$$a_1v_1+\cdots+a_nv_n \tag{8.54}$$

telle que $a_1+\cdots+a_n=1$ *et* $a_i\in[0,1]$ *appartient à* A.

Démonstration. Nous prouvons la proposition pour $n=3$. Nous devons trouver des nombres $t_1,t_2\in[0,1]$ tels que

$$t_2\big(t_1v_1+(1-t_1)v_2\big)+(1-t_2)v_3=av_1+bv_2+cv_3. \tag{8.55}$$

La réponse est immédiatement donnée par

$$t_2a=1-c \tag{8.56a}$$
$$t_1=a/t_2. \tag{8.56b}$$

Étant donné que $c\in[0,1]$ nous avons $t_2\in[0,1]$. En ce qui concerne t_1 nous avons

$$t_1=\frac{a}{t_2}\leqslant\frac{1-c}{1-c}=1. \tag{8.57}$$

□

3. Définition 7.136.
4. Définition 7.134.
5. Définition 7.134.

8.7 Barycentre

Soit $\mathcal{E}$ un espace affine sur le $\mathbb{K}$-espace vectoriel E. Un couple (A, λ) avec $A \in \mathcal{E}$ et $\lambda \in \mathbb{K}$ est un **point pondéré**.

Lemme-Définition 8.31 ([217]).
Soit une famille de points pondérés $\{(A_i, \lambda_i)\}_{i=1\ldots r}$. Si $\sum_i \lambda_i \neq 0$, alors il existe un unique $G \in \mathcal{E}$ tel que

$$\sum_{i=1}^{r} \lambda_i \overrightarrow{GA_i} = 0. \tag{8.58}$$

*Le point G donné par le lemme 8.31 est le **barycentre** des points pondérés (A_i, λ_i).*

Notons que l'on peut toujours supposer que $\sum_i \lambda_i = 1$ parce que le barycentre ne change pas lorsque tous les λ_i sont multipliés par un même nombre.

Définition 8.32 (Combinaison convexe).
*Des nombres $\lambda_1, \ldots, \lambda_n$ vérifiant $\sum_i \lambda_i = 1$ forment une **combinaison convexe**.*

Le théorème suivant donne quelques caractérisations équivalentes du barycentre.

Théorème 8.33 ([217]).
Soient $\{(A_i, \lambda_i)\}_{i=1,\ldots,r}$ une famille de points pondérés. Les conditions suivantes sur le point $G \in \mathcal{E}$ sont équivalentes.

(1) Le point G est le barycentre de la famille.
(2) Pour tout $\alpha \in \mathbb{R}^*$, $\sum_i (\alpha\lambda_i)\overrightarrow{GA_i} = 0$.
(3) Il existe $A \in \mathcal{E}$ tel que $\left(\sum_i \lambda_i\right)\overrightarrow{AG} = \sum_i \lambda_i \overrightarrow{AA_i}$.
(4) Pour tout $B \in \mathcal{E}$, nous avons $\left(\sum_i \lambda_i\right)\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}$.

Définition 8.34.
*Si $A, B \in \mathcal{E}$, le **segment** $[AB]$ est l'ensemble des barycentres de A et B pondérés par des poids positifs (ouvert ou fermé suivant que l'on accepte que l'un ou l'autre des poids soit nul).*

Lorsque tous les λ_i sont égaux, nous parlons d'**isobarycentre**. Autrement dit, l'isobarycentre des points A_i est le barycentre des points pondérés $(A_i, 1)$.

8.7.1 Sous-espaces affines

Proposition 8.35.
Une partie $\mathcal{F}$ de $\mathcal{E}$ est un sous-espace affine si et seulement si elle est stable par barycentrisation.

Démonstration. Soit $\mathcal{F}$ un sous-espace affine de direction F et $A_1, \ldots, A_n$ des points de $\mathcal{F}$. Nous devons voir que le barycentre des points A_i pondérés de n'importe quelles masses appartient à $\mathcal{F}$. Pour ce faire nous faisons appel à la caractérisation (4) du théorème 8.33 : pour tout $B \in \mathcal{F}$,

$$\overrightarrow{BG} = \sum_i \lambda_i \overrightarrow{BA_i}. \tag{8.59}$$

Puisque B et A_i sont dans $\mathcal{F}$, nous avons $\overrightarrow{BA_i} \in F$ et donc $\overrightarrow{BG} \in F$. Mais comme $B \in \mathcal{F}$, le point G est à son tour dans $\mathcal{F}$.

Réciproquement, nous supposons que $\mathcal{F}$ est stable par barycentrisme. Nous voudrions montrer que l'ensemble

$$F = \{\overrightarrow{AB} \text{ tel que } A, B \in \mathcal{F}\} \tag{8.60}$$

est un sous-espace vectoriel. Soit $A \in \mathcal{F}$. Nous commençons par prouver que les vecteurs de la forme $\overrightarrow{AX}$ $(X \in \mathcal{F})$ forment un espace vectoriel. Considérons $\overrightarrow{AX} + \overrightarrow{AY}$ qui est un élément de E ; il existe donc $V \in \mathcal{E}$ tel que

$$\overrightarrow{AV} = \overrightarrow{AX} + \overrightarrow{AY}. \tag{8.61}$$

Par les relations de Chasles,

$$\overrightarrow{AV} = \overrightarrow{AV} + \overrightarrow{VX} + \overrightarrow{AV} + \overrightarrow{VY}, \tag{8.62}$$

donc

$$0 = \overrightarrow{VX} - \overrightarrow{VA} + \overrightarrow{VY}, \tag{8.63}$$

ce qui prouve que V est un barycentre de X, A, Y, et donc que $V \in \mathcal{F}$. De la même manière si $W \in \mathcal{E}$ est défini par $\overrightarrow{AW} = \mu\overrightarrow{AX}$, alors

$$\overrightarrow{AW} = \mu\overrightarrow{AX} = \mu(\overrightarrow{AW} + \overrightarrow{WX}), \tag{8.64}$$

ce qui signifie que

$$(1-\mu)\overrightarrow{AW} + \mu\overrightarrow{XW} = 0 \tag{8.65}$$

et que W est un barycentre.

Afin de montrer que (8.60) est bien un espace vectoriel, nous devons considérer $A, B, X, Y \in \mathcal{F}$ et prouver que $\overrightarrow{AX} + \overrightarrow{BY} \in F$. Nous avons

$$\begin{aligned}
\overrightarrow{AX} + \overrightarrow{BY} &= \overrightarrow{AX} + \overrightarrow{BA} + \overrightarrow{AY} && \text{(8.66a)}\\
&= \overrightarrow{AV} + \overrightarrow{BA} && V \text{ est celui donné plus haut} \quad \text{(8.66b)}\\
&= \overrightarrow{AV} - \overrightarrow{AB} && \text{(8.66c)}\\
&= \overrightarrow{AV} + \overrightarrow{AW} && W \text{ est donné par } \mu = -1. \quad \text{(8.66d)}\\
&= \overrightarrow{AV'}. && \text{(8.66e)}
\end{aligned}$$

□

Proposition 8.36 ([217]).
Soient $A_0, \dots, A_r$ des points de $\mathcal{E}$. L'ensemble des barycentres de ces points (avec des masses de somme 1) est le sous-espace affine engendré par les A_i que nous nommons $\mathcal{F}$.

Démonstration. Soit G le barycentre associé aux poids λ_i. Nous avons

$$G = A_0 + \overrightarrow{A_0G} = A_0 + \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i}. \tag{8.67}$$

Notons que les vecteurs $\overrightarrow{A_0A_i}$ sont dans la direction du sous-espace affine engendré par les A_i par (8.47). Donc G est bien dans $\mathcal{F}$.

Inversement si X est dans $\mathcal{F}$, on a

$$X = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i} \tag{8.68}$$

parce que $\sum_i \lambda_i \overrightarrow{A_0A_i}$ est un élément général de la direction de $\mathcal{F}$. Donc

$$\overrightarrow{A_0X} = \sum_i \lambda_i \overrightarrow{A_0A_i}, \tag{8.69}$$

et en utilisant la relation de Chasles sur chacun des $\overrightarrow{A_0A_i}$,

$$\overrightarrow{A_0X} = \sum_i \lambda_i \big(\overrightarrow{A_0X} + \overrightarrow{XA_i}\big). \tag{8.70}$$

De là nous concluons que

$$\big(1 - \sum_i \lambda_i\big)\overrightarrow{A_0X} + \sum_i \lambda_i \overrightarrow{A_iX} = 0, \tag{8.71}$$

ce qui signifie précisément que X est un barycentre des A_i. □

Proposition 8.37.
Soient $r+1$ point $A_0, \dots, A_r$ dans $\mathcal{E}$. Le sous-espace affine engendré par les A_i est au plus de dimension r.

Démonstration. La direction de l'espace engendré $\mathrm{Aff}\{A_i\}$ est l'espace

$$\mathrm{Span}\{\overrightarrow{A_0A_i}_{i=1,\dots,r}\} \tag{8.72}$$

qui est engendré par r vecteurs et donc est au plus de dimension r. □

En deux mots, la proposition suivante signifie que le barycentre des barycentres est le barycentre.

Proposition 8.38 (Associativité des barycentres[218])**.**
Soit $I = \{0, 1, \dots, n\}$ et une partition $I = J_0 \cup \dots \cup J_r$. Soient des points $a_0, \dots, a_n \in \mathcal{E}$ et $\lambda_0, \dots, \lambda_m$ des nombres tels que $\sum_i \lambda_i \neq 0$. Nous supposons que $\mu_k = \sum_{i\in J_k} \lambda_i \neq 0$ pour tout k, et enfin nous nommons b_k le barycentre de la famille $\{(a_i, \lambda_i), i \in J_k\}$.

Alors le barycentre de la famille $\{(b_k, \mu_k)\}_{k=1,\dots,n}$ est le barycentre de la famille $\{(a_i, \lambda_i)\}_{i\in I}$.

Démonstration. Nous nommons b le barycentre des b_k pondérés par les μ_k, donc par définition

$$\begin{aligned}
0 &= \sum_{k=0}^{r} \mu_k \overrightarrow{bb_k} && (8.73a)\\
&= \sum_k \sum_{i\in J_k} \lambda_i \overrightarrow{bb_k} && (8.73b)\\
&= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i (\overrightarrow{ba_i} + \overrightarrow{a_ib_k}) && (8.73c)\\
&= \sum_{k=0}^{r} \sum_{i\in J_k} \lambda_i \overrightarrow{ba_i} + \sum_{k=0}^{r} \underbrace{\sum_{i\in J_k} \lambda_i \overrightarrow{a_ib_k}}_{=0} && (8.73d)\\
&= \sum_{i\in I} \lambda_i \overrightarrow{ba_i}. && (8.73e)
\end{aligned}$$

Donc b est bien barycentre des a_i avec les poids λ_i. □

8.7.2 Enveloppe convexe

Définition 8.39.
*Soit A une partie d'un espace vectoriel E. L'**enveloppe convexe** de A, notée $\mathcal{C}\mathrm{onv}(A)$ est l'intersection de tous les convexes contenant A.*

L'enveloppe convexe est un convexe. En effet soit C un convexe contenant A et $x, y \in \mathcal{C}\mathrm{onv}(A)$; alors x et y sont dans C et par conséquent le segment $[x, y]$ est inclus dans C. Ce segment étant inclus dans tout convexe contenant A, il est inclus dans $\mathcal{C}\mathrm{onv}(A)$.

Proposition 8.40 ([219])**.**
Soit C un convexe dans l'espace affine $\mathcal{E}$ et une famille de points pondérés $\{(a_i, \lambda_i)\}_{i=1,\dots,r}$ dont tous les poids sont positifs (et non tous nuls). Alors le barycentre est aussi dans C.

En d'autre termes, un convexe est stable par barycentrage à poids positifs[6]*.*

Démonstration. Nous prouvons par récurrence. D'abord pour $r = 2$. Le barycentre des points pondérés (a_1, λ_1), (a_2, λ_2) est le point b tel que

$$\lambda_1 \overrightarrow{ba_1} + \lambda_2 \overrightarrow{ba_2} = 0. \tag{8.74}$$

6. Sauf si on prend tous les poids nuls ; mais contre ce genre d'idées, on ne peut rien faire.

Par définition, ce qui est noté $\overrightarrow{ab}$ n'est rien d'autre que $b-a$; en déballant (8.74), nous trouvons

$$\lambda_1(a_1-b)+\lambda_2(a_2-b)=0 \tag{8.75}$$

et donc

$$b=\frac{\lambda_1}{\lambda_1+\lambda_2}a_1+\frac{\lambda_2}{\lambda_1+\lambda_2}a_2, \tag{8.76}$$

qui est bien un point du segment $[a_1,a_2]$ parce que c'est une combinaison à coefficients positifs de somme 1.

Nous passons maintenant à la vraie récurrence avec un ensemble de points pondérés

$$A_r=\{(a_1,\lambda_1),\dots,(a_r,\lambda_r)\} \tag{8.77}$$

de masse totale non nulle ; et en vous laissant deviner ce que va désigner A_{r-1}. Si une des masses est nulle (disons λ_r), alors le barycentre de A_r est le même que celui de A_{r-1} et l'hypothèse de récurrence nous enseigne que ledit barycentre est dans C. Nous supposons donc que $\lambda_i\neq 0$ pour tout i. Dans ce cas le théorème d'associativité des barycentres 8.38 dit que le barycentre de A_r est le barycentre entre le barycentre de A_{r-1} et (a_r,λ_r), qui sont deux points de C par hypothèse de récurrence. □

Si E est un espace vectoriel et si $x_i\in E$ et $\lambda_i\in\mathbb{R}$, alors le barycentre des couples (x_i,λ_i) est le point g tel que $\sum_i\lambda_i\overrightarrow{gx_i}$, c'est-à-dire $\sum_i\lambda_i(x_i-g)=0$ ou encore

$$\sum_i\lambda_i x_i=\sum_i\lambda_i g. \tag{8.78}$$

Donc, quitte à diviser tous les λ_i par la somme, nous pouvons supposer que la somme des poids est 1. C'est pourquoi lorsque nous parlerons de barycentre dans un espace vectoriel sans contexte affine, nous allons toujours supposer $\sum_i\lambda_i=0$ et avoir le barycentre

$$g=\sum_i\lambda_i x_i. \tag{8.79}$$

Proposition 8.41.
Soit E, un espace vectoriel et $A\subset E$. L'enveloppe convexe $\mathcal{C}\mathrm{onv}(A)$ est l'ensemble des barycentres de familles finies de points affublés de masses positives.

Démonstration. Nous notons $\mathcal{B}$ l'ensemble des dits barycentres. Par la proposition 8.40, ces barycentres sont dans l'enveloppe convexe et donc $\mathcal{B}\subset\mathcal{C}\mathrm{onv}(A)$. A contrario, si nous prouvons que $\mathcal{B}$ était convexe, alors nous aurions $\mathcal{C}\mathrm{onv}(A)\subset\mathcal{B}$ parce que l'enveloppe convexe est l'intersection des convexes contenant A.

Soient $a,b\in\mathcal{B}$, c'est-à-dire que l'on a $a_0,\dots,a_n$ et $b_0,\dots,b_m$ dans A ainsi que les nombres strictement positifs $\lambda_0,\dots,\lambda_n$ et $\mu_0,\dots,\mu_m$ tels que

$$a=\sum_i\lambda_i a_i \qquad \sum_{i=1}^n\lambda_i=1 \tag{8.80a}$$

$$b=\sum_j\mu_j b_j \qquad \sum_{j=1}^n\mu_j=1 \tag{8.80b}$$

Un point du segment $[a,b]$ est de la forme $p=ta+(1-t)b$ avec $t\in[0,1]$. En développant,

$$p=\sum_{i=0}^n(t\lambda_i)a_i+\sum_{j=0}^m(1-t)\mu_j b_j. \tag{8.81}$$

C'est le barycentre de la famille $\{(a_i,\lambda_i),(b_j,\mu_j)\}$, parce que la somme des coefficients vaut bien 1 :

$$\sum_i(t\lambda_i)+\sum_j(1-t)\mu_j=t+(1-t)=1. \tag{8.82}$$

□

Théorème 8.42 (Carathéodory[89]).
Dans un espace affine de dimension n, l'enveloppe convexe[7] *de A est l'ensemble des barycentres à coefficients positifs ou nuls de familles de $n+1$ points.*

Démonstration. Soit $x \in \mathcal{C}\mathrm{onv}(A)$; on sait par la proposition 8.41 que x est barycentre de points de A avec des coefficients positifs :

$$x = \sum_{k=1}^{p} \lambda_k x_k \tag{8.83}$$

avec $\sum_k \lambda_k = 1$. Nous supposons que $p > n+1$ (sinon le théorème est réglé), et nous allons faire une récurrence à l'envers en montrant qu'on peut aussi écrire x sous forme d'un barycentre de strictement moins de p points.

Étant donné que $p-1 > n$, la famille $\{x+i-x_1\}_{i=2,\dots,p}$ est liée et il existe donc $\alpha_1, \dots, \alpha_p \in \mathbb{R}$ tels que $\sum_{i=2}^{p} \alpha_i(x_i - x_1) = 0$, c'est-à-dire telle que

$$\sum_{i=2}^{p} \alpha_i x_i = \sum_{i=2}^{p} \alpha_i x_1. \tag{8.84}$$

Nous posons $\alpha_1 = -\sum_{i=2}^{p} \alpha_1$. Remarquons qu'alors $\sum_{i=1}^{p} \alpha_i x_i = 0$ parce que

$$\sum_{i=1}^{p} \alpha_i x_i = \alpha_1 x_1 + \sum_{i=2}^{p} \alpha_i x_i = \alpha_1 x_1 + \sum_{i=2}^{p} \alpha_i x_1 = \sum_{i=1}^{p} \alpha_i x_1 = 0. \tag{8.85}$$

Par conséquent ça ne coûte rien de récrire (8.83) sous la forme

$$x = \sum_{i=1}^{p} (\lambda_i + t\alpha_i) x_i. \tag{8.86}$$

Les α_i ne sont pas tous nuls, mais leur somme est nulle, donc il y en a au moins un négatif. Nous notons

$$\tau = \min\{-\frac{\lambda_i}{\alpha_i} \text{ tel que } \alpha_i < 0\}, \tag{8.87}$$

et J l'ensemble de i pour lesquels ce minimum est atteint. Nous considérons aussi les nombres $\mu_i = \lambda_i + \tau\alpha_i$. Plusieurs remarques.

(1) Si $j \in J$, alors $\mu_j = 0$

(2) Si $\alpha_i > 0$ alors $\mu_i \geqslant 0$, mais si $\alpha_i < 0$ alors

$$\lambda_i + \tau\alpha_i \geqslant \lambda_i + (-\frac{\lambda_i}{\alpha_i})\alpha_i = 0m \tag{8.88}$$

donc $\mu_i \geqslant 0$ quand même.

(3) $\sum_{i=1}^{p} \mu_i = 1$, toujours parce que $\sum_{i=1}^{p} \alpha_i = 0$.

Avec tout ça, nous avons

$$\sum_{i \notin J} \mu_i x_i = \sum_{i=1}^{p} \mu_i x_i = x. \tag{8.89}$$

Et voilà, nous avons écrit x comme un barycentre à coefficients positifs de moins de p éléments parce que J n'est pas vide. □

Corolaire 8.43.
Dans un espace affine de dimension finie, l'enveloppe convexe d'un compact est compacte.

7. Définition 8.39.

Démonstration. Soit A une partie compacte de l'espace vectoriel E, et $\mathcal{C}\mathrm{onv}(A)$ son enveloppe convexe. Nous allons montrer que toute suite dans $\mathcal{C}\mathrm{onv}(A)$ admet une sous-suite convergente en écrivant un point de $\mathcal{C}\mathrm{onv}(A)$ comme le théorème de Carathéodory 8.42 nous le suggère. Pour cela nous considérons le simplexe

$$\Lambda = \left\{ \lambda \in \mathbb{R}^{n+1} \text{ tel que } \sum_{k=1}^{n+1} \lambda_k = 1 \text{ et } \lambda_k \geqslant 0 \forall k \right\}. \tag{8.90}$$

Montrons en passant que Λ est compact. Si $\lambda_k \in \Lambda$ est une suite, alors chacun des λ_k est un $(n+1)$-uple de nombres dans $[0,1]$:

$$k \mapsto (\lambda_k)_i \tag{8.91}$$

est une suite qui possède une sous-suite convergente. En passant $n+1$ fois à une sous-suite, nous tombons sur une suite convergente vers $\lambda \in \Lambda$, grâce à la convergence composante par composante. De plus pour chaque k nous avons $\sum_{i=1}^{n+1} (\lambda_k)_i = 1$, et en passant à la limite, la somme étant une application continue, $\sum_i \lambda_i = 1$.

Considérons l'application

$$\begin{aligned} f\colon \Lambda \times A^{n+1} &\to \mathcal{C}\mathrm{onv}(A) \\ (\lambda, x) &\mapsto \sum_{k=1}^{n+1} \lambda_k x_k. \end{aligned} \tag{8.92}$$

C'est une application continue parce qu'elle est bilinéaire en dimension finie ; son image est contenue dans $\mathcal{C}\mathrm{onv}(A)$ par la proposition 8.40, et elle est surjective par le théorème de Carathéodory 8.42. Bref, $\mathcal{C}\mathrm{onv}(A) = f(\Lambda \times A^{n+1})$ est donc l'image d'un compact par une application continue ; elle est donc compacte par le théorème 7.186. □

Notons que sans le théorème de Carathéodory, peut être que le nombre de points utiles pour décomposer les différents a_k n'était pas borné ; dans ce cas nous aurions du prendre une infinité de sous-suites et rien n'aurait été sûr.

8.7.3 Applications affines et barycentre

Proposition 8.44 ([220]).
Une application $f\colon \mathcal{E} \to \mathcal{E}'$ entre deux espaces affines est affine si et seulement si pour tout système $\{(A_i, \lambda_i)\}_{i=1,\dots,k}$ de barycentre G et de poids total non nul, le point $f(G)$ est barycentre du système $\{\big(f(A_i), \lambda_i\big)\}$.

Démonstration. En deux parties.

(i) **Si f est affine** Par définition d'un barycentre,

$$\sum_i \lambda_i \overrightarrow{GA_i} = 0. \tag{8.93}$$

Nous considérons un point arbitraire $O \in \mathcal{E}$ et nous écrivons $A_i = O + x_i$, $G = O + x_g$. Ensuite nous utilisons le lemme 8.14 pour le calcul suivant :

$$\begin{aligned} \sum_i \lambda_i \overrightarrow{f(G)f(A_i)} &= \sum_i \lambda_i u_f(x_i - x_g) && \text{(8.94a)} \\ &= u_f\big(\sum_i \lambda_i (x_i - x_g)\big) && \text{(8.94b)} \\ &= u_f\big(\sum_i \lambda_i \overrightarrow{GA_i}\big) && \text{(8.94c)} \\ &= u_f(0) = 0. && \text{(8.94d)} \end{aligned}$$

Donc $f(G)$ est bien le barycentre du nouveau système.

(ii) **Si f conserve les barycentres** Nous définissons u par $f(O+x) = f(O) + u(x)$. À priori, ce u dépend de O et n'est pas linéaire.

(i) **u est linéaire** Soient $M, N \in \mathcal{E}$ et les éléments $x_m, x_n \in E$ tels que $\overrightarrow{OM} = x_m$ et $\overrightarrow{ON} = x_n$. Nous définissons enfin P par

$$\overrightarrow{OP} = \alpha\overrightarrow{OM} + \beta\overrightarrow{ON}, \tag{8.95}$$

et $P = O + x_p$. En décomposant $\overrightarrow{MO}$ et $\overrightarrow{NO}$ par les relations de Chasles de la proposition 8.5(1) nous avons

$$(\alpha + \beta - 1)\overrightarrow{PO} - \alpha\overrightarrow{PM} - \beta\overrightarrow{PN} \tag{8.96}$$

et donc P est barycentre du système

$$\{(\alpha + \beta - 1, O), (\alpha, M), (\beta, N)\}. \tag{8.97}$$

Le point $f(P)$ sera barycentre du système

$$\Big\{\big(\alpha + \beta - 1, f(O)\big), \big(\alpha, f(M)\big), \big(\beta, f(N)\big)\Big\}. \tag{8.98}$$

Cela signifie que

$$(\alpha + \beta - 1)\overrightarrow{f(P)f(O)} - \alpha\overrightarrow{f(P)f(M)} - \beta\overrightarrow{f(P)f(N)} = 0. \tag{8.99}$$

En y substituant $\overrightarrow{f(P)f(O)} = u(-x_p)$, $\overrightarrow{f(P)f(M)} = u(x_m - x_p)$ et $\overrightarrow{f(P)f(N)} = u(x_m m - x_p)$ ainsi que $x_p = \alpha x_m + \beta x_b$ nous trouvons

$$u(\alpha x_m + \beta x_n) = \alpha u(x_m) + \beta u(x_n). \tag{8.100}$$

Donc u est linéaire.

(ii) **u ne dépend pas du point O** Il n'est pas besoin de démontrer cela parce que la définition 8.12 ne le demande pas. Note : c'est le lemme 8.13 qui dit que c'est par ailleurs vrai.

□

8.8 Repères, coordonnées cartésiennes et barycentriques

Définition 8.45.
*On dit que les points $A_0, \dots, A_r \in \mathcal{E}$ sont **affinement indépendants** si le sous-espace affine engendré est de dimension r.*

Proposition 8.46 ([217]).
Pour $r+1$ points $A_0, \dots, A_r$ dans $\mathcal{E}$, les propriétés suivantes sont équivalentes.

(1) Les A_i sont affinement indépendants.

(2) Pour tout $i = 0, \dots, r$, le point A_i n'est pas dans $\mathrm{Aff}\{A_0, \dots, A_i, \dots, A_r\}$.

(3) Les points $A_0, \dots, A_{r-1}$ sont affinement indépendants et $A_r \notin \mathrm{Aff}\{A_0, \dots, A_{r+1}\}$.

(4) Il existe i tel que les vecteurs $\overrightarrow{A_kA_i}$ $(k \neq i)$ sont linéairement indépendants.

(5) Pour tout $i \in \{1, \dots, r\}$, les vecteurs $\overrightarrow{A_kA_i}$ $(k \neq i)$ sont linéairement indépendants.

Notons à propos de la condition (3) que l'existence d'un i tel que $A_i \notin \mathrm{Aff}\{A_0, \dots, A_i, \dots, A_r\}$ n'implique pas l'indépendance des $r+1$ points. En effet dans $\mathbb{R}^2$ nous considérons les 4 points $A_0 = (0,0)$, $A_1 = (1,0)$, $A_2 = (2,0)$ et $A_3 = (0,1)$. Évidemment le point A_3 n'est pas dans l'espace engendré par les trois autres ; il n'empêche que ces points ne sont pas affinement indépendants parce que la direction est de dimension 2 au lieu de 3.

Définition 8.47.
Soit $\mathcal{E}$ un espace affine de dimension n et $\mathcal{F}$ un sous-espace affine de dimension k. Un ***repère affine*** *de $\mathcal{F}$ est la donnée de $k+1$ points affinement indépendants de $\mathcal{F}$.*

Si $\{A_0, \dots, A_n\}$ est un repère affine, le point A_0 est l'**origine**. C'est un choix complètement arbitraire ; et c'est bien cet arbitraire qui nous amènera à considérer les coordonnées barycentriques au lieu des coordonnées cartésiennes.

Soit $M \in \mathcal{E}$; par définition nous avons

$$M = A_0 + \overrightarrow{A_0M}. \tag{8.101}$$

Mais nous savons que les vecteurs $\overrightarrow{A_0A_i}$ forment une base de E, nous avons donc des nombres λ_i tels que

$$\overrightarrow{A_0M} = \sum_{i=1}^{n} \lambda_i \overrightarrow{A_0A_i}. \tag{8.102}$$

Les nombres λ_i ainsi construits sont les **coordonnées cartésiennes** du point M dans le repère $\{A_0, \dots, A_n\}$ d'origine A_0.

À partir de ces coordonnées, le point $M \in \mathcal{E}$ se retrouve par la formule

$$M = A_0 + \sum_{i=1}^{n} \lambda_i \overrightarrow{A_0A_i}. \tag{8.103}$$

Proposition 8.48 ([1])**.**
La paire $\big(O, \{e_1, \dots, e_n\}\big)$ est un repère cartésien de $\mathcal{E}$ si et seulement si $\{O, O+e_1, \dots, O+e_n\}$ est un repère affine.

Démonstration. En deux parties.

(i) <u>**Sens direct**</u> Vue la proposition 8.46, il suffit de prouver que les vecteurs $\overrightarrow{O(O+e_i)}$ sont linéairement indépendants. Mais $\overrightarrow{O(O+e_i)} = e_i$, donc oui, ils sont linéairement indépendants.

(ii) <u>**Sens inverse**</u> Il s'agit d'utiliser la même proposition 8.46 qui est encore applicable parce que c'est une équivalence.

□

Soient (A, e_i) et (A', e'_i) deux repères cartésiens pour l'espace affine $\mathcal{E}$. Soit (a_{ij}) la matrice de changement de base entre $\{e_i\}$ et $\{e'_i\}$ dans E. Nous voudrions trouver les x_i en termes des x'_i.

Pour cela nous considérons un point M dans $\mathcal{E}$ et nous l'écrivons dans les deux bases. Cela fournit l'égalité

$$A + \sum_i x_i e_i = A' + \sum_i x'_i e'_i. \tag{8.104}$$

Nous considérons les coordonnées (a_i) de A' dans le repère (A, e_i), c'est-à-dire

$$A' = A + \sum_i a_i e_i. \tag{8.105}$$

En substituant $e'_i = \sum_k a_{jk} e_k$ et (8.105) dans (8.104) nous trouvons

$$\sum_k x_k e_k = \sum_k a_k e_k + \sum_{jk} a_{jk} x'_j e_k, \tag{8.106}$$

et par conséquent

$$x_k = a_k + \sum_j a_{jk} x'_j. \tag{8.107}$$

Les coordonnées barycentriques sont données par la proposition suivante.

Proposition 8.49 ([217]).
Soient $A_0, \dots, A_r$ des points affinement indépendants dans $\mathcal{E}$ et $\mathcal{F} = \mathrm{Aff}\{A_0, \dots, A_r\}$. Tout point $M \in \mathcal{F}$ s'écrit de façon unique comme barycentre[8] *des A_i affectés de poids λ_i tels que $\sum_{i=0}^{r} \lambda_i = 1$.*

Démonstration. Nous avons vu plus haut (définition 8.47) que l'affine indépendance des points A_i assurait que $(A_0, \dots, A_r)$ était un repère de $\mathcal{F}$.

En ce qui concerne l'existence de l'écriture de M comme barycentre, nous savons que les sous-espace affines sont exactement les ensembles de barycentres (proposition 8.36), c'est-à-dire que si on a des points dans un sous-espace affine, alors les barycentres de ces points est encore dans le sous-espace affine.

L'unicité est comme suit. Si M est barycentre des A_i avec poids λ_i, nous écrivons la caractérisation (4) du théorème 8.33 avec $B = A_0$:

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} \tag{8.108}$$

où la somme à droite s'étend à priori de 0 à r, mais comme $\overrightarrow{A_0A_0} = 0$, nous l'avons limitée à 1. Si M s'écrit comme barycentre de deux façons différentes, nous aurions

$$\overrightarrow{A_0M} = \sum_{i=1}^{r} \lambda_i \overrightarrow{A_0A_i} = \sum_{i=1}^{r} \mu_i \overrightarrow{A_0A_i} \tag{8.109}$$

avec $\sum_i \lambda_i = \sum_i \mu_i = 1$. Étant donné que les points $A_0, \dots, A_r$ forment un repère, les vecteurs $\overrightarrow{A_0A_i}$ sont linéairement indépendants (point (5) de la proposition 8.46) et donc $\lambda_i = \mu_i$ pour $i = 1, \dots, r$. La condition de somme des poids égale à 1 impose alors immédiatement $\lambda_0 = \mu_0$. □

Définition 8.50.
Soit un espace affine $\mathcal{E}$ de dimension n. Soient des points affinement indépendants $A_1, \dots, A_n$. Pour $M \in \mathcal{E}$, la proposition 8.49 indique qu'il existe un unique choix de λ_i tel que

$$\begin{cases} \sum_i \lambda_i = 1 & (8.110a) \\ \sum_i \lambda_i \overrightarrow{MA_i} = 0. & (8.110b) \end{cases}$$

Ces λ_i sont les ***coordonnées barycentriques*** *de M dans le repère $\{A_i\}_{i=1,\dots,n}$.*

8.51.
Soit $\mathbb{R}^2$ et les points non alignés A, B, C. Les coordonnées barycentriques (α, β, γ) dans ce système correspondent à l'unique $X \in \mathbb{R}^2$ tel que

$$\alpha\overrightarrow{XA} + \beta\overrightarrow{XB} + \gamma\overrightarrow{XC} = 0. \tag{8.111}$$

Exemple 8.52.
Soient les points $A = (3, 1)$, $B = (-1, 2)$ et $C = (0, -1)$ dans $\mathbb{R}^2$. Nous allons montrer qu'il forment un repère affine de $\mathbb{R}^2$. L'espace engendré par ces trois points est l'espace des

$$A + \alpha\overrightarrow{AB} + \beta\overrightarrow{AC}, \tag{8.112}$$

et la direction correspondante est l'espace vectoriel donné par $\alpha\overrightarrow{AB} + \beta\overrightarrow{AC}$ qui est de dimension deux. Donc l'espace affine engendré par A, B et C est de dimension 2. △

Exemple 8.53.
Dans le repère (A, B, C), quel est le point de coordonnées barycentriques $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$? D'abord nous vérifions que

$$\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1. \tag{8.113}$$

8. Définition 8.31.

Ensuite nous cherchons $X \in \mathbb{R}^2$ tel que

$$\frac{1}{6}\overrightarrow{AX} + \frac{1}{3}\overrightarrow{BX} + \frac{1}{2}\overrightarrow{CX} = 0, \tag{8.114}$$

c'est-à-dire

$$\frac{1}{6}\begin{pmatrix} x-3 \\ y-1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} x+1 \\ y-2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x \\ y+1 \end{pmatrix} = 0. \tag{8.115}$$

Nous trouvons immédiatement $x = 1/6$ et $y = 1/3$. Le point cherché est donc le point $\begin{pmatrix} 1/6 \\ 1/3 \end{pmatrix}$. △

Lemme 8.54 ([1])**.**
Une application affine $f\colon \mathcal{E} \to \mathcal{E}$ qui préserve les points d'une base affine de $\mathcal{E}$ est l'identité.

Démonstration. Une base affine de $\mathcal{E}$ consiste en $n+1$ points $\{A_0, \dots, A_n\}$ affinement indépendants. Nous utilisons la proposition 8.48 pour dire que $\big(A_0, \{\overrightarrow{A_0A_i}\}_{i=1,\dots,n}\big)$ est un repère cartésien.

En utilisant la formule du lemme 8.13,

$$f(A_i) = f(A_0 + \overrightarrow{A_0A_i}) = f(A_0) + u(\overrightarrow{A_0A_i}). \tag{8.116}$$

Donc $A_i = A_0 + u(\overrightarrow{A_0A_i})$, ce qui signifie que

$$u(\overrightarrow{A_0A_i}) = \overrightarrow{A_0A_i} \tag{8.117}$$

Par ailleurs, tout point M [9] de $\mathcal{E}$ peut être écrit sous la forme

$$M = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i}. \tag{8.118}$$

En appliquant f, et en utilisant (8.117),

$$f(M) = f(A_0) + \sum_i \lambda_i u(\overrightarrow{A_0A_i}) = A_0 + \sum_i \lambda_i \overrightarrow{A_0A_i} = M. \tag{8.119}$$

Donc tout point de $\mathcal{E}$ est fixé par f, ce qui signifie que f est l'identité. □

8.8.1 Équation de droite

Proposition-Définition 8.55.
Soit $\mathcal{E}$ un espace affine de dimension trois muni d'un repère barycentrique [10] *(A_1, A_2, A_3). Nous notons $D(a, b, c)$ l'ensemble des éléments de $\mathcal{E}$ dont les coordonnées barycentriques (normalisées) (x, y, z) vérifient $ax + by + cz = 0$, c'est à dire l'ensemble des $M \in \mathcal{E}$ tels que*

$$\begin{cases} x + y + z = 1 & \text{(8.120a)} \\ x\overrightarrow{MA_1} + y\overrightarrow{MA_2} + \overrightarrow{MA_3} = 0. & \text{(8.120b)} \end{cases}$$

Alors L'ensemble $D(a, b, c)$ est un sous-espace de dimension 1 de $\mathcal{E}$.

*Nous nommons **droite affine** une telle partie de $\mathcal{E}$.*

Idée de preuve : ne pas oublier la condition $x + y + z = 1$ parce que la somme des coordonnées barycentriques doit valoir 1.

Exemple 8.56.
La droite $D(1, 1, 1,)$ n'existe pas parce que ce serait $x + y + z = 0$, qui est incompatible avec $x + y + z = 1$. △

9. Même les points qui ne s'appellent pas « M » en fait.
10. Définition 8.50.

Les droites $D(a,b,c)$ et $D(a',b',c')$ s'intersectent selon les solutions du système

$$\begin{cases} x+y+z=1 & (8.121a)\\ ax+by+cz=0 & (8.121b)\\ a'x+b'y+c'z=0 & (8.121c)\end{cases}$$

Donc deux droites affines ont un unique point d'intersection si et seulement si

$$d=\begin{vmatrix}1 & 1 & 1\\ a & b & c\\ a' & b' & c'\end{vmatrix}\neq 0. \tag{8.122}$$

Elles seront parallèles ou confondues si et seulement si $d=0$.

8.8.2 Associativité, coordonnées barycentriques dans un triangle

Lemme 8.57 ([221])**.**
Soient trois points non alignés A, B, C ainsi que des nombres α, β, γ tels que $\alpha+\beta\neq 0$ et $\alpha+\beta+\gamma\neq 0$.

Soit H le barycentre du système $\{(A,\alpha),(B,\beta)\}$ et G le barycentre de $\{(A,\alpha),(B,\beta),(C,\gamma)\}$.

Alors G est barycentre de $\{(H,\alpha+\beta),(C,\gamma)\}$.

Démonstration. Vues les définitions de H et G nous avons

$$\alpha\overrightarrow{HA}+\beta\overrightarrow{HB}=0 \tag{8.123a}$$
$$\alpha\overrightarrow{GA}+\beta\overrightarrow{GB}+\gamma\overrightarrow{GC}=0. \tag{8.123b}$$

En utilisant les relations de Chasles nous introduisons H dans la seconde relation :

$$\alpha(\overrightarrow{GH}+\overrightarrow{HA})+\beta(\overrightarrow{GH}+\overrightarrow{HB})+\gamma\overrightarrow{GC}=0 \tag{8.124a}$$
$$(\alpha+\beta)\overrightarrow{GH}+\underbrace{\alpha\overrightarrow{HA}+\beta\overrightarrow{HB}}_{=0}+\gamma\overrightarrow{GC}=0 \tag{8.124b}$$
$$(\alpha+\beta)\overrightarrow{GH}+\gamma\overrightarrow{GC}=0. \tag{8.124c}$$

□

Les coordonnées barycentriques dans un triangle (et plus généralement en fait) permettent de faire des projections.

Proposition 8.58.
Soient trois points non alignés A, B, C ainsi qu'un point N de coordonnées barycentriques (α,β,γ) dans le système (A,B,C). Si P est l'intersection $(AN)\cap(BC)$ alors les coordonnées de P sont $(0,\beta,\gamma)$.

Démonstration. Un dessin de la situation :

Dire que les coordonnées de N sont (α, β, γ) signifie que

$$\alpha\overrightarrow{NA} + \beta\overrightarrow{NB} + \gamma\overrightarrow{NC} = 0. \tag{8.125}$$

Nous voudrions montrer que le point P est bien le point de coordonnées $(0, \beta, \gamma)$. Soit donc le point P tel que

$$\beta\overrightarrow{PB} + \gamma\overrightarrow{PC} = 0 \tag{8.126}$$

et montrons que ce point est l'intersection $(BC) \cap (NA)$.

D'abord la relation (8.126) nous dit immédiatement que P est sur la droite (BC). Ensuite, en utilisant les relations de Chasles pour introduire N :

$$\beta(\overrightarrow{PN} + \overrightarrow{NB}) + \gamma(\overrightarrow{PN} + \overrightarrow{NC}) = 0. \tag{8.127}$$

Nous remplaçons $\beta\overrightarrow{NB} + \gamma\overrightarrow{NC}$ par $-\alpha\overrightarrow{NA}$ pour obtenir :

$$(\beta + \gamma)\overrightarrow{PN} - \alpha\overrightarrow{NA} = 0. \tag{8.128}$$

Cela montre que les vecteurs $\overrightarrow{PN}$ et $\overrightarrow{NA}$ sont colinéaires, et donc que P, N et A sont alignés. □

8.9 Applications affines sur $\mathbb{R}^n$

Soit $v \in \mathbb{R}^n$; nous notons $\tau_v \colon \mathbb{R}^n \to \mathbb{R}^n$ la translation donnée par $\tau_v(x) = x + v$. Le groupe de toutes les translations de $\mathbb{R}^n$ est noté $T(n)$ et est isomorphe au groupe abélien $(\mathbb{R}^n, +)$.

Nous avons déjà discuté de la structure d'un espace vectoriel (en particulier $\mathbb{R}^n$) comme espace affine en 8.6.

Lemme 8.59.
Décomposition d'une application affine.

(1) Une application $f \colon E \to E$ est affine si et seulement si il existe $v \in E$ et une application linéaire α sur E telle que $f = \tau_v \circ \alpha$.

(2) Dans ce cas, le choix de (v, α) est unique.

(3) Si f est bijective, alors α est bijective.

Démonstration. Nous supposons d'abord que f est affine. Alors il existe une application linéaire u_f sur E telle que

$$f(M + x) = f(M) + u_f(x) = (\tau_{f(M)} + u_f)(x) \tag{8.129}$$

pour tout x et M. De plus l'application u_f ne dépend ni de M ni de x (c'est la proposition 8.16(1)). En posant $M = 0$ nous avons :

$$f(x) = (\tau_{f(0)} \circ u_f)(x). \tag{8.130}$$

Dans l'autre sens nous supposons avoir $v \in E$ et α linéaire sur E telles que

$$f(M) = (\tau_v \circ \alpha)(M). \tag{8.131}$$

Notons qu'il y a un abus de notation entre α qui est linéaire sur l'espace *vectoriel* E et l'application α qui est une application sur l'espace *affine* E. Cet abus est légitime parce que les deux espaces sont identiques en tant qu'ensembles. Ce qui est vraiment abuser par contre, c'est de se poser ce genre de questions.

Nous avons :

$$\begin{aligned} f(M + x) = \tau_v\big(\alpha(M + x)\big) = \alpha(M + x) + v &= \alpha(M) + v + \alpha(x) \\ &= (\tau_v \circ \alpha)(M) + \alpha(x) = f(M) + \alpha(x). \end{aligned} \tag{8.132}$$

Donc la fonction f vérifie la définition 8.12. La partie (1) est prouvée.

Pour prouver l'unicité de la partie (2), nous supposons que $\tau_v \circ \alpha = \tau_w \circ \alpha$. En appliquant cela à 0 nous trouvons $v = w$. Nous avons donc $\tau_v \circ \alpha = \tau_v \circ \beta$. Comme τ_v est inversible, nous en déduisons $\alpha = \beta$.

Enfin le point (3) est relativement évident du fait que τ_v, elle, est surement bijective. □

Corolaire 8.60.
Une application affine qui conserve l'origine est linéaire.

Démonstration. Conserver l'origine demande de poser $v = 0$ dans l'expression du lemme 8.59. □

Proposition 8.61.
Soit une application affine $f\colon \mathbb{R}^n \to \mathbb{R}^n$. L'ensemble des points fixes

$$\mathrm{Fix}(f) = \{x \in \mathbb{R}^n \text{ tel que } f(x) = x\} \tag{8.133}$$

est soit vide soit un sous-espace affine de $\mathbb{R}^n$.

Démonstration. Soit $f = \tau_v \circ \alpha$; nous avons $x \in \mathrm{Fix}(f)$ si et seulement si

$$x = \tau_v\big(\alpha(x)\big) = \alpha(x) + v, \tag{8.134}$$

autrement dit, en considérant l'application linéaire $\beta = \mathrm{Id} - \alpha$, si et seulement si $\beta(x) = v$. Nous écrivons $\mathrm{Fix}(f) = \beta^{-1}(v)$. Supposons que cet ensemble soit non vide et considérons $x_0 \in \beta^{-1}(v)$. Nous avons

$$\begin{aligned}
\beta^{-1}(v) &= \{x \in \mathbb{R}^n \text{ tel que } \beta(x) = \beta(x_0)\} && (8.135a)\\
&= \{x \text{ tel que } \beta(x - x_0) = 0\} && (8.135b)\\
&= \{x \text{ tel que } x - x_0 \in \ker(\beta)\} && (8.135c)\\
&= \ker(\beta) + x_0 && (8.135d)\\
&= \tau_{x_0}\big(\ker(\beta)\big). && (8.135e)
\end{aligned}$$

Mais comme $\ker(\beta)$ est un sous-espace vectoriel, $\beta^{-1}(v)$ est le translaté d'un sous-espace vectoriel, c'est-à-dire un sous-espace affine. □

8.9.1 Structure de groupe pour les applications affines

Proposition-Définition 8.62 ([1])**.**
L'ensemble des applications affines bijectives de $\mathbb{R}^n$ forment un groupe pour la composition. Les lois de groupe sont données par les formules suivantes :

(1) Le neutre est l'identité.

(2) Le produit est donné par

$$(\tau_v \circ \alpha)(\tau_w \circ \beta) = \tau_{\alpha(w)+v} \circ \alpha\beta. \tag{8.136}$$

(3) L'inverse est donné par

$$(\tau_v \circ \alpha)^{-1} = \tau_{-\alpha^{-1}(v)} \circ \alpha^{-1}. \tag{8.137}$$

Ce groupe est noté $\mathrm{Aff}(\mathbb{R}^n)$.

Démonstration. Pour l'identité, oui, composer par l'identité est neutre.

Le fait que la formule (8.136) soit vraie est un simple calcul :

$$(\tau_v \circ \alpha) \circ (\tau_w \circ \beta)(x) = (\alpha\beta)(x) + \alpha(w) + v = \big(\tau_{\alpha(w)+v} \circ \alpha\beta\big)x. \tag{8.138}$$

Le fait que la formule (8.136) donne bien un produit pour tous les éléments de $\mathrm{Aff}(\mathbb{R}^n)$ est le lemme 8.59.

En ce qui concerne l'inverse, c'est un calcul :

$$\begin{aligned}
(\tau_{-\alpha^{-1}(v)}\alpha^{-1})(\tau_v\alpha)(x) &= (\tau_{-\alpha^{-1}(v)\alpha^{-1}})\big(\alpha(x) + v\big) && (8.139a)\\
&= \tau_{-\alpha^{-1}(v)}\big(x + \alpha^{-1}(v)\big) && (8.139b)\\
&= x. && (8.139c)
\end{aligned}$$

□

Si $f\colon \mathbb{R}^n \to \mathbb{R}^n$ est une application affine, la proposition 8.59 affirme qu'il existe une application linéaire u telle que

$$f(x+y) = f(x) + u(y). \tag{8.140}$$

En écrivant cela pour $x = 0$,

$$f(y) = f(0) + u(y), \tag{8.141}$$

ou encore $f = \tau_{f(0)} \circ u$.

Proposition 8.63.
L'ensemble $\mathrm{Aff}(\mathbb{R}^n)$ *est isomorphe au produit semi-direct*[11]

$$\mathrm{Aff}(\mathbb{R}^n) \simeq T(n) \times_{\boldsymbol{Ad}} \mathrm{GL}(n,\mathbb{R}) \tag{8.142}$$

où $\boldsymbol{Ad}$ *est l'action adjointe, c'est-à-dire*

$$\begin{aligned} \boldsymbol{Ad}\colon \mathrm{GL}(n,\mathbb{R}) &\to \mathrm{Aut}\big(T(n)\big) \\ \alpha &\mapsto \big(\tau_v \mapsto \alpha \circ \tau_v \circ \alpha^{-1}\big). \end{aligned} \tag{8.143}$$

Démonstration. L'application que nous allons montrer être un isomorphisme est ψ qui à $f = \tau_v \circ \alpha$ fait correspondre le couple $(\tau_v, \alpha) \in T(n) \times \mathrm{GL}(n,\mathbb{R})$.

(i) **Égalité d'ensembles** Il faut que $\mathrm{Aff}(\mathbb{R}^n)$ soit en bijection avec $T(n) \times \mathrm{GL}(n,\mathbb{R})$. En effet si $f \in \mathrm{Aff}(\mathbb{R}^n)$, la décomposition $f = \tau_v \circ \alpha$ est unique. D'abord en appliquant à 0, $f(0) = \tau_v\big(\alpha(v)\big) = v$. Donc v est fixé par la valeur de $f(0)$. Ensuite $\alpha = f \circ \tau_v^{-1}$, donc α fixé.

(ii) **L'action adjointe fonctionne** Il faut vérifier que $\alpha \circ \tau_v \circ \alpha^{-1}$ est bien dans $T(n)$. Pour cela, en agissant sur $x \in \mathbb{R}^n$ nous trouvons

$$\alpha\tau_v\alpha^{-1}(x) = \alpha\big(\alpha^{-1}(x) + v\big) = x + \alpha(v) = \tau_{\alpha(v)}(x). \tag{8.144}$$

Le fait que $\mathbf{Ad}(\alpha)$ soit un automorphisme est toujours correct.

(iii) **Morphisme** Il faut vérifier que l'application ψ est un morphisme de groupe. D'abord la loi de groupe sur $\mathrm{Aff}(\mathbb{R}^n)$ est donnée par

$$(\tau_v \circ \alpha) \circ (\tau_w \circ \beta) = \tau_{v+\alpha(w)} \circ (\alpha \circ \beta). \tag{8.145}$$

Ensuite la loi de groupe du produit semi-direct est donnée par

$$(\tau_v, \alpha) \cdot (\tau_w, \beta) = \big(\tau_v\, \mathbf{Ad}(\alpha)\tau_w, \alpha\beta\big) = \big(\tau_v\tau_{\alpha(w)}, \alpha\beta\big) = \big(\tau_{\alpha(w)+v}, \alpha\beta\big). \tag{8.146}$$

Nous avons donc bien

$$\psi\big((\tau_v, \beta) \cdot (\tau_w, \beta)\big) = \psi(\tau_v, \beta) \circ \psi(\tau_w, \beta). \tag{8.147}$$

□

8.10 Isométries

Définition 8.64 (Isométrie d'espace affine)**.**
Si $\mathcal{E}$ *est un espace affine muni d'une distance* d*, une isométrie de* $\mathcal{E}$ *est une application* $f\colon \mathcal{E} \to \mathcal{E}$ *préservant* d*.*

Notons que toutes les applications affines ne sont pas des isométries : par exemple les homothéties.

11. Définition 2.47.

Proposition 8.65.
Si $\mathcal{E}$ est modelé sur un espace euclidien $(E, \|.\|)$ alors la formule

$$d(A, B) = \|\overrightarrow{AB}\| \tag{8.148}$$

définit une distance[12] *sur $\mathcal{E}$.*

Démonstration. Étant donné la proposition 8.3, la formule a un sens parce qu'à A et B donnés dans $\mathcal{E}$, il est associé un unique vecteur $\overrightarrow{AB} \in E$.

Pour vérifier que d est une distance, nous vérifions les points de la définition 7.97 et nous utilisons les propriétés correspondantes dans la définition 7.136 d'une norme.

(1) $d(A, B) = \|\overrightarrow{AB}\| \geqslant 0$.

(2) Si $d(A, B) = 0$, alors $\|\overrightarrow{AB}\| = 0$, ce qui implique que $\overrightarrow{AB} = 0$. Nous avons donc

$$B = A + \overrightarrow{AB} \qquad \text{proposition 8.3} \tag{8.149a}$$
$$= A + 0 \tag{8.149b}$$
$$= A \qquad \text{lemme 8.4.} \tag{8.149c}$$

(3) En utilisant la proposition 8.5(3),

$$d(A, B) = \|\overrightarrow{AB}\| = \| - \overrightarrow{BA}\| = \|BA\| = d(B, A) \tag{8.150}$$

(4) En utilisant les relation de Chasles 8.5(1) ainsi que l'inégalité triangulaire 7.136(4)

$$d(A, B) = \|\overrightarrow{AB}\| = \|\overrightarrow{AC} + \overrightarrow{CB}\| \leqslant \|\overrightarrow{AC}\| + \|\overrightarrow{CB}\| = d(A, C) + d(C, B). \tag{8.151}$$

□

Nous parlons d'isométries affines ou linéaires dans le thème 75.

12. Définition 7.97.

Chapitre 9

Espaces vectoriels (encore)

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 ([222])**.**
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. Soient $\{e_i\}$, une base de E, une n-forme linéaire alternée $f\colon E \to \mathbb{K}$ ainsi que des vecteurs $(v_1, \dots, v_n)$ 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 $\alpha_{i1} = \alpha_{i2}$:

$$\begin{aligned}
\Pi(v, v, v_3, \dots, v_n) &= \sum_{\sigma \in S_n} \epsilon(\sigma)\alpha_{\sigma(1)1}\alpha_{\sigma(2)2}\alpha_{\sigma(3)3} \dots \alpha_{\sigma(n)n} && (9.6a)\\
&= \sum_{\sigma \in S_n} \epsilon(\sigma\tau)\alpha_{\sigma\tau(1)1}\alpha_{\sigma\tau(2)2}\alpha_{\sigma\tau(3)3} \dots \alpha_{\sigma\tau(n)n} \qquad \text{où } \tau = (12) && (9.6b)\\
&= -\sum_{\sigma \in S_n} \epsilon(\sigma)\alpha_{\sigma(1)1}\alpha_{\sigma(2)2}\alpha_{\sigma(3)3} \dots \alpha_{\sigma(n)n} && (9.6c)\\
&= -\Pi(v, v, v_3, \dots, v_n). && (9.6d)
\end{aligned}$$

□

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 9.4 (Déterminant d'une famille de vecteurs[9])**.**
*Le **déterminant** de la famille de vecteurs $(v_1, \dots, v_n)$ dans la base B est l'élément de $\mathbb{K}$*

$$\det_{(e_1,\dots,e_n)}(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} e^*_{\sigma(i)}(v_i) \tag{9.7}$$

où

- *la somme porte sur le groupe symétrique,*
- *le nombre $\epsilon(\sigma)$ est la signature*[2] *de la permutation σ,*
- *les éléments $\{e_i\}$ forment la base canonique de $\mathbb{K}^n$.*
- *les éléments $\{e^*_i\}$ sont la base duale de $\{e_i\}$.*

Nous le notons $\det_{(e_1,\dots,e_n)}(v_1, \dots, v_n)$.

9.5.
La base $\{e_i\}$ est la base canonique de $\mathbb{K}^n$, et l'élément e^*_k est la forme linéaire définie par

$$\begin{aligned}
e^*_k \colon \mathbb{K}^n &\to \mathbb{K}\\
\sum_i x_i e_i &\mapsto x_k.
\end{aligned} \tag{9.8}$$

Il n'est pas sous-entendu que $\mathbb{K}^n$ ait un produit scalaire. Il n'est donc pas autorisé de dire que $\{e_i\}$ est une base orthonormée et que $e^*_k(x) = \langle e_k, x\rangle$. Ce genre d'égalités sont vraies dans le cas $\mathbb{K} = \mathbb{R}$, mais n'ont pas de sens en général.

Le lemme 11.5 va un peu parler du cas où $\mathbb{K}^n$ est muni d'une base orthonormée.

Lemme 9.6 ([9])**.**
Les propriétés du déterminant. Soit B une base de E.

1. C'est ici que joue l'hypothèse sur la caractéristique de $\mathbb{K}$.
2. Définition 1.202.

(1) L'application $\det_B \colon E^n \to \mathbb{K}$ *est* n*-linéaire.*

(2) L'application $\det_B \colon E^n \to \mathbb{K}$ *est* n*-linéaire est antisymétrique et alternée*[3].

(3) Pour toute base, $\det_B(B) = 1$.

(4) Le déterminant ne change pas si on remplace un vecteur par une combinaison linéaire des autres :

$$\det_B(v_1, \dots, v_n) = \det_B\big(v_1 + \sum_{s=2}^{n} a_s v_s, v_2, \dots, v_n\big). \tag{9.9}$$

(5) Si on permute les vecteurs,

$$\det_B(v_1, \dots, v_n) = \epsilon(\sigma) \det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)}). \tag{9.10}$$

(6) Si B' *est une autre base :*

$$\det_B = \det_B(B') \det_{B'} \tag{9.11}$$

(7) Nous avons aussi la formule $\det_B(B') \det_{B'}(B) = 1$.

(8) 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.

(i) **(1)** En posant $v_1 = x_1 + \lambda x_2$ nous avons

$$\det_B(x_1 + \lambda x_2, v_2, \dots, v_n) = \sum_\sigma \epsilon(\sigma) \prod_{i=1}^{n} e^*_{\sigma(i)}(v_i) \tag{9.12a}$$

$$= \sum_\sigma \epsilon(\sigma) \Big(e^*_{\sigma(1)}(x_1 + \lambda x_2)\Big) \prod_{i=2}^{n} e^*_{\sigma(i)}(v_i). \tag{9.12b}$$

À partir de là, la linéarité de $e^*_{\sigma(1)}$ montre que $\det_B$ est linéaire en son premier argument. Pour les autres arguments, le même calcul tient.

(ii) **(2)** Nous prouvons à présent que det est alternée. Si votre corps est de caractéristique différente de deux, vous pouvez lire la proposition 9.7.

Supposons $v_k = v_l$, et considérons la permutation $\beta = (k, l)$. Nous savons par la proposition 5.32 que $S_n = A_n \cup A_n\beta$. Cela nous permet de décomposer la somme sur S_n en deux parties :

$$\sigma_{\sigma \in S_n} (-1)^\sigma \prod_i \epsilon^*_{\sigma(i)}(v_i) = \sum_{\sigma \in A_n} (-1)^\sigma \prod_i \epsilon^*_{\sigma(i)}(v_i) + \sum_{\sigma \in A_n} (-1)^{\sigma\beta} \prod_i \epsilon^*_{(\sigma\beta)(i)}(v_i). \tag{9.13}$$

D'abord $(-1)^\sigma = 1$ et $(-1)^{\sigma\beta} = -1$. Ensuite, pour un $\sigma \in A_n$ donné, nous avons

$$\prod_i \epsilon^*_{(\sigma\beta)(i)}(v_i) = \epsilon^*_{(\sigma\beta)(k)}(v_k) \epsilon^*_{(\sigma\beta)(l)}(v_l) \prod i \overset{i \neq k}{\neq} l \; \epsilon^*_{(\sigma\beta)(i)}(v_i) \tag{9.14a}$$

$$= \epsilon^*_{\sigma(l)}(v_k) \epsilon^*_{\sigma(k)}(v_l) \prod_{\substack{i \neq k \\ i \neq l}} \epsilon^*_{\sigma(i)}(v_i) \tag{9.14b}$$

$$= \epsilon^*_{\sigma(l)}(v_l) \epsilon^*_{\sigma(k)}(v_k) \prod_{\substack{i \neq k \\ i \neq l}} \epsilon^*_{\sigma(i)}(v_i) \tag{9.14c}$$

$$= \prod_i \epsilon^*_{\sigma(i)}(v_i). \tag{9.14d}$$

Donc les deux termes de la somme (9.13) ne diffèrent que par un signe. Elle est donc nulle, et la forme déterminant est alternée.

La fonction det est antisymétrique parce que alternée, voir le lemme 9.2.

3. Alternée, définition 9.1. En caractéristique 2, alternée n'est pas équivalent à symétrique.

(iii) **(3)** Nous avons

$$\det_B(B) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \underbrace{e^*_{\sigma(i)}(e_i)}_{=\delta_{\sigma(i),i}}. \tag{9.15}$$

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.

(iv) **(4)** Vu que $\det_B$ est linéaire en tous ses arguments,

$$\det_B\left(v_1 + \sum_{s=2}^{n} a_s v_s, v_2, \dots, v_n\right) = \det_B(v_1, \dots, v_n) + \sum_{s=2}^{n} a_s \det_B(v_s, v_2, \dots, v_n). \tag{9.16}$$

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.

(v) **(5)** 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 :

$$\det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(w_i) \tag{9.17a}$$

$$= \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(v_{\sigma(i)}) \tag{9.17b}$$

$$= \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma^{-1}\sigma'(i)}(v_i) \tag{9.17c}$$

$$= \sum_{\sigma'} \epsilon(\sigma\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(v_i) \tag{9.17d}$$

$$= \epsilon(\sigma) \det_B(v_1, \dots, v_n). \tag{9.17e}$$

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 1.205).

(vi) **(6)** 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' :

$$\det_B(B') = \lambda \det_{B'}(B'), \tag{9.18}$$

donc $\lambda = \det_B(B')$.

(vii) **(7)** Il suffit d'appliquer l'égalité précédente à B en nous souvenant que $\det_B(B) = 1$.

(viii) **(8)** 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 la proposition 4.17. Il y a donc moyen de remplacer un des vecteurs par une combinaison linéaire des autres. Le déterminant s'annule alors.

□

Proposition 9.7.
Si la caractéristique du corps de base n'est pas deux, le déterminant est antisymétrique et alterné.

Démonstration. Si la caractéristique du corps de base n'est pas deux, une forme antisymétrique est alternée (lemme 9.2).

Pour prouver que le déterminant est antisymétrique, remarquez que permuter v_k et v_l revient à calculer le nombre $\det_B(v_{\sigma_{kl}(1)}, \dots, v_{\sigma_{kl}(n)})$ au lieu de $\det_B(v_1, \dots, v_n)$. Cela revient à changer la somme $\sum_\sigma$ en $\sum_{\sigma\circ\sigma_{kl}}$. Cela ajoute 1 à $\epsilon(\sigma)$ vu que l'on ajoute une permutation.

Donc le déterminant est antisymétrique. Nous en déduisons qu'il est alterné parce que, en permutant trivialement v_1 et v_1, nous obtenons $\det_B(v_1, v_1) = -\det_B(v_1, v_1)$. Si le corps est de caractéristique différente de deux, cela implique que $\det_B(v_1, v_1) = 0$. □

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 le théorème 14.264.

Lemme-Définition 9.8.
Si $f \colon E \to E$ est un endomorphisme, et si les parties B et B' sont deux bases, alors[4]

$$\det_B\big(f(B)\big) = \det_{B'}\big(f(B')\big). \tag{9.19}$$

Ce nombre, indépendant de la base choisie est nommé le ***déterminant*** *de f et est noté $\det(f)$.*

Démonstration. L'application

$$\begin{aligned} \varphi \colon E^n &\to \mathbb{K} \\ v_1, \dots, v_n &\mapsto \det_B\big(f(v_1), \dots, f(v_n)\big) \end{aligned} \tag{9.20}$$

est n-linéaire et alternée ; il existe donc $\lambda \in \mathbb{K}$ tel que $\varphi = \lambda \det_B$. En appliquant cela à B :

$$\det_B\big(f(B)\big) = \lambda \det_B(B) = \lambda. \tag{9.21}$$

Nous avons donc déjà prouvé que $\lambda = \det_B\big(f(B)\big)$, c'est-à-dire

$$\det_B\big(f(v)\big) = \det_B\big(f(B)\big) \det_B(v). \tag{9.22}$$

Nous allons maintenant introduire B' là où il y a du v en utilisant les formules (9.11) :

$$\det_B\big(f(v)\big) = \det_B(B') \det_{B'}\big(f(v)\big) \tag{9.23a}$$

$$\det_B(v) = \det_B(B') \det_{B'}(v). \tag{9.23b}$$

Nous obtenons

$$\det_{B'}\big(f(v)\big) = \det_B\big(f(B)\big) \det_{B'}(v). \tag{9.24}$$

Et on applique cela à $v = B'$:

$$\det_{B'}\big(f(B')\big) = \det_B\big(f(B)\big) \underbrace{\det_{B'}(B')}_{=1}. \tag{9.25}$$

□

Proposition 9.9.
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[5] *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.

4. Définition de $\det_B(B')$, 9.4.
5. Endomorphisme inversible, définition 4.34.

(1) Nous considérons l'application

$$\begin{aligned}\varphi\colon E^n &\to \mathbb{K}\\ v &\mapsto \det_B\big(f(v)\big).\end{aligned} \tag{9.26}$$

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 :

$$\lambda = \det(f). \tag{9.27}$$

Calculons à présent $\varphi\big(g(B)\big)$: d'une part,

$$\varphi\big(g(B)\big) = \det_B\big((f\circ g)(B)\big) \tag{9.28}$$

et d'autre part,

$$\varphi\big(g(B)\big) = \lambda \det_B\big(g(B)\big) = \lambda \det(g) \tag{9.29}$$

En égalisant et en reprenant la la valeur déjà trouvée de λ,

$$\det\big(f\circ g)(B)\big) = \det(f)\det(g), \tag{9.30}$$

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 9.6(8)).
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, $1 = \det(f\circ f^{-1}) = \det(f)\det(f^{-1})$.

□

Proposition 9.10.
Soient deux espaces vectoriels E et F de dimension finies n et m sur le corps $\mathbb{K}$ munis de bases $\{e_i\}$ et $\{f_\alpha\}$. À une matrice $A \in \mathbb{M}(m\times n, \mathbb{K})$ nous associons l'application linéaire[6]

$$f_A(x) = \sum_{i\alpha} A_{\alpha i} x_i f_\alpha. \tag{9.31}$$

Alors, en ce qui concerne les déterminants[7]*, nous avons*

(1) $\det(f_A) = \det(A)$

(2) $\det(f_{AB}) = \det(f_A)\det(f_B)$

Démonstration. Nous devons étudier la formule

$$\det(f_A) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^n e^*_{\sigma(i)}\big(f_A(e_i)\big). \tag{9.32}$$

En premier lieu nous avons

$$f_A(e_i) = \sum_{jk} A_{jk}(e_i)_k e_j = \sum_j A_{ji} e_j. \tag{9.33}$$

Nous avons alors

$$e^*_{\sigma(i)}\big(f_A(e_i)\big) = \sum_j A_{ji} \underbrace{e^*_{\sigma(i)}(e_j)}_{\delta_{j\sigma(i)}} = A_{\sigma(i)i}. \tag{9.34}$$

6. Dont nous avons déjà beaucoup parlé entre autres dans la proposition 4.70.
7. Définition 9.8 pour les applications linéaires et 4.74 pour les matrices.

Au final,

$$\det(f_A) = \sum_\sigma \epsilon(\sigma) \prod_{i=1}^n A_{\sigma(i)i} = \det(A^t) = \det(A) \tag{9.35}$$

où la dernière égalité est autorisée par le lemme 4.76.

Cela prouve la formule $\det(f_A) = \det(A)$.

En ce qui concerne la seconde formule, il s'agit de se souvenir de la proposition 4.70 qui donne $f_{AB} = f_A \circ f_B$, et ensuite de la proposition 9.9(1) qui donne $\det(f_A \circ f_B) = \det(f_A)\det(f_B)$. □

9.1.4 Déterminant de Vandermonde

Proposition 9.11 ([101]).
*Le **déterminant de Vandermonde** est le polynôme en n variables donné par*

$$V(T_1, \dots, T_n) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ T_1 & T_2 & \dots & T_n \\ \vdots & \ddots & \ddots & \vdots \\ T_1^{n-1} & T_2^{n-1} & \dots & T_n^{n-1} \end{pmatrix} = \prod_{1 \leqslant i < j \leqslant n} (T_j - T_i). \tag{9.36}$$

Notez que l'inégalité du milieu est stricte (sinon d'ailleurs l'expression serait nulle).

Démonstration. Nous considérons le polynôme

$$f(X) = V(T_1, \dots, T_{n-1}, X) \in \big(\mathbb{K}[T_1, \dots, T_{n-1}]\big)[X]. \tag{9.37}$$

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 [8] nous pouvons factoriser les $X - T_i$, c'est-à-dire qu'il existe $\alpha \in \mathbb{K}[T_1, \dots, T_{n-1}]$ tel que

$$f = \alpha \prod_{i=1}^{n-1} (X - T_i). \tag{9.38}$$

Nous trouvons α en écrivant $f(0)$. D'une part la formule (9.38) nous donne

$$f(0) = \alpha(-1)^{n-1} T_1 \dots T_{n-1}. \tag{9.39}$$

D'autre part la définition donne

$$f(0) = \det \begin{pmatrix} 1 & \cdots & 1 & 1 \\ T_1 & & T_{n-1} & 0 \\ \vdots & & \vdots & \vdots \\ T_1^{n-1} & \cdots & T_{n-1}^{n-1} & 0 \end{pmatrix} \tag{9.40a}$$

$$= (-1)^{n-1} \det \begin{pmatrix} T_1 & \dots & T_{n-1} \\ \vdots & \ddots & \vdots \\ T_1^{n-1} & \dots & T_{n-1}^{n-1} \end{pmatrix} \tag{9.40b}$$

$$= (-1)^{n-1} T_1 \dots T_{n-1} \det \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ T_1^{n-1} & \cdots & T_{n-1}^{n-1} \end{pmatrix} \tag{9.40c}$$

$$= (-1)^{n-1} T_1 \dots T_{n-1} V(T_1, \dots, T_{n-1}) \tag{9.40d}$$

En égalisant avec (9.39), nous trouvons $\alpha = V(T_1, \dots, T_{n-1})$, et donc

$$f = V(T_1, \dots, T_{n-1}) \prod_{j \leqslant n-1} (X - T_j) \tag{9.41}$$

8. Proposition 3.142.

Enfin, une récurrence montre que

$$\begin{aligned} V(T_1,\dots,T_n) &= f(T_n) && (9.42a)\\ &= V(T_1,\dots,T_{n-1})\prod_{j\leqslant n-1}(T_n-T_j) && (9.42b)\\ &= \prod_{k\leqslant n}\prod_{j\leqslant k-1}(T_k-T_j) && (9.42c)\\ &= \prod_{1\leqslant j<k\leqslant n}(T_i-T_j). && (9.42d)\end{aligned}$$

□

Exemple 9.12.
Le déterminant de Vandermonde (proposition 9.11) 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 9.13.
Un polynôme semi-symétrique $f \in \mathbb{K}[T_1,\dots,T_n]$ se décompose de façon unique en

$$f = P + VQ \tag{9.43}$$

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 $\{1,\dots,n\}$, alors $\sigma_1 \cdot f = \sigma_2 \cdot f$ parce que l'élément $\sigma_2^{-1}\sigma_1$ est pair (proposition 1.205), 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

$$g = \tau \cdot f = P - VQ. \tag{9.44}$$

Donc $f + g = 2P$ et $f - g = 2VQ$. Cela donne P et Q en termes 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,

$$\sigma \cdot (f + g) = \sigma \cdot f + \sigma\tau \cdot f = g + f \tag{9.45}$$

parce que $\sigma\tau$ est pair. De la même façon,

$$\sigma \cdot (f - g) = g - f = \epsilon(\sigma)(f - g). \tag{9.46}$$

Dans les deux cas nous concluons en utilisant le fait que toute permutation est un produit de transpositions (proposition 1.200) et que ϵ est un homomorphisme.

Soient maintenant deux entiers $h < k$ dans $\{1,\dots,n\}$ et l'anneau

$$\big(\mathbb{K}[T_1,\dots,\hat{T}_k,\dots,T_n]\big)[T_k]. \tag{9.47}$$

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 $T_k - T_h$:

$$f - g = (T_k - T_h)q + r \tag{9.48}$$

où $\deg_{T_k} r < 1$, c'est-à-dire que r ne dépends pas de T_k. Nous revoyons maintenant l'égalité (9.48) 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

$$-(f-g) = -(T_k - T_h)\tau_{kh} \cdot q + \tau_{kh} \cdot r \tag{9.49}$$

où $\tau_{kh} \cdot r$ ne dépend pas de T_h. Nous appliquons à (9.49) l'application

$$\begin{aligned} t\alpha \colon \mathbb{K}[T_1, \dots, T_n] &\to \mathbb{K}[T_1, \dots, \hat{T}_k, \dots, T_n] \\ \alpha(PT_1, \dots, \hat{T}_k, \dots, T_n) &= P(T_1, \dots, T_h, \dots, T_n). \end{aligned} \tag{9.50}$$

Cette application vérifie $\alpha\big(\tau_{kh} \cdot r\big) = \alpha(r)$ et nous avons

$$-\alpha(f-g) = \alpha(r). \tag{9.51}$$

Puis en appliquant α à la relation $f - g = (T_k - T_h)q + r$, nous trouvons

$$\alpha(f-g) = \alpha(r), \tag{9.52}$$

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

$$f - g = 2Q \prod_{h<k} \prod_{k \leqslant n} (T_k - T_h) = 2Q(T_1, \dots, T_n)V(T_1, \dots, T_n), \tag{9.53}$$

où nous avons utilisé la formule du déterminant de Vandermonde de la proposition 9.11.

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

$$\sigma \cdot (f-g) = \epsilon(\sigma)(f-g) = \epsilon(\sigma)2QV, \tag{9.54}$$

Mais en appliquant σ à l'équation (9.53),

$$\sigma \cdot (f-g) = 2(\sigma \cdot V)(T_1, \dots, , T_n)(\sigma \cdot Q)(T_1, \dots, T_n) \tag{9.55a}$$

$$= 2\epsilon(\sigma)V(T_1, \dots, T_n)(\sigma \cdot Q)(T_1, \dots, T_n). \tag{9.55b}$$

Nous égalisons cela avec (9.54) et nous souvenant que l'anneau $\mathbb{K}[T_1, \dots, T_n]$ est intègre par le théorème 3.126. Ensuite nous simplifions par $2\epsilon(\sigma)V$ pour obtenir

$$Q = \sigma \cdot Q, \tag{9.56}$$

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,

$$f = P + VQ. \tag{9.57}$$

□

9.1.5 Déterminant de Gram

Si $x_1, \dots, x_r$ sont des vecteurs d'un espace vectoriel, alors le **déterminant de Gram** est le déterminant

$$G(x_1, \dots, x_r) = \det\big(\langle x_i, x_j \rangle\big). \tag{9.58}$$

Notons que la matrice est une matrice symétrique.

Proposition 9.14.
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

$$d(x, F)^2 = \frac{G(x, x_1, \dots, x_n)}{G(x_1, \dots, x_n)}. \tag{9.59}$$

9.1.6 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

$$D_n = \det\left(\frac{1}{a_i + b_j}\right). \tag{9.60}$$

Proposition 9.15 ([223]).
Le déterminant de Cauchy est donné par la formule

$$D_n = \frac{\prod_{i<j}(a_j - a_i)\prod_{i<j}(b_j - b_i)}{\prod_{ij}(a_i + b_j)}. \tag{9.61}$$

9.1.7 Matrice de Sylvester

La définition est pompée de wikipédia. Soient P et Q deux polynômes non nuls, de degrés respectifs m et n :

$$P(x) = p_0 + p_1 x + \cdots + p_n x^n \tag{9.62a}$$
$$Q(x) = q_0 + q_1 x + \cdots + q_m x^m. \tag{9.62b}$$

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 :

$$\begin{pmatrix} p_n & p_{n-1} & \cdots & p_1 & p_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{9.63}$$

(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 :

$$\begin{pmatrix} q_m & q_{m-1} & \cdots & q_1 & q_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{9.64}$$

(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

$$S_{p,q} = \begin{pmatrix} p_4 & p_3 & p_2 & p_1 & p_0 & 0 & 0 \\ 0 & p_4 & p_3 & p_2 & p_1 & p_0 & 0 \\ 0 & 0 & p_4 & p_3 & p_2 & p_1 & p_0 \\ q_3 & q_2 & q_1 & q_0 & 0 & 0 & 0 \\ 0 & q_3 & q_2 & q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_3 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & 0 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}. \tag{9.65}$$

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 9.16 ([224]).
Si P et Q sont deux polynômes de degrés n et m à coefficients dans l'anneau $\mathbb{A}$, alors pour tout $\lambda \in \mathbb{A}$,

$$\operatorname{res}(\lambda P, Q) = \lambda^m \operatorname{res}(P, Q) \tag{9.66a}$$
$$\operatorname{res}(P, \lambda Q) = \lambda^n \operatorname{res}(P, Q). \tag{9.66b}$$

Démonstration. Cela est simplement un comptage du 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 (6.91) peut être traitée avec une matrice de Sylvester. Soient P et Q, deux polynômes donnés et à résoudre l'équation

$$xP + yQ = 0 \tag{9.67}$$

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 (9.67) sous la forme

$$S_{PQ}^t \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = 0. \tag{9.68}$$

Pour s'en convaincre, écrivons pour les polynômes de l'exemple (9.65) :

$$\begin{pmatrix} p_4 & 0 & 0 & q_3 & 0 & 0 & 0 \\ p_3 & p_4 & 0 & q_2 & q_3 & 0 & 0 \\ p_2 & p_3 & p_4 & q_1 & q_2 & q_3 & 0 \\ p_1 & p_2 & p_3 & q_0 & q_1 & q_2 & q_3 \\ p_0 & p_1 & p_2 & 0 & q_0 & q_1 & q_2 \\ 0 & p_0 & p_1 & 0 & 0 & q_0 & q_1 \\ 0 & 0 & p_0 & 0 & 0 & 0 & q_0 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \\ x_0 \\ y_3 \\ y_2 \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_2p_4 + y_2q_3 \\ p_3x_2 + p_4x_1 + q_2y_3 + q_3y_2 \\ \\ \vdots \\ \\ \end{pmatrix} \tag{9.69}$$

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 9.17.
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 20.98 sur l'intégration de fractions rationnelles.

Exemple 9.18.
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

$$S_{P,P'} = \begin{pmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{pmatrix}. \tag{9.70}$$

Le résultant est alors

$$\operatorname{res}(P, P') = -a(b^2 - 4ac). \tag{9.71}$$

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 6.44 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 9.19.
Si

$$P(X) = a_p \prod_{i=1}^{p} (X - \alpha_i) \tag{9.72a}$$

$$Q(X) = b_q \prod_{j=1}^{q} (X - \beta_i) \tag{9.72b}$$

alors nous avons les expressions suivantes pour le résultant :

$$\operatorname{res}(P,Q) = a_p^q b_q^p \prod_{i=1}^{p}\prod_{j=1}^{q}(\beta_j - \alpha_i) = b_q^p \prod_{j=1}^{q} P(\beta_j) = (-1)^{pq} a_p^q \prod_{i=1}^{p} Q(\alpha_i). \tag{9.73}$$

Démonstration. Si P et Q ne sont pas premiers entre eux, d'une part la proposition 9.17 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 (9.73) 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

$$\mathbb{A} = \mathbb{Z}[\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q]. \tag{9.74}$$

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

$$R = \lambda(\alpha_1, \ldots, \beta_q) \prod_{i=1}^{p}\prod_{j=1}^{r}(\beta_j - \alpha_i). \tag{9.75}$$

Comptons les degrés. Pour donner une idée de ce calcul de degré, voici comment se présente, au niveau des dimensions, le déterminant :

$$\begin{array}{ccccccccc}
 & \overleftrightarrow{\quad p+1 \quad} & & & & \overleftrightarrow{q-1} & & & \\
a_p & a_{p-1} & \cdots & a_0 & & 0 & \cdots & 0 & \updownarrow q \\
 & & a_p & & & & a_1 & & a_0 \\
0 & \cdots & 0 & & & & & & \\
 & & & & \overleftrightarrow{\quad p+q \quad} & & & &
\end{array} \tag{9.76}$$

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 (9.76). 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 9.16 nous permet de conclure.

□

9.1.8 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

$$P = X^n + \sum_{k=0}^{n-1} a_k X^k. \tag{9.77}$$

Théorème 9.20 (Kronecker[89])**.**
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 $R = X^n + \sum_{k=0}^{n-1} b_k X^k$ un élément de K_n dont nous notons $\beta_1, \dots, \beta_n$ les racines dans $\mathbb{C}$. Les relations coefficients-racines stipulent que

$$b_k = \sum_{1 \leqslant i_1 < \dots < i_{n-k} \leqslant n} \prod_{j=1}^{n-k} \beta_{i_j}. \tag{9.78}$$

En prenant le module et en se souvenant que $|\beta_l| \leqslant 1$ pour tout l, nous trouvons que

$$|b_k| \leqslant \binom{n}{n-k}. \tag{9.79}$$

Mais comme $b_k \in \mathbb{Z}$, nous avons

$$b_k \in \Big\{ -\binom{n}{n-k}, -\binom{n}{n-k} + 1, \dots, 0, \cdots, \binom{n}{n-k} \Big\} \tag{9.80}$$

qui est de cardinal $\binom{n}{n-k} + 1$. Nous avons donc

$$\mathrm{Card}(K_n) \leqslant \prod_{k=0}^{n-1} \Big(1 + \binom{n}{n-k}\Big) < \infty. \tag{9.81}$$

La conclusion jusqu'ici est que K_n est un ensemble fini.

Pour chaque $k \in \mathbb{N}^*$ nous considérons les polynômes

$$P_k = \prod_{i=1}^{n} (X - \alpha_i^k) \tag{9.82a}$$

$$Q_k = X^k - Y \in \mathbb{Z}[X, Y], \tag{9.82b}$$

et puis nous considérons le résultant $R_k = \mathrm{res}_X(P, Q_k) \in \mathbb{Z}[Y]$:

$$R_k = \mathrm{res}_X(P, Q_k) = \begin{pmatrix} 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & & & \\ 0 & \cdots & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 \\ & & & & & & & & \\ 1 & 0 & \dots & 0 & -Y & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & -Y & 0 & \dots & 0 \\ & & \ddots & & & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y \end{pmatrix} \tag{9.83}$$

Cela est un polynôme en Y dont le terme de plus haut degré est $(-1)^n Y^n$. Les petites formules de la proposition 9.19 nous permettent d'exprimer $R_k(Y)$ en termes des racines de P :

$$R_k(Y) = \prod_{i=1}^{n} Q_k(\alpha_i) = \prod_{i=1}^{n}(\alpha_i^k - Y) = (-1)^n \prod_{i=1}^{n}(Y - \alpha_i^k) = (-1)^n P_k(Y). \tag{9.84}$$

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) [9]. 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

$$\alpha^{2^l/2^k} = \alpha^{2^{l-k}} \tag{9.85}$$

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 (9.85) est racine de P_{2^l}, il est aussi racine de P_{2^k}. Donc

$$\left(\alpha^{2^{l-k}}\right)^{2^{l-k}} = \alpha^{2^{2(l-k)}} \tag{9.86}$$

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

$$\alpha^{2^{n(l-k)}} = \alpha^{2^{m(l-k)}} \tag{9.87}$$

pour certains entiers $m > n$, alors

$$\alpha^{2^{n(l-k)} - 2^{m(l-k)}} = 1, \tag{9.88}$$

ce qui prouve 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. □

9.2 Orientation

9.2.1 Cas vectoriel

Proposition-Définition 9.21 ([225]).
Soient deux bases $\mathcal{B}$ et $\mathcal{B}'$ d'un espace vectoriel réel E. Nous définissons la relation $\mathcal{B} \sim \mathcal{B}'$ si et seulement si $\det_{\mathcal{B}}(\mathcal{B}') > 0$ [10].

Cela est une relation d'équivalence [11] *sur l'ensemble des bases de E, et les classes sont les* ***orientations*** *de E.*

Démonstration. Tout est dans le lemme 9.6. D'abord quand $\mathcal{B}$ et $\mathcal{B}'$ sont des bases, $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$ ensuite, nous passons en revue les points qu'il faut pour être une relation d'équivalence.

(1) $\mathcal{B} \sim \mathcal{B}$ parce que $\det_{\mathcal{B}}(\mathcal{B}) = 1 > 0$.

(2) Vu que $\det_{\mathcal{B}}(\mathcal{B}') = \frac{1}{\det_{\mathcal{B}'}(\mathcal{B})}$, les deux sont positifs en même temps ou pas du tout.

(3) Si $\mathcal{B} \sim \mathcal{B}'$ et $\mathcal{B}' \sim \mathcal{B}''$, alors en utilisant la formule

$$\det_{\mathcal{B}}(\mathcal{B}'') = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B}''), \tag{9.89}$$

nous voyons que $\det_{\mathcal{B}}(\mathcal{B}'') > 0$.

□

9. Ici dans [89], 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.

10. Définition 9.4.

11. Définition 1.29.

Lemme 9.22.
Soit un espace vectoriel réel E. L'ensemble des bases de E possède exactement deux orientations[12]

Démonstration. Nous considérons une base $\mathcal{B} = (e_1, \dots, e_n$[13] à partir de laquelle nous définissons une autre base : $\mathcal{B}' = (-e_1, e_2, \dots, e_n)$. Nous allons prouver que ces deux bases ne sont pas équivalentes, et que toute base de E est équivalente soit à $\mathcal{B}$ soit à $\mathcal{B}'$.

(i) **Au moins deux classes** Le fait que $\det_{\mathcal{B}}(\mathcal{B}') = -1$ vient du fait que $\det_{\mathcal{B}}(\mathcal{B}') = 1$ et que l'application $\det_{\mathcal{B}}$ est n-linéaire ; en multipliant par -1 le premier argument, la valeur du déterminant est multipliée par -1.

Donc les bases $\mathcal{B}$ et $\mathcal{B}'$ ne sont pas équivalentes et il existe au moins deux classes.

(ii) **Au plus deux classes** Nous montrons à présent que toute base est équivalente soit à $\mathcal{B}$ soit à $\mathcal{B}'$. Supposons que $\mathcal{B}''$ ne soit pas équivalente à $\mathcal{B}$, c'est-à-dire que $\det_{\mathcal{B}}(\mathcal{B}'') < 0$. Nous utilisons encore la formule (9.11),

$$\underbrace{\det_{\mathcal{B}}(\mathcal{B}'')}_{<0} = \underbrace{\det_{\mathcal{B}}(\mathcal{B}')}_{<0} \det_{\mathcal{B}'}(\mathcal{B}''), \tag{9.90}$$

et nous déduisons que $\det_{\mathcal{B}'}(\mathcal{B}'') > 0$.

□

9.23.
Vu qu'il n'y a que deux classes d'équivalence parmi les bases, nous pouvons utiliser le vocable « avoir la même orientation que » ou « avoir l'orientation contraire de ». Ce n'est pas ambigu.

Proposition 9.24 ([225])**.**
Si $\mathcal{B}$ est une base de l'espace vectoriel E de dimension n, et si τ est une transposition[14] *de S_n, alors la base $\tau(\mathcal{B})$ est de sens contraire.*

Démonstration. Le lemme 9.6(2) dit que $\det_{\mathcal{B}}$ est une forme anti-symétrique ; donc

$$\det_{\mathcal{B}}(\mathcal{B}') = -\det_{\mathcal{B}}\big(\tau(\mathcal{B})'\big). \tag{9.91}$$

Si l'un est positif, l'autre est négatif. Elles ont donc des orientations contraires. □

Corolaire 9.25.
Si $\mathcal{B}$ est une base de l'espace vectoriel E de dimension n, et si $\sigma \in S_n$, la base $\sigma(\mathcal{B})$ a même orientation que $\mathcal{B}$ si et seulement si $\sigma \in A_n$.

Démonstration. Notons c_1 la classe d'orientation de $\mathcal{B}$ et c_2 l'autre classe. La permutation σ se décompose en produit de transpositions dont la parité est fixée (proposition 1.201). Posons $\sigma = \tau_k \dots \tau_1$.

En posant $\mathcal{B}_0 = \mathcal{B}$ et $\mathcal{B}_{l+1} = \tau_{l+1}(\mathcal{B}_l)$, pour tout l, la base $\mathcal{B}_l$ est d'orientation contraire à celle de la base $\mathcal{B}_{l-1}$. Une base sur deux a l'orientation de $\mathcal{B}$ et l'autre sur deux a l'orientation contraire.

Donc $\sigma(\mathcal{B})$ a la même orientation que $\mathcal{B}$ si et seulement si k est pair. Mais $\sigma \in A_n$ si et seulement si k est pair. C'est bon. □

Proposition-Définition 9.26 ([225])**.**
Soit un espace vectoriel réel, et un endomorphisme f de E. Deux définitions.

(1) L'endomorphisme f est ***direct*** *si sont déterminant est strictement positif.*

(2) L'endomorphisme ***préserve l'orientation*** *si il transforme toutes base de E en une base de même orientation.*

Un endomorphisme est direct si et seulement si il préserve l'orientation.

12. Définition 9.21.
13. Nous notons (e_1, e_2) et non $\{e_1, e_2\}$ parce que l'ordre est important.
14. Définition 1.198.

Démonstration. En deux sens.

(i) **Direct implique préserve l'orientation** Soit une base $\mathcal{B}$ de E et un endomorphisme direct u. D'abord, u est inversible du fait que son déterminant est non nul par la proposition 9.9(2). Donc u transforme une base en une base par le lemme 4.7.

La définition 9.8 du déterminant de u est que

$$\det(u) = \det_{\mathcal{B}}\big(u(\mathcal{B})\big) > 0. \tag{9.92}$$

Donc $\mathcal{B}$ et $u(\mathcal{B})$ ont même orientation.

(ii) **Préserve l'orientation implique direct** Le fait que u préserve l'orientation signifie en particulier qu'il transforme une base en une base et qu'il est inversible par le lemme 4.7.

Donc si $\mathcal{B}$ est une base, $u(\mathcal{B})$ est encore une base et nous avons, parce que $\mathcal{B}$ et $u(\mathcal{B})$ ont même orientation,

$$0 < \det_{\mathcal{B}}\big(u(\mathcal{B})\big) = \det(u). \tag{9.93}$$

□

9.2.2 Cas affine

Définition 9.27.
Soit un espace affine $\mathcal{E}$ modelé sur E. Les repères cartésiens[15] *$(O,\mathcal{B})$ et $(O',\mathcal{B}')$ ont* ***même orientation*** *si les bases $\mathcal{B}$ et $\mathcal{B}'$ ont même orientation.*

Les classes d'équivalence (il y en a deux) sont les orientations de $\mathcal{E}$.

Une application affine $f\colon \mathcal{E} \to \mathcal{E}$ ***préserve l'orientation*** *si sa partie linéaire*[16] *préserve l'orientation.*

9.3 Hermitien, orthogonal, adjoint

9.28.
Une des choses à retenir de la définition de l'opérateur adjoint est que la notion de A^* dépend du produit scalaire considéré.

Il se fait que le plus souvent, sur $\mathbb{R}^n$, nous considérons le produit scalaire usuel et la base canonique. De ce fait, les notions d'opérateur adjoint et d'opérateur transposés se confondent avec la notion de matrice transposée. Ce sont pourtant, en général, trois notions distinctes.

Proposition-Définition 9.29 (Définition de la transposée[1])**.**
Soient deux espaces vectoriels euclidiens ou hermitiens E et F et une application linéaire $A\colon E \to F$.

(1) Il existe une unique application linéaire $B\colon F \to E$ telle que

$$\langle Ax, y\rangle_F = \langle x, By\rangle_E \tag{9.94}$$

pour tout $x \in E$ et $y \in F$.

(2) Si $\{e_i\}$ est une base orthonormée de E et $\{f_\alpha\}$ est une base orthonormée de F, alors la matrice de A et B pour ces bases sont liées par

$$B_{i\alpha} = A_{\alpha i}. \tag{9.95}$$

L'application B ainsi définie set nommée ***adjoint*** *de A et sera notée $B = A^*$.*

15. Définition 8.7.
16. Définition 8.13.

Démonstration. Pour l'unicité, nous écrivons la condition avec $x = e_j$ pour obtenir :

$$\langle Ae_j, y\rangle = \langle e_j, By\rangle = (By)_j \tag{9.96}$$

c'est-à-dire que les coefficients $B(y)_j$ de $B(y)$ dans la base canonique sont fixés par la condition.

Pour l'existence, il suffit de vérifier que poser

$$B(y) = \sum_j \langle Ae_j, y\rangle e_j \tag{9.97}$$

fonctionne. Pour cela il faut utiliser la bilinéarité du produit scalaire et le fait que $\langle x, e_j\rangle = x_j$. Nous avons :

$$\langle x, B(y)\rangle = \langle x, \sum_j \langle A(e_j), y\rangle e_j\rangle \tag{9.98a}$$

$$= \sum_j \langle A(e_j), y\rangle\langle x, e_j\rangle \tag{9.98b}$$

$$= \sum_j \langle A(x_j e_j), y\rangle \tag{9.98c}$$

$$= \langle A(x), y\rangle. \tag{9.98d}$$

En ce qui concerne la matrice de l'application B ainsi définie, nous écrivons la condition (9.94) avec $y = e'_\alpha$ et $x = e_i$, de telle sorte que

$$A(x) = A(e_i) = \sum_\beta A_{\beta i} e'_\beta \tag{9.99}$$

et

$$B(y) = B(e'_\alpha) = \sum_j B_{j\alpha} e_j. \tag{9.100}$$

Alors nous avons :

$$\sum_\beta A_{\beta i}\langle e'_\beta, e'_\alpha\rangle = \sum_j B_{j\alpha}\langle e_i, e_j\rangle, \tag{9.101}$$

donc

$$A_{\alpha i} = B_{i\alpha}. \tag{9.102}$$

□

9.30.
À cause de l'expression (9.95) pour la matrice de A^*, cette application est souvent appelé **transposé** de A et noté A^t. Nous allons cependant voir plus tard (définition9.176) que la transposée de A est une application $A^t\colon F^* \to E^*$. Il nous arrivera cependant d'écrire des égalités comme $\langle Ax, y\rangle = \langle x, A^t y\rangle$.

Proposition 9.31.
En ce qui concerne le déterminant,

$$\det(A^*) = \det(A)^* \tag{9.103}$$

où l'étoile à droite dénote la conjugaison complexe dans $\mathbb{C}$.

Démonstration. Écrivons l'expression explicite (9.7) du déterminant. Le tout avec la base canonique :

$$\det(A) = \det_{(e_1,\dots,e_n)}(Ae_1, \dots, Ae_n) = \sum_{\sigma\in S_n} \epsilon(\sigma)\prod_{i=1}^n e^*_{\sigma(i)}(Ae_i). \tag{9.104}$$

Mais nous pouvons développer :

$$e^*_{\sigma(i)}(Ae_i) = \langle e_{\sigma(i)}, Ae_i\rangle = \langle A^* e_{\sigma(i)}, e_i\rangle = \langle e_i, A^* e_{\sigma(i)}\rangle^* = e_i^*(A^* e_{\sigma(i)})^*. \tag{9.105}$$

Notez que dans la dernière expression, les trois * ont trois significations différentes. Par conséquent,

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} e_i^*(A^* e_{\sigma(i)})^*. \tag{9.106}$$

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 :

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} e_{\sigma(i)}^*(A^* e_i)^* = \det(A^*)^*. \tag{9.107}$$

Nous avons donc $\det(A) = \det(A^*)^*$, c'est-à-dire $\det(A)^* = \det(A^*)$. Pour information, la dernière étoile est la conjugaison complexe. □

Proposition 9.32 ([1]).
Si $A\colon E_2 \to E_3$ et $B\colon E_1 \to E_2$ sont des applications linéaires, alors

$$(AB)^* = B^* A^* \tag{9.108}$$

où la « multiplication » est la composition.

Démonstration. L'existence de $(AB)^*$, de A^* et de B^* ne donne pas lieu à débat parce que la proposition 9.29 ne souffre pas de discussions. La propriété que $(AB)^*$ est unique a avoir est que

$$\langle ABx, y\rangle = \langle x, (AB)^* y\rangle \tag{9.109}$$

pour tout $x \in E_1$ et $y \in E_3$. Or l'application B^*A^* possède également cette propriété parce que

$$\langle x, B^*A^*y\rangle = \langle Bx, A^*y\rangle = \langle ABx, y\rangle. \tag{9.110}$$

La partie unicité de la proposition 9.29 nous impose donc d'accepter que les applications $(AB)^*$ et B^*A^* sont en réalité les mêmes [17]. □

9.33.
Un grand moment d'utilisation de la notion d'adjoint pour un opérateur non carré sera la définition d'une intégrale sur une variété ; en particulier dans la proposition 20.9.

Lemme 9.34.
Si E est un espace euclidien, un endomorphisme $f\colon E \to E$ est autoadjoint si et seulement si pour tout $x, y \in E$ nous avons $\langle x, f(y)\rangle = \langle f(x), y\rangle$.

Démonstration. Dans le sens direct, nous avons

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle = \langle x, f(y)\rangle. \tag{9.111}$$

La première égalité est la définition de f^* et la seconde est l'hypothèse $f = f^*$.

Dans l'autre sens, l'hypothèse est que l'endomorphisme f vérifie $\langle x, f(y)\rangle = \langle f(x), y\rangle$. Mais la proposition 9.29(1) spécifie que f^* est l'unique endomorphisme à satisfaire cette égalité. Donc $f = f^*$. □

9.3.1 Opérateur orthogonal, matrice orthogonale

Définition 9.35.
*Un opérateur est **orthogonal** lorsque $A^* = A^{-1}$ où A^* est l'adjoint de A définit en 9.29.*

Définition 9.36.
*Une matrice U est **orthogonale** si $U^t = U^{-1}$. Le **groupe orthogonal** noté $\mathrm{O}(n)$ est l'ensemble des matrices orthogonales $n \times n$.*

17. Et ce même si vous croyez les avoir déjà vu ensemble dans la même pièce.

Lemme 9.37.
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[18].

Démonstration. Soit la base canonique $\{e_i\}_{i=1,\dots,n}$ de $\mathbb{R}^n$. Nous avons

$$\langle AA^* e_i, e_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}, \tag{9.112}$$

donc $\big((AA^*)e_i\big)_j = \delta_{ij}$, ou encore $(AA^*)_{ij} = \delta_{ij}$, ce qui signifie que la matrice AA^* est l'identité. □

Proposition 9.38 (Thème 75)**.**
À propos de matrices orthogonales.

(1) L'ensemble des matrices 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 linéaires*[19] *de* $\mathbb{R}^n$.

Démonstration. Si A et B sont orthogonales, alors

$$(AB)(AB)^t = ABB^tA^t = A\mathbb{1}A^t = \mathbb{1}. \tag{9.113}$$

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 9.31 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,

$$\langle A^t Ax, y\rangle = \langle x, y\rangle \tag{9.114}$$

pour tout $x, y \in \mathbb{R}^n$. En prenant $y = e_i$ nous trouvons

$$(A^t Ax)_i = x_i, \tag{9.115}$$

ce qui signifie que pour tout x, $A^t Ax = x$, ou encore que $A^t A$ est l'identité.

Réciproquement si $A^t A$ est l'identité nous avons

$$\langle x, y\rangle = \langle A^t Ax, y\rangle = \langle Ax, Ay\rangle, \tag{9.116}$$

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 $\mathrm{O}(3)$ et $\mathrm{SO}(3)$, ce sera pour la proposition 18.213 et ce qui s'ensuit.

Définition 9.39.
Le sous-groupe des matrices orthogonales de déterminant 1 *est le groupe* ***spécial orthogonal*** *noté* $\mathrm{SO}(n)$.

9.4 Topologie

9.4.1 Boules et sphères

Un espace vectoriel normé (définition 7.136) vient avec sa topologie métrique (théorème 7.98). Sphères et boules fermées viennent dans la définition 7.114.

Définition 9.40.
Une partie A de V est dite ***bornée*** *si il existe un réel R tel que $A \subset B(0_V, R)$.*

18. Définition 9.36.
19. Au sens où, parmi les applications linéaires, les isométries sont les éléments de $\mathrm{O}(n)$. À part ça, il y a aussi les translations, mais c'est une autre histoire qui vous sera contée une autre fois.

Une partie est donc bornée si elle est contenue dans une boule de rayon fini.

Exemple 9.41.
Dans $\mathbb{R}$, les boules sont les intervalles ouverts et fermés tandis que la sphère est donnée par les points extrêmes des intervalles :

$$\begin{aligned} B(a,r) &=]a-r, a+r[, \\ \bar{B}(a,r) &= [a-r, a+b], \\ S(a,r) &= \{a-r, a+r\}. \end{aligned} \tag{9.117}$$

△

Exemple 9.42.
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

$$|x| + |y| = r. \tag{9.118}$$

Pour la norme $\|.\|_2$, l'équation de la sphère de rayon r est

$$\sqrt{x^2 + y^2} = r, \tag{9.119}$$

et pour la norme supremum, la sphère de rayon r a pour équation

$$\max\{|x|, |y|\} = r. \tag{9.120}$$

Elles sont dessinées sur la figure 9.1

(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.

△

Proposition 9.43.
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 $d(Q,a) < r$.

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 », comme sur la figure suivante :

Plus précisément, nous considérons le point

$$P = x + \frac{v}{N} \tag{9.121}$$

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

$$d(P, x) = \|P - x\| = \frac{\|v\|}{N} \tag{9.122}$$

peut être rendu aussi petit que l'on veut par un choix approprié de N. Montrons maintenant que $d(a, P) > d(a, x)$:

$$\begin{aligned} d(a, P) &= \|a - x - \frac{v}{N}\| \\ &= \|a - x + \frac{a}{N} - \frac{x}{N}\| \\ &= \|\big(1 + \frac{1}{N}(a - x)\big)\| \\ &> \|a - x\| = d(a, x). \end{aligned} \tag{9.123}$$

Nous laissons en exercice le soin de trouver un point Q tel que $d(Q, a) < r$ et $d(Q, x) < \delta$. □

9.4.2 Ouverts, fermés, intérieur et adhérence

Définition 9.44.
*Soit $(V, \|.\|)$ un espace vectoriel normé et A, une partie de V. Un point a est dit **intérieur** à A si 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 9.45.
Trouver l'intérieur d'un intervalle dans $\mathbb{R}$ consiste à « ouvrir là où c'est fermé ».

(1) $\mathrm{Int}\big([0, 1[\big) =]0, 1[$.
Prouvons d'abord que $]0, 1[\subset \mathrm{Int}([0, 1[)$. 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 9.2). Cela prouve que a est dans l'intérieur de $[0, 1[$.

FIGURE 9.2 – 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 $\mathrm{Int}\big([0, 1[\big) \subset]0, 1[$. Vu que l'intérieur d'un ensemble est inclus dans l'ensemble, nous savons déjà que $\mathrm{Int}\big([0, 1[\big) \subset [0, 1[$. 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) $\mathrm{Int}\Big([0, \infty[\Big) =]0, \infty[$.

(3) $\mathrm{Int}\big(]2, 3[\big) =]2, 3[$.

△

Exemple 9.46.
Les intérieurs des boules et sphères sont importantes à savoir.

(1) $\mathrm{Int}\big(B(a,r)\big) = B(a,r)$. 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) $\mathrm{Int}\big(\bar{B}(a,r)\big) = B(a,r)$. 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 9.43.

(3) $\mathrm{Int}\big(S(a,r)\big) = \emptyset$. 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 9.47.
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 $A \subset \mathrm{Int}(A)$. Par convention, nous disons que l'ensemble vide $\emptyset$ 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 $\mathrm{Int}(A) = A$.

Définition 9.48.
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 agréables.

Exemple 9.49.
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 9.50.
Soit V un espace vectoriel normé.

(1) L'ensemble V lui-même et le vide sont à la fois fermés et ouverts.
(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 incluses 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[20], et l'union

$$\mathcal{O} = \bigcup_{i\in I} \mathcal{O}_i. \tag{9.124}$$

20. 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.

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 $(\mathcal{O}_k)_{k\in\{1,\dots,n\}}$, et $a \in \mathcal{O}$ où

$$\mathcal{O} = \bigcap_{k=1}^{n} \mathcal{O}_k. \tag{9.125}$$

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 $r = \min\{r_1,\dots,r_n\}$ 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

$$B(a,r) \subset \bigcap_{k=1}^{n} B(a,r_k) \subset \bigcap_{k=1}^{n} \mathcal{O}_k = \mathcal{O}, \tag{9.126}$$

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 :

$$\bigcap_{i=1}^{\infty}]-\frac{1}{n},\frac{1}{n}[= \{0\}. \tag{9.127}$$

Chacun des ensembles $]-\frac{1}{n},\frac{1}{n}[$ est ouvert, mais le singleton $\{0\}$ est fermé (pourquoi ?).

Nous reportons à la proposition 1.393 la preuve du fait que tout ensemble borné de $\mathbb{R}$ possède un infimum et un supremum.

Définition 9.51.
*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 9.52.
Les ensembles 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 $\mathbb{R}$:

— $]1,3[$;
— $]1,3]$;

— $[0,5[\setminus\{3\}$.

△

Proposition 9.53.
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 :

$$\bigcup_{n=1}^{\infty}[-1+\frac{1}{n},1-\frac{1}{n}]=]-1,1[. \tag{9.128}$$

Chacun des intervalles dont on prend l'union est fermé tandis que l'union est ouverte.

Lemme 9.54.
Soit A, une partie de l'espace vectoriel normé V. Un point $a\in V$ est adhérent[21] *à A dans V si et seulement si pour tout $\varepsilon>0$,*

$$B(a,\varepsilon)\cap A\neq\emptyset. \tag{9.129}$$

Un point peut être adhérent à A sans faire partie de A, et nous avons toujours $A\subset\mathrm{Adh}(A)$.

Exemple 9.55.
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 $\bar{A}\subset A$. Cela est fondamentalement le sens de la définition 7.28.

(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 9.49.

△

Exemple 9.56.
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 9.57.
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 9.58.
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 9.59.
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) $\overline{[0,2[}=[0,2]$.

21. Définition 7.28.

(2) $\overline{]3,\infty]} = [3,\infty[$.

△

Proposition 9.60.
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 9.61.
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 (1.1.5), 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 $a \in \operatorname{Int}(V\backslash A)$. Cela prouve la première affirmation.

Pour prouver la seconde affirmation, nous appliquons la première au complémentaire de A : $\complement(\overline{\complement A}) = \operatorname{Int}(\complement\complement A)$. En prenant le complémentaire des deux membres nous trouvons successivement

$$\begin{aligned} \complement\complement(\overline{\complement A}) &= \complement \operatorname{Int}(\complement\complement A), \\ \overline{\complement A} &= \complement \operatorname{Int}(A), \end{aligned} \tag{9.130}$$

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 9.62.
Soient A et B deux parties de l'espace vectoriel normé V.

(1) Pour les inclusions, si $A \subset B$, alors $\operatorname{Int}(A) \subset \operatorname{Int}(B)$ et $\bar{A} \subset \bar{B}$.

(2) Pour les unions, $\overline{A \cup B} = \bar{A} \cup \bar{B}$ et $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$.

(3) Pour les intersections, $\operatorname{Int}(A) \cap \operatorname{Int}(B) = \operatorname{Int}(A \cap B)$ et $\operatorname{Int}(A) \cup \operatorname{Int}(B) \subset \operatorname{Int}(A \cup B)$.

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 rayons ε_1 et ε_2 tels que

$$\begin{aligned} B(a, \varepsilon_1) \cap A &= \varnothing, \\ B(a, \varepsilon_2) \cap B &= \varnothing. \end{aligned} \tag{9.131}$$

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

$$\overline{\complement A \cup \complement B} = \overline{\complement A} \cup \overline{\complement B}. \tag{9.132}$$

En utilisant les propriétés du lemme 1.26, le membre de gauche devient

$$\overline{\complement A \cup \complement B} = \overline{\complement(A \cap B)} = \complement \operatorname{Int}(A \cap B), \tag{9.133}$$

tandis que le membre de droite devient

$$\overline{\complement A} \cup \overline{\complement B} = \complement \operatorname{Int}(A) \cup \complement \operatorname{Int}(A) = \complement\Big(\operatorname{Int}(A) \cap \operatorname{Int}(B)\Big). \tag{9.134}$$

En égalisant le membre de droite de (9.133) avec celui de (9.134) et en passant au complémentaire nous trouvons

$$\operatorname{Int}(A \cap B) = \operatorname{Int}(A) \cap \operatorname{Int}(B), \tag{9.135}$$

comme annoncé.

La dernière affirmation provient du fait que $\operatorname{Int}(A) \subset \operatorname{Int}(A \cup B)$ et de la propriété équivalente pour B.

□

Remarque 9.63.
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 9.64.
*La **frontière** d'un sous-ensemble A de l'espace vectoriel normé V est l'ensemble des points $a \in V$ tels que*

$$\begin{aligned} B(a, r) \cap A &\neq \varnothing, \\ B(a, r) \cap \complement A &\neq \varnothing, \end{aligned} \tag{9.136}$$

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 9.65.
La frontière d'une partie A d'un espace vectoriel normé V s'exprime sous la forme

$$\partial A = \bar{A} \backslash \operatorname{Int}(A). \tag{9.137}$$

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 9.65 est celle qu'en pratique nous utilisons le plus souvent. Dans certains textes, elle est prise comme définition de la frontière.

Lemme 9.66.
La frontière de A peut également s'exprimer des façons suivantes :

$$\partial A = \bar{A} \cap \complement \operatorname{Int}(A) = \bar{A} \cap \overline{\complement A}, \tag{9.138}$$

Démonstration. En partant de $\partial A = \bar{A}\backslash \operatorname{Int}(A)$, la première égalité est une application de la propriété (2) du lemme 1.26. La seconde égalité est alors la proposition 9.61. □

Exemple 9.67.
Dans $\mathbb{R}$, la frontière d'un intervalle est la paire constituée des points extrêmes. En effet

$$\partial[a,b[= \overline{[a,b[}\backslash \operatorname{Int}\big([a,b[\big) = [a,b]\backslash]a,b[= \{a,b\}. \tag{9.139}$$

Toujours dans $\mathbb{R}$ nous avons

$$\partial\mathbb{R} = \bar{\mathbb{R}}\backslash \operatorname{Int}(\mathbb{R}) = \mathbb{R}\backslash\mathbb{R} = \varnothing, \tag{9.140}$$

et

$$\partial\mathbb{Q} = \bar{\mathbb{Q}}\backslash \operatorname{Int}(\mathbb{Q}) = \mathbb{R}\backslash\varnothing = \mathbb{R}. \tag{9.141}$$

△

Exemple 9.68.
Dans $\mathbb{R}^n$, nous avons

$$\partial B(a,r) = \partial\bar{B}(a,r) = S(a,r). \tag{9.142}$$

Cela est un boulot pour la proposition 9.43. 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 $d(a,x) \geqslant r$. Nous avons donc $d(a,x) = r$. △

Remarque 9.69.
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 $\partial\bar{A} = \varnothing$.

9.4.3 Point isolé, point d'accumulation

Lemme 9.70.
Soit un espace vectoriel normé V ainsi qu'une partie D de V.

(1) Un point $a \in D$ est dit isolé[22] *dans D relativement à V si il existe un $\varepsilon > 0$ tel que*

$$B(a,\varepsilon) \cap D = \{a\}. \tag{9.143}$$

(2) Un point $a \in V$ point d'accumulation[23] *de D si pour tout $\varepsilon > 0$,*

$$\Big(B(a,\varepsilon)\backslash\{a\}\Big) \cap D \neq \varnothing. \tag{9.144}$$

Exemple 9.71.
Considérons la partie suivante de $\mathbb{R}^2$:

$$D = \{(x,y) \text{ tel que } x^2 + y^2 < 1\} \cup \{(1,1)\}. \tag{9.145}$$

Comme on peut le voir sur la figure 9.3, 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.

22. Définition 7.40.
23. Définition 7.39.

FIGURE 9.3 – L'ensemble décrit par l'équation (9.145). Le point P est un point isolé de D, tandis que les points S et Q sont des points d'accumulation.

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 9.72.
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 9.73.
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 9.74.
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.

9.4.4 Des exemples

Exemple 9.75.
Nous considérons l'ensemble.

$$A_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 - 5x + 6 < y \leqslant 2\}. \tag{9.146}$$

Si un point $(x, y) \in \mathbb{R}^2$ est tel que $x^2 - 5x + 6 < y$, alors dans une boule centrée en (x, y) (de rayon r_1), l'inégalité reste vraie (parce que la fonction $x^2 - 5x + 6 - y$ est une fonction continue). De la même manière, si nous avons $y < 2$ en (x, y), alors nous avons encore l'inégalité dans une boule de rayon r_2. En prenant $r = \min\{r_1, r_2\}$, les deux inégalités restent vraies dans la boule de rayon r.

Donc les points (x, y) tels que $x^2 - 5x + 6 < y < 2$ sont dans l'intérieur de A_1.

Pour les mêmes raisons, autour d'un point (x, y) tel que $x^2 - 5x + 6 > y$, nous pouvons trouver une boule dans laquelle l'inégalité reste stricte. Ces points ne sont donc pas dans l'adhérence de A_1. Un point qui vérifie $x^2 - 5x + 6 = y = 2$ est par contre dans l'adhérence parce que dans toute boule, on pourra trouver un x tel que $x^2 - 5x + 6 < y$, et un y. L'adhérence est donc donnée par les inéquations

$$\bar{A}_1 \equiv x^2 - 5x + 6 \leqslant y \leqslant 2. \tag{9.147}$$

La frontière est donnée par les points de l'adhérence qui ne sont pas dans l'intérieur de A_1. Attention : **ne pas dire** que la frontière est alors donnée simplement en remplaçant les inégalités par des égalités : $\partial A_1 \equiv x^2 - 5x + 6 = y = 2$. Quel est cet ensemble ?

Trouver la frontière demande un peu plus de travail. Le point marqué sur la figure 9.4 est sur la frontière parce que toute boule intersecte l'intérieur et l'extérieur. Cela est dû au fait que, sur

ce point, nous ayons $x^2 - 5x + 6 = y$ en même temps que $y < 2$. Donc si on prend une boule assez petite, on conserve $y < 2$, mais on obtient des points tels que $x^2 - 5x + 6 < y$.

En voyant le dessin, la chose à faire pour écrire la frontière est de trouver les deux points d'intersection entre la parabole et la droite horizontale. Ces points sont les points (x, y) qui satisfont au système

$$\begin{cases} x^2 - 5x + 6 = y & \text{(9.148a)} \\ y = 2. & \text{(9.148b)} \end{cases}$$

En substituant la seconde équation dans la première, il vient $x^2 - 5x + 6 = 2$, ce qui nous donne à résoudre le polynôme du second degré $x^2 - 5x + 4 = 0$. Les éventuelles solutions entières sont les diviseurs de 4. Par chance [24], on voit que $x = 1$ et $x = 4$ sont des solutions. Le théorème 3.148 nous assure qu'il n'y a pas d'autres racines. Les deux points d'intersection sont les points $P = (1, 2)$ et $Q = (4, 2)$. Les points de la frontière de A_1 sont donc donnés par

$$\begin{aligned} \partial A_1 = \{(x, y) \in \mathbb{R}^2 \text{ tels que } x^2 - 5x + 6 = y \text{ et } 1 \leqslant x \leqslant 4\} \\ \cup \{(x, y) \in \mathbb{R}^2 \text{ tels que } y = 2 \text{ et } 1 \leqslant x \leqslant 4\}. \end{aligned} \tag{9.149}$$

FIGURE 9.4 – En hachuré : l'intérieur ; en trait plein : la frontière. L'adhérence est l'union des deux. Exemple9.75.

Notez que les points de la parabole qui sont sur la frontière ne font pas partie de l'ensemble A_1 lui-même, tandis que ceux de la frontière qui sont sur la droite horizontale en font partie sauf $(4, 2)$ et $(1, 2)$.

L'intérieur de A_1 n'étant pas égal à A_1, cet ensemble n'est pas ouvert ; de la même manière, vu que $\bar{A}_1 \neq A_1$, l'ensemble n'est pas fermé. L'ensemble A_1 est par contre borné parce qu'il est contenu par exemple dans la boule de centre $(0, 0)$ et de rayon 5. Les points d'accumulation de A_1 sont les points de sa fermeture. △

Exemple 9.76.
Nous étudions

$$A_2 = \{(x, y) \in \mathbb{R}^2 \mid x + 1 < y < 2x\}. \tag{9.150}$$

Pour les mêmes raisons que dans l'exemple 9.75 l'intérieur est donné par

$$\text{Int}(A_2) \equiv x + 1 < y < 2x; \tag{9.151}$$

L'adhérence est donnée par

$$\overline{A_2} \equiv x + 1 \leqslant y \leqslant 2x, \tag{9.152}$$

Pour la frontière, les deux droites dont il est question dans la définition de A_2 (les droites $y = x + 1$ et $y = 2x$) se coupent en $x = 1$ (refaire soi-même le dessin de la figure 9.5). Lorsque $x < 1$, les conditions $x + 1 < y$ et $y < 2x$ sont incompatibles : aucun point de A_2 n'est dans la partie $x < 1$

24. Sinon, il aurait fallu utiliser la proposition 10.99.

du plan. Lorsque $x > 1$, alors les points situés *entre* les deux droites font partie de A_2. La frontière est donc donnée par ces deux droites pour $x \geqslant 1$.

Étant donné que $\mathrm{Int}(A_2) = A_2$, cet ensemble est ouvert (et donc pas fermé par la proposition 9.50(4)). Il n'est par contre pas borné parce qu'il contient des point (x, y) avec des x arbitrairement grands. △

FIGURE 9.5 – Notez que le point d'angle fait partie de la frontière, mais pas de l'ensemble. Exemple 9.76.

Proposition 9.77 ([226]).
Tout partie de $\mathbb{R}$ sans point d'accumulation est dénombrable.

Démonstration. Soit une partie S de $\mathbb{R}$ ne contenant pas de points d'accumulation. Pour chaque $s \in S$, il existe un $\epsilon_s > 0$ tel que

$$B(s, \epsilon_s) \cap S = \{s\}. \tag{9.153}$$

Rien ne garantit cependant que $B(s, \epsilon_s) \cap B(t, \epsilon_t) = \varnothing$, alors que nous en aurons besoin pour la suite.

Le nombre

$$\inf\{|s - v| \text{ tel que } v \in S\backslash\{s\}\} \tag{9.154}$$

est au moins égal à ϵ_s et est don strictement positif. Nous posons

$$r_s = \frac{\inf\big\{|s - v| \text{ tel que } v \in S\backslash\{s\}\big\}}{4}, \tag{9.155}$$

et

$$I_s = B(s, r_s). \tag{9.156}$$

Nous avons maintenant $I_s \cap I_t = \varnothing$. Soit en effet $u \in I_s \cap I_t$. Alors

$$|s - t| \leqslant |s - u| + |u - t| \tag{9.157a}$$
$$\leqslant r_s + t_t \tag{9.157b}$$
$$\leqslant 2r_s \tag{9.157c}$$

où nous avons supposé $r_s \geqslant r_t$. Si ce n'est pas le cas, changer s et t dans ce qui suit ; les deux point ont des rôles symétrique. Nous avons donc

$$|s - t| \leqslant 2r_s = \frac{\inf\big\{|s - v| \text{ tel que } v \in S\backslash\{s\}\big\}}{2} \leqslant \frac{|s - t|}{2}. \tag{9.158}$$

Donc $|s-t|=0$ et $s=t$.

Cela pour dire que I_s ne possède d'intersection avec I_t que si $s=t$.

Nous définissons alors une application

$$\begin{aligned}\varphi\colon S &\to \mathbb{Q}\\ s &\mapsto q_s\end{aligned} \tag{9.159}$$

où q_s est un choix de rationnel dans I_s. C'est le lemme 1.375 qui nous permet de choisir un tel rationnel.

La construction des intervalles I_s garantit que φ est une injection. Le fait qu'il existe une injection de $\mathbb{Q}$ vers S et le fait que $\mathbb{Q}$ est dénombrable impliquent que S est au plus dénombrable. □

9.5 Valeur propre et vecteur propre

9.5.1 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

$$Ae_1,\quad Ae_2\quad\text{et}\quad Ae_3. \tag{9.160}$$

La matrice d'une application A se forme en mettant simplement les vecteurs Ae_1, Ae_2 et Ae_3 en colonne. Donc la matrice

$$A=\begin{pmatrix}3&0&0\\0&0&1\\0&1&0\end{pmatrix} \tag{9.161}$$

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 :

$$\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+2y+3z\\4x+5y+6z\\7x+8y+9z\end{pmatrix}. \tag{9.162}$$

Une chose intéressante est de savoir quelles sont les directions invariantes de la transformation linéaire. Par exemple, on peut lire sur la matrice (9.161) 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 si 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

$$(A-\lambda\mathbb{1})\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}. \tag{9.163}$$

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.

9.5.2 Dans le vif du sujet

Définition 9.78.
*Soit un $\mathbb{K}$-espace vectoriel E et un endomorphisme $T\colon E \to E$. Un **vecteur propre** de T est un vecteur $v \neq 0$ tel que $T(v) = \lambda v$ pour un certain $\lambda \in \mathbb{K}$. Dans ce cas, λ est la **valeur propre** de v.*

*L'**espace propre** de T pour la valeur λ*[25] *est l'ensemble des vecteurs propres de T pour la valeur propre λ, et le vecteur nul.*

Définition 9.79.
*L'ensemble des valeurs propres de l'endomorphisme T est son **spectre** et est noté* $\mathrm{Spec}(T)$.

Remarque 9.80.
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 9.81.
Soient un espace vectoriel E, un endomorphisme $T \in \mathrm{End}(E)$, ainsi que ses sous-espaces propres $\{E_\lambda\}_{\lambda \in \mathrm{Spec}(T)}$. Toute somme finie de la forme

$$E_{\lambda_1} + \ldots + E_{\lambda_p} \tag{9.164}$$

est directe[26].

Démonstration. Nous utilisons le lemme 4.132. Soient $v_i \in E_{\lambda_i}$ un choix de vecteurs tels que

$$\sum_{i=1}^{p} v_i = 0. \tag{9.165}$$

Soit un entier j_0 entre 1 et p. Nous allons montrer que $v_{j_0} = 0$. Pour cela nous remarquons d'abord que, pour tout $i \neq j_0$,

$$\prod_{k \neq j_0} (\lambda_i - \lambda_k) = 0. \tag{9.166}$$

Nous appliquons l'opérateur $\prod_{k \neq j_0}(T - \lambda_k \mathbb{1})$ à l'égalité (9.165) :

$$0 = \sum_{i=1}^{p} \prod_{k \neq j_0} (T - \lambda_k) v_i \tag{9.167a}$$

$$= \sum_{i=1}^{p} \prod_{k \neq j_0} (\lambda_i - \lambda_k) v_i \tag{9.167b}$$

$$= \prod_{k \neq j_0} (\lambda_{j_0} - \lambda_k) v_{j_0}. \tag{9.167c}$$

Justifications.

— Pour (9.167b). Pour chaque k et i nous avons $(T - \lambda_k)v_i = Tv_i - \lambda_k v_i = \lambda_i v_i - \lambda_k v_i$ parce que v_i est un vecteur propre de T pour la valeur propre λ_i.

— Pour (9.167c). Dans la somme, seul le terme $i = j_0$ est non nul, à cause de (9.166).

Donc $v_{j_0} = 0$ parce que le produit $\prod_{k \neq j_0}(\lambda_{j_0} - \lambda_k)$, lui, est non nul. □

¡¡ Avertissement/question à la lectrice !! 9.82
L'énoncé de la proposition 9.83 me semble douteux, et la référence bibliographique est un lien mort.

Si vous savez un énoncé exact avec une preuve, écrivez-moi.

Proposition 9.83 ([227])**.**
Soit E, un espace vectoriel sur un corps infini et $(F_k)_{k=1,\ldots,r}$, des sous-espaces vectoriels de E tels que $\bigcup_{i=1}^{r} F_i = E$. Alors $E = F_k$ pour un certain k.

25. Nous laissons au lecteur le soin de vérifier que c'est bien un sous-espace vectoriel de E.
26. Définition 4.130.

9.6 Polynômes 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

$$\mathbb{M}(A) \to \mathbb{M}\big(A[X]\big). \tag{9.168}$$

9.6.1 Polynômes d'endomorphismes

Soit $u \in \mathrm{End}(E)$ où E est un $\mathbb{K}$-espace vectoriel. Nous considérons l'application

$$\begin{aligned} \varphi_u \colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(u). \end{aligned} \tag{9.169}$$

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 qu'endomorphisme de E.

Lemme 9.84.
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

$$(PQ)(u) = P(u) \circ Q(u). \tag{9.170}$$

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

$$\sum_l a_l b_{k-l}. \tag{9.171}$$

Par conséquent $(PQ)(u)$ contient $\sum_l a_l b_{k-l} u^k$. Par ailleurs $P(u) \circ Q(u)$ est donné par

$$\sum_i a_i u^i \left(\sum_j b_j u^j \right)(x) = \sum_{ij} a_i b_j u^{i+j}(x). \tag{9.172}$$

Le coefficient du terme en u^k est bien le même que celui donné par (9.171). □

Théorème 9.85 (Décomposition des noyaux ou lemme des noyaux[228])**.**
Soit u un endomorphisme du $\mathbb{K}$-espace[27] *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*[28]*. Alors*

$$E = \ker P_1(u) \oplus \dots \oplus \ker P_n(u). \tag{9.173}$$

De plus les projecteurs associés à cette décomposition sont des polynômes en u.

Démonstration. Dans ce qui suit, nous allons beaucoup utiliser le fait que $\mathbb{K}[X]$ soit commutatif (lemme 1.308). Nous posons

$$Q_i = \prod_{j \neq i} P_j. \tag{9.174}$$

(i) **Utilisation de Bézout** Par le lemme 6.45 ces polynômes sont étrangers entre eux et le théorème de Bézout (théorème 6.44) donne l'existence de polynômes R_i tels que

$$R_1 Q_1 + \dots + R_n Q_n = 1. \tag{9.175}$$

27. Le corps $\mathbb{K}$ est commutatif comme tous les corps dans le Frido.
28. Définition 3.133.

(ii) **Une première somme, pas directe** Si nous appliquons cette égalité à u et ensuite à $x \in E$ nous trouvons

$$\sum_{i=1}^{n}(R_iQ_i)(u)(x) = x, \tag{9.176}$$

et en particulier si nous posons $E_i = \text{Image}\big(R_iQ_i(u)\big)$ nous avons

$$E = \sum_{i=1}^{n} E_i. \tag{9.177}$$

Cette dernière somme n'est éventuellement pas une somme directe.

(iii) **Q_iQ_j est multiple de P** Si $i \neq j$, en utilisant la commutativité de $\mathbb{K}[X]$,

$$Q_iQ_j = \Big(\prod_{k\neq i} P_k\Big)\Big(\prod_{l\neq j} P_l\Big) = \Big(\prod_{\substack{k\neq i\\k\neq j}} P_k\Big)P_j\Big(\prod_{l\neq j} P_l\Big) = \prod_{\substack{k\neq i\\k\neq j}} P_k \prod_{k} P_l = S_{ij}P, \tag{9.178}$$

où S_{ij} est un polynôme. Nous voyons que Q_iQ_j est multiple de P.

(iv) **Une somme directe** Toujours avec $i \neq j$, en utilisant le lemme 9.84,

$$(R_iQ_i)(u) \circ (R_jQ_j)(u) = \big(R_iQ_iR_jQ_j\big)(u) = \big(R_iR_j \underbrace{Q_iQ_j}_{=S_{ij}P}\big)(u) = (R_iR_jS_{ij})(u) \circ P(u) = 0 \tag{9.179}$$

Nous pouvons voir E comme un $\mathbb{K}$-module et appliquer le théorème 1.288. Les opérateurs $R_iQ_i(u)$ ont l'identité comme somme et sont orthogonaux, et nous avons donc la décomposition en somme directe :

$$E = \bigoplus_{i=1}^{n} R_iQ_i(u)E. \tag{9.180}$$

(v) $R_iQ_i(u)E \subset \ker P_i(u)$ Attention : utilisation massive du lemme 9.84. Un élément de $R_iQ_i(u)E$ est de la forme $(R_iQ_i)(u)x$ avec $x \in E$. Nous appliquons l'endomorphisme $P_i(u)$ à cet élément, et nous vérifions que nous obtenons zéro :

$$\begin{aligned}
P_i(u)\big((R_iQ_i)(u)x\big) &= (P_iR_iQ_i)(u)x && (9.181a)\\
&= (R_i \underbrace{P_iQ_i}_{=P})(u)x && (9.181b)\\
&= (R_iP)(u)x && (9.181c)\\
&= \big(R_i(u) \circ \underbrace{P(u)}_{=0}\big)x && (9.181d)\\
&= 0. && (9.181e)
\end{aligned}$$

Par conséquent $\text{Image}(R_iQ_i(u)) \subset \ker P_i(u)$.

(vi) **Et la somme qu'il nous fallait** Le fait que la somme (9.180) soit directe n'est en fait pas crucial. En effet, vu que chacun des termes est inclus à $\ker P_i(u)$, nous avons la somme (pas directe à priori)

$$E = \sum_{i=1}^{n} R_iQ_i(u)E \subset \sum_{i=1}^{n} \ker P_i(u). \tag{9.182}$$

Mais cette fois, nous prouvons qu'elle est directe en utilisant la caractérisation du lemme 4.132(4). Supposons que, pour un certain k,

$$x \in \ker P_k(u) \cap \big(\sum_{j\neq k} \ker P_j(u)\big). \tag{9.183}$$

Nous allons montrer que $x = 0$.

(i) $\underline{Q_i(u)x = 0 \textbf{ si } i \neq k}$ Si $i \neq k$, nous avons

$$Q_i(u)x = \Big(\prod_{\substack{j\neq i\\ j\neq k}} P_j\Big)P_k(u)x = 0 \tag{9.184}$$

parce que $x \in \ker P_k(u)$.

(ii) $\underline{Q_k(u)x = 0}$ Nous savons qu'il existe $z_l \in \ker P_l(u)$ tel que $x = \sum_{l\neq k} z_l$. Nous avons alors

$$Q_k(u)x = \Big(\prod_{j\neq k} P_j\Big)\sum_{l\neq k} z_l = \sum_{l\neq k}\Big(\prod_{j\neq k} P_j(u)\Big)z_l = 0 \tag{9.185}$$

parce que parmi les $P_j(u)$ $(j \neq k)$, il y a $P_l(u)$ qui annule z_l.

(iii) **Et finalement** Nous avons prouvé que $Q_i(u)x = 0$ pour tout i. La formule de Bézout (9.175) donne alors

$$\sum_i R_i \subset Q_i(u)x_{=0} = x \tag{9.186}$$

et donc $x = 0$.

(vii) **Les projecteurs**

□

9.86.
Ce résultat est utilisé pour prouver que toute représentation est décomposable en représentations irréductibles, proposition 16.10 ainsi que pour le théorème 9.204 qui dit que si le polynôme minimal d'un endomorphisme est scindé à racine simple alors il est diagonalisable.

Corolaire 9.87.
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

$$F = \bigoplus_{i=1}^{r} \ker M_i^{\alpha_i}(f) \cap F \tag{9.187}$$

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 (9.85) s'applique et

$$E = E_1 \oplus \ldots \oplus E_r. \tag{9.188}$$

Nous pouvons décomposer $x \in F$ en termes de cette somme :

$$x = x_1 + \cdots + x_r \tag{9.189}$$

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 (9.189), $\texttt{proj}_i(x) = x_i$. Puisque $x \in F$, le membre de gauche est encore dans F et $x_i \in E_i \cap F$. Nous avons donc

$$F \subset \bigoplus_{i=1}^{r} F_i. \tag{9.190}$$

L'inclusion inverse est immédiate parce que $F_i \subset F$ pour chaque i. □

Lemme 9.88.
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 $P(u)$ pour la valeur propre $P(\lambda)$.

Démonstration. C'est un simple calcul de $P(u)x$ en ayant noté[29] $P(X) = \sum_{k=0}^{n} c_k X^n$:

$$P(u)x = \sum_{k=0}^{n} c_k u^k(x) = \sum_{k=0}^{n} c_k \lambda^k x = P(\lambda)x. \tag{9.191}$$

□

9.6.2 Polynôme minimal et minimal ponctuel

Nous avons déjà vu la définition de polynôme minimal en 6.60. Le lemme suivant permet de parler de polynôme minimal d'endomorphisme.

Lemme 9.89 ([1])**.**
Si E est un $\mathbb{K}$-espace vectoriel, l'ensemble $\mathrm{End}(E)$ *des endomorphismes de E est une extension du corps $\mathbb{K}$.*

Lemme 9.90.
Soit un endomorphisme $f\colon E \to E$ d'un $\mathbb{K}$-espace vectoriel de dimension finie. Il existe un unique polynôme annulateur unitaire de degré minimum[30]*.*

Tout endomorphisme de $\mathbb{K}$-espace vectoriel de dimension finie possède un polynôme minimal[31]*.*

Démonstration. Pour l'unicité, soient P et Q deux polynômes annulateurs de f de même degré minimum N et ayant tous deux 1 comme coefficient de x^N. Alors $P-Q$ est de degré $N-1$ tout en étant encore annulateur. Vu que nous avions dit que N était le degré minimum, le seul polynôme annulateur de degré $N-1$ est le polynôme nul. Donc $P-Q=0$.

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 $\sum_{k=0}^{N} a_k f^k = 0$. Le polynôme $P(X) = \sum_{k=0}^{N} a_k X^k$ 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.

L'existence d'un polynôme minimal est maintenant seulement dû au fait que, avec les notations de la définition 6.60, l'idéal I_f n'est pas réduit à $\{0\}$. □

Remarque 9.91.
La preuve donnée ci-dessus montre que $\deg(\mu) \leqslant \dim(E)^2$. Comme conséquence du théorème de Cayley-Hamilton 9.114 nous verrons qu'en réalité le degré du polynôme minimal est majoré par la dimension de l'espace.

Proposition 9.92 (Exemple en dimension infinie[1])**.**
L'endomorphisme de dérivation sur l'espace des fonctions dérivables $\mathbb{R} \to \mathbb{R}$ n'a pas de polynôme minimal.

Dans la suite, l'endomorphisme f du $\mathbb{K}$-espace vectoriel E de dimension n est fixé. Pour $x \in E$ nous notons

$$E_x = \{P(f)x \text{ tel que } P \in \mathbb{K}[X]\}. \tag{9.192}$$

Nous considérons le morphisme d'algèbres

$$\begin{aligned} \varphi\colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(f) \end{aligned} \tag{9.193}$$

29. En complète violation de ce qu'on disait dans 1.309.
30. Degré minimum au sens où il existe peut-être d'autres polynômes annulateurs, mais ils seront de degré plus élevé.
31. Définition 6.60.

et si $x \in E$ est donné nous considérons le morphisme de $\mathbb{K}$-espaces vectoriels

$$\begin{aligned} \varphi_x \colon \mathbb{K}[X] &\to E \\ P &\mapsto P(f)x. \end{aligned} \tag{9.194}$$

Les noyaux de ces applications sont des idéaux, entre autres par le lemme 9.84. Ils ont donc un unique générateur unitaire (chacun) par le théorème 6.40. En termes de vocabulaire, l'ensemble

$$\ker(\varphi) = \{P \in \mathbb{K}[X] \text{ tel que } P(f) = 0\} \tag{9.195}$$

est l'**idéal annulateur** de f et un polynôme P tel que $P(f) = 0$ est un polynôme annulateur de f.

Proposition-Définition 9.93.
La partie $\ker(\varphi_x)$ *est un idéal de* $\mathbb{K}[X]$ *qui possède un unique générateur unitaire.*

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. Puisque $\mu \in \ker(\varphi_x)$ pour tout x nous avons $\mu_x \mid \mu$ pour tout x.

Exemple 9.94 (Pas en dimension infinie)**.**
En dimension infinie, il n'y a pas toujours de polynôme annulateur. Si E est un espace vectoriel de dimension infinie 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 soit.

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 15.87 donnera un habillage à peine subtil à cet exemple. △

Proposition 9.95.
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 $\ker(\varphi) = \{Q \in \mathbb{K}[X] \text{ tel que } Q(f) = 0\}$ est un idéal par le lemme 9.84. Le polynôme minimal de f est un élément de degré plus bas dans I et par conséquent $I = (\mu_f)$ par le théorème 6.40. Nous concluons que μ_f divise tous les éléments de I. □

La proposition suivante permet de caractériser le polynôme minimal.

Proposition 9.96 ([113])**.**
Soit une application linéaire f *sur un* $\mathbb{K}$*-espace vectoriel. Il existe un unique polynôme unitaire*[32] $P \in \mathbb{K}[X]$ *tel que*

(1) $P(f) = 0$ *;*

(2) l'application

$$\begin{aligned} \varphi \colon \frac{\mathbb{K}[X]}{(P)} &\to \operatorname{End}(E) \\ \bar{Q} &\mapsto Q(f) \end{aligned} \tag{9.196}$$

est injective.

Démonstration. En ce qui concerne l'existence, il existe le polynôme minimal de f qui satisfait les conditions. Pour l'unicité nous y travaillons maintenant.

Supposons que l'application (9.196) 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 6.40(3) nous dit alors que $P = \mu$ parce que μ est également générateur unitaire. □

32. À mon avis, « unitaire » manque dans [113].

Lemme 9.97 ([229]).
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 9.85 nous donne la somme directe

$$E = \bigoplus_{i=1}^{r} \ker\big(P_i^{\alpha_i}(f)\big). \tag{9.197}$$

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 $Q(f)x_i \in E_j \cap E_i = \{0\}$ si $j \neq i$. 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

$$0 = \mu_a(f)a = \mu_a(f)x_1 + \dots + \mu_a(f)x_r. \tag{9.198}$$

Mais $\mu_a(f)x_i \in E_i$, et la somme des E_i est directe, donc l'annulation de la somme (9.198) 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 concluons que le produit divise encore μ_a :

$$\prod_{i=1}^{r} \mu_{x_i} \mid \mu_a, \tag{9.199}$$

c'est-à-dire $\mu \mid \mu_a$. Comme nous avons aussi $\mu_a \mid \mu$, nous déduisons $\mu_a = \mu$. □

Définition 9.98 (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** si 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 9.99.
Soit E un espace vectoriel de dimension finie et un endomorphisme cyclique[33] *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 9.95. Étant donné que v est cyclique, tout élément de E s'écrit sous la 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

$$P(f)x = \big(P(f) \circ Q(f)\big)v = \big(Q(f) \circ P(f)\big)v = 0 \tag{9.200}$$

où nous avons utilisé le lemme 9.84. □

Lemme 9.100 ([229]).
Soit $a \in E$ un vecteur cyclique pour f, tel que $\mu_a = \mu$. Alors E_a est un sous-espace stable par 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 $e_1 = a$, $e_2 = f(a)$,..., $e_l = f^{l-1}(a)$ forment une base de E_a. Nous pouvons la compléter en une

33. Voir la définition 9.98.

base $\{e_1, \dots, e_n\}$ de E. Et nous posons[34]

$$G = \{x \in E \text{ tel que } e_l^*\big(f^k(x)\big) = 0, \forall k \geqslant 0\} \tag{9.201a}$$
$$= \bigcap_{k\geqslant 0} \ker\{e_l^* \circ f^k\} \tag{9.201b}$$
$$= \bigcap_{k=0}^{l-1} \ker(e_l^* \circ f^k). \tag{9.201c}$$

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

$$\sum_{j=0}^{l-1} \lambda_j \big(e_l^* \circ f^j\big) = 0 \tag{9.202}$$

et montrons que $\lambda_j = 0$ pour tout j est l'unique solution. Soit $x \in E$ et appliquons l'opération (9.202) au vecteur $f^i(x)$; le résultat est zéro :

$$0 = \sum_{j=0}^{l-1} \lambda_j (e_l^* \circ f^i \circ f^j) = (e_l^* \circ f^i) P(u) \tag{9.203}$$

où nous avons posé $P(X) = \sum_{j=0}^{l-1} \lambda_j X^j$. Appliquons cela à a : pour tout i nous avons

$$(e_l^* \circ f^i)\big(P(f)a\big) = 0. \tag{9.204}$$

Mais par définition de E_a, l'élément $P(f)a$ est dans E_a. Nous en déduisons que

$$P(f)a \in G \cap E_a = \{0\}, \tag{9.205}$$

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 9.101.
L'endomorphisme f d'un espace vectoriel est ***semi-simple*** *si tout sous-espace stable par f possède un supplémentaire stable.*

Lemme 9.102.
Si le polynôme minimal d'un endomorphisme est irréductible, alors cet endomorphisme est semi-simple[35].

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 problème. Sinon nous considérons $u_1 \in E \backslash F$ et

$$E_{u_1} = \{P(f)u_1 \text{ tel que } P \in \mathbb{K}[X]\}, \tag{9.206}$$

qui est un espace stable par f.

34. ici, comme presque partout, e_l^* est le dual de e_l, c'est-à-dire l'application linéaire sur E donnée par $e_l^*(e_i) = \delta_{li}$, voir la définition 4.120.
35. Définition 9.101.

Montrons que $E_{u_1} \cap F = \{0\}$. Pour cela nous étudions l'idéal

$$I_{u_1} = \{P \in \mathbb{K}[X] \text{ tel que } P(f)u_1 = 0\}. \tag{9.207}$$

C'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 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 $y \neq 0$, cela 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 des coefficient de Bézout (théorème 6.44) $A, B \in \mathbb{K}[X]$ tels que

$$AP + BP_{u_1} = 1. \tag{9.208}$$

Nous appliquons cette égalité à f et puis à u_1 :

$$u_1 = A(f) \circ \underbrace{P(f)u_1}_{=y} + B(f) \circ \underbrace{P_{u_1}(u_1)}_{=0} = A(f)y. \tag{9.209}$$

Mais $y \in F$, donc $A(f)y \in F$. Nous aurions donc $u_1 \in F$, ce qui est impossible par choix. Nous savons 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

$$E = F \oplus E_{u_1} \oplus \ldots \oplus E_{u_k} \tag{9.210}$$

où chacun des E_{u_i} sont stables. □

Théorème 9.103.
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 $\mu_f = M^2 N$ où l'on a noté $M = M_i$. Nous étudions l'espace

$$F = \ker M(f) \tag{9.211}$$

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 $M(f)$,

$$M(f)\big(MN(f)x\big) = \mu_f(f)x = 0, \tag{9.212}$$

ce qui signifie que $MN(f)x \in F$. Mais puisque S est stable par f nous avons aussi $MN(f)x \in S$. Finalement $MN(f)x \in F \cap S = \{0\}$. Autrement dit, $MN(f)$ s'annule sur S.

Prenons maintenant $y \in F$. Nous avons

$$MN(f) = N(f)\big(M(f)y\big) = 0 \tag{9.213}$$

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 $\mu_f = M^2 N$.

Nous passons au sens inverse. Soit $m_f = M_1 \ldots M_r$ 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

$$E_i = \ker(M_i(f)) \tag{9.214}$$

et $f_i = f|_{E_i}$. Par le lemme 9.87 nous avons

$$F = \bigoplus_{i=1}^{r}(F \cap E_i). \tag{9.215}$$

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 polynôme 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 9.102.

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 :

$$E_i = S_i \oplus (F \cap E_i). \tag{9.216}$$

Étant donné que sur chaque S_i nous avons $f|_{S_i} = f_i$, l'espace $S = S_1 \oplus \ldots \oplus S_r$ est stable par f. Par conséquent, nous avons

$$\begin{aligned}
E &= E_1 \oplus \ldots \oplus E_r && (9.217a)\\
&= \big(S_1 \oplus (F \cap E_1)\big) \oplus \ldots \oplus \big(S_r \oplus (F \cap E_r)\big) && (9.217b)\\
&= \big(\bigoplus_{i=1}^{r} S_i\big) \oplus \big(\bigoplus_{i=1}^{r} F \cap E_i\big) && (9.217c)\\
&= S \oplus F, && (9.217d)
\end{aligned}$$

ce qui montre que F a bien un supplémentaire stable par f et donc que f est semi-simple. □

Exemple 9.104 (L'espace engendré par $\mathbb{1}$, A, A^2,...)**.**
Soit A une matrice, et

$$E = \mathrm{Span}\{A^k \text{ tel que } k \in \mathbb{N}\}. \tag{9.218}$$

Nous montrons que $\dim(E)$ est le degré du polynôme minimal de A.

D'abord l'idéal annulateur de A est engendré par le polynôme minimal[36] que nous notons $\mu = \sum_{k=0}^{p} a_k X^k$. La partie $\{\mathbb{1}, \ldots, A^{p-1}\}$ est libre parce qu'une combinaison linéaire nulle de ces éléments serait un polynôme annulateur en A de degré plus petit que p. Donc $\dim(E) \geqslant p$.

La partie $\{\mathbb{1}, A, \ldots, A^p\}$ 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 $\{\mathbb{1}, A, \ldots, A^{p-1}\}$ est génératrice (alors $\dim(E) \leqslant p$). Soit un entier $q \geqslant p$ et de division euclidienne[37] $np + r = q$ avec $r < p$. Nous avons $A^q = A^{np}A^r$. D'une part

$$A^{np} = (A^p)^n = f(A)^n \tag{9.219}$$

est de degré $n(p-1)$. Par conséquent

$$A^q = f(A)^n A^r \tag{9.220}$$

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 $\{\mathbb{1}, A, \ldots, A^{p-1}\}$ est génératrice de E.

La dimension de E est donc p, le degré du polynôme minimal. △

Proposition 9.105.
Soit f un endomorphisme d'un espace vectoriel de dimension finie. Nous avons l'isomorphisme d'espace vectoriel

$$\mathbb{K}[f] \simeq \frac{\mathbb{K}[X]}{(\mu_f)} \tag{9.221}$$

La dimension en est $\deg(\mu_f)$.

36. Proposition 9.95.
37. Théorème 1.224.

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

$$\bar{P} = \{P + S\mu\}_{S\in\mathbb{K}[X]}. \tag{9.222}$$

Nous allons montrer que l'application suivante fournit l'isomorphisme :

$$\begin{aligned}\psi\colon \frac{\mathbb{K}[X]}{(\mu)} &\to \mathbb{K}[f]\\ \bar{P} &\mapsto P(f).\end{aligned} \tag{9.223}$$

(i) <u>ψ **est bien définie**</u> Si $Q \in \bar{P}$ alors $Q = P + S\mu$ pour un certain $S \in \mathbb{K}[X]$. Du coup nous avons

$$\psi(\bar{Q}) = P(f) + (S\mu)(f). \tag{9.224}$$

Mais $\mu(f) = 0$ donc le deuxième terme est nul. Donc $\psi(\bar{P})$ est bien défini.

(ii) <u>**Injectif**</u> 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 $\bar{P} = 0$.

(iii) <u>**Surjectif**</u> 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 corolaire 6.41 en parle déjà : une base est donnée par les projections de $1, X, \dots, X^{\deg(\mu_f)-1}$. □

9.6.3 Polynôme caractéristique

Définition 9.106.
Soit un anneau commutatif A. Si $u \in \mathbb{M}(n, A)$, nous définissons le ***polynôme caractéristique de*** *u :*

$$\chi_u(X) = \det(u - X\mathbb{1}_n). \tag{9.225}$$

Nous définissons de même le polynôme caractéristique d'un endomorphisme $u\colon E \to E$.

Remarque 9.107.
Quelques remarques à propos du signe[38].

— Certains auteurs définissent le polynôme caractéristique par $\det(X-u)$ au lieu de $\det(u-X)$.

— Wikipédia francophone[230] prend la définition $\det(X-u)$ (donc opposée de la notre). Allez lire la page de discussion.

— Sur les wikipédias en d'autres 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 9.108.
Le polynôme caractéristique χ_u est unitaire en dimension paire et a pour degré la dimension de l'espace vectoriel E.

Théorème 9.109.
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.

38. Attention : je crois qu'il y a des incohérences dans le Frido à propos de ce choix

Théorème 9.110.
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 9.9(1) ou encore que λ est une racine du polynôme caractéristique de u.

(2) $\Leftrightarrow$ (3). C'est une application directe du théorème 9.109 qui précise que le polynôme caractéristique a les mêmes racines dans $\mathbb{K}$ que le polynôme minimal. □

Exemple 9.111.
Sur $\mathbb{R}^2$, nous considérons la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ qui a pour polynôme caractéristique[39] 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 :

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.226}$$

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[40].

Proposition 9.112 ([231])**.**
Soit un espace vectoriel E de dimension finie pour lequel il existe un endomorphisme $f\colon E \to E$ tel que $(f \circ f)(x) = -x$ pour tout $x \in E$. Alors la dimension de E est paire.

Démonstration. Cherchons les valeurs propres de f en résolvant l'équation $f(x) = \lambda x$. Nous appliquons f à cette égalité :

$$-x = \lambda f(x) = \lambda^2 x. \tag{9.227}$$

Donc λ ne peut pas être réel. Nous avons montré que f n'a pas de valeur propre réelle. 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 E est n nous avons :

$$\det(f^2) = \det(-\mathrm{Id}) = (-1)^n. \tag{9.228}$$

Donc $(-1)^n$ est positif, ce qui montre que n est pair.

Proposition 9.113 ([227])**.**
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

$$p_{f,x} = \dim E_{f,x} = \deg(\mu_{f,x}) \tag{9.229}$$

où $\mu_{f,x}$ est le générateur unitaire de $I_{f,x}$.

39. Définition 9.106.
40. Définition 9.106.

(3) Le polynôme caractéristique de $f|_{E_{f,x}}$ est $\mu_{f,x}$.

(4) Nous avons

$$\chi_{f|_{E_{f,x}}}(f)x = \mu_{f,x}(f)x = 0. \tag{9.230}$$

Démonstration. Le fait que $E_{f,x}$ soit stable par f est classique. Le point (4) est 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

$$B = \{f^k(x) \text{ tel que } 0 \leqslant k \leqslant p_{f,x} - 1\} \tag{9.231}$$

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

$$\mu_{f,x}(X) = X^{p_{f,x}} - \sum_{i=0}^{p_{f,x}-1} a_i X^i. \tag{9.232}$$

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

$$0 = f^{p_{f,x}}(x) - \sum_{i=0}^{p_{f,x}-1} a_i f^i(x) \tag{9.233}$$

nous trouvons

$$f^{p_{f,x}+k}(x) = \sum_{i=0}^{p_{f,x}-1} a_i f^{i+k}(x), \tag{9.234}$$

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

$$\mu_{f,x}(f)x = 0. \tag{9.235}$$

En y appliquant f^k et en profitant de la commutativité des polynômes sur les endomorphismes (proposition 9.84), nous avons

$$0 = f^k\big(\mu_{f,x}(f)x\big) = \mu_{f,x}(f)f^k(x), \tag{9.236}$$

de telle sorte que $\mu_{f,x}(f)$ est nul sur B et donc est nul sur $E_{f,x}$. Autrement dit,

$$\mu_{f,x}\big(f|_{E_{f,x}}\big) = 0. \tag{9.237}$$

Montrons que $\mu_{f,x}$ est même minimal pour $f|_{E_{f,x}}$.

Supposons avoir Q, un polynôme non nul de degré $p_{f,x} - 1$ annulant $f|_{E_{f,x}}$. En particulier $Q(f)x = 0$. Cela signifie que B est un système lié, alors que nous avons montré que c'était un système libre. Contradiction. 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 9.114 (Cayley-Hamlilton)**.**
Le polynôme caractéristique est un polynôme annulateur.

Une démonstration plus simple via la densité des diagonalisables est donnée en théorème 13.25.

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 $f|_{E_{f,x}}$ divise χ_f, c'est-à-dire qu'il existe un polynôme Q_x tel que

$$\chi_f = Q_x \chi_{f,x}, \tag{9.238}$$

et donc aussi

$$\chi_f(f)x = Q_x(f)\big(\chi_{f,x}(f)x\big) = 0 \tag{9.239}$$

parce que la proposition 9.113 nous indique que $\chi_{f,x}$ est un polynôme annulateur de $f|_{E_{f,x}}$. □

Corolaire 9.115.
Le degré du polynôme minimal est majoré par la dimension de l'espace.

Démonstration. Le polynôme minimal divise le polynôme caractéristique parce qu'il engendre l'idéal des polynômes annulateurs par la proposition 9.95. Or le degré du polynôme caractéristique est la dimension de l'espace par le lemme 9.108. □

Exemple 9.116 (Calcul de l'inverse d'un endomorphisme)**.**
Le théorème de Cayley-Hamilton donne un moyen de calculer l'inverse d'un endomorphisme inversible pourvu que l'on connaisse son polynôme caractéristique. En effet, supposons que

$$\chi_f(X) = \sum_{k=0}^{n} a_k X^k. \tag{9.240}$$

Nous aurons alors

$$0 = \chi_f(f) = \sum_{k=0}^{n} a_k f^k. \tag{9.241}$$

Nous appliquons f^{-1} à cette dernière égalité en sachant que $f^{-1}(0) = 0$:

$$0 = a_0 f^{-1} + \sum_{k=1}^{n} a_k f^{k-1}, \tag{9.242}$$

et donc

$$f^{-1} = -\frac{1}{\det(f)} \sum_{k=1}^{n} a_k f^{k-1} \tag{9.243}$$

où nous avons utilisé le fait que $a_0 = \chi_f(0) = \det(f)$. △

Proposition 9.117.
Si $(X - z)^l$ $(l \geqslant 1)$ est la plus grande puissance de $(X - z)$ dans le polynôme caractéristique d'un endomorphisme f alors

$$1 \leqslant \dim(E_z) \leqslant l. \tag{9.244}$$

C'est-à-dire que nous avons au moins un vecteur propre pour chaque racine du polynôme caractéristique.

Démonstration. Si $(X - z)$ divise χ_f alors en posant $\chi_f = (X - z)P(X)$ nous avons

$$\det(f - X\mathbb{1}) = (X - z)P(X), \tag{9.245}$$

ce qui, évalué en $X = z$, donne $\det(f - z\mathbb{1}) = 0$. L'annulation du déterminant étant équivalente à l'existence d'un noyau non trivial, nous avons $v \neq 0$ dans E tel que $(f - z\mathbb{1})v = 0$. Cela donne $f(v) = zv$ et montre que v est vecteur propre de f pour la valeur propre z. Et aussi que $\dim(E_z) \geqslant 1$.

Si $\dim(E_z) = k$ alors le théorème de la base incomplète 4.12 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 f est de la forme

$$\begin{pmatrix} z & & & * \\ & \ddots & & \vdots \\ & & z & * \\ & & & * \end{pmatrix} \tag{9.246}$$

où les étoiles représentent des blocs à priori non nuls. En tout cas, sous cette forme, il est visible que $(X - z)^k$ divise χ_f. □

9.7 Formes bilinéaires et quadratiques

Plus à propos de formes bilinéaires dans le thème 14.

Définition 9.118 ([232]).
*Soient trois espaces vectoriels E, F et V sur le même corps commutatif $\mathbb{K}$. Une application $b\colon E \times F \to V$ est **bilinéaire** si elle est séparément linéaire en ses deux variables, c'est-à-dire si*

(1) $b(u_1 + u_2, v) = b(u_1, v) + b(u_2, v)$,

(2) $b(u, v_1 + v_2) = b(u, v_1) + b(u, v_2)$

(3) $b(\lambda u, v) = b(u, \lambda v) = \lambda b(u, v)$

pour tout $u, u_1, u_2 \in E$, $v, v_1, v_2 \in F$ et pour tout $\lambda \in \mathbb{K}$.

*Dans le cas $E = F$ et $V = \mathbb{K}$, nous parlons de **forme bilinéaire** sur E.*

*Nous parlons de forme bilinéaire **symétrique** si de plus $b(u, v) = b(v, u)$.*

9.119.
Une application bilinéaire $E \times E \to \mathbb{K}$ n'est pas une application linéaire ; la distinction est importante. La linéarité est

$$b(\lambda u, \lambda v) = b\big(\lambda(u, v)\big) = \lambda b(u, v) \tag{9.247}$$

et la bilinéarité est

$$b(\lambda u, v) = b(u, \lambda v) = \lambda b(u, v). \tag{9.248}$$

En réalité la seule forme qui soit à la fois linéaire et bilinéaire est la forme identiquement nulle : la condition

$$b(\lambda u, \lambda v) = \lambda^2 b(u, v) = \lambda b(u, v) \tag{9.249}$$

pour tout $\lambda \in \mathbb{K}$ implique $b(u, v) = 0$.

Exemple 9.120 ([233]).
L'application

$$\begin{aligned} b\colon \mathbb{M}(n, \mathbb{K}) \times \mathbb{M}(n, \mathbb{K}) &\to \mathbb{K} \\ (A, B) &\mapsto \operatorname{Tr}(AB) \end{aligned} \tag{9.250}$$

est une forme bilinéaire symétrique.

La vérification est un calcul :

$$\operatorname{Tr}(BA) = \sum_i (BA)_{ii} = \sum_{ik} B_{ik} A_{ki} = \sum_{ik} A_{ki} A_{ik} = \sum_k (AB)_{kk} = \operatorname{Tr}(AB). \tag{9.251}$$

△

9.7.1 Dégénérescence d'une forme bilinéaire

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 9.121.
Une forme bilinéaire est ***non dégénérée*** $b(x,z)=0$ *pour tout* z *implique* $x=0$.

Lemme 9.122.
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

$$b(x-y,z)=0 \tag{9.252}$$

pour tout z, ce qui implique $x-y=0$. □

9.7.2 Topologie

La topologie considérée sur $Q(E)$ est celle de la norme

$$N(q)=\sup_{\|x\|_E=1}|q(x)|, \tag{9.253}$$

qui du point de vue de $S(n,\mathbb{R})$ est

$$N(A)=\sup_{\|x\|_E=1}|x^tAx|. \tag{9.254}$$

Notons que à droite, c'est la valeur absolue usuelle sur $\mathbb{R}$.

Proposition 9.123.
Soit $\{e_i\}$ *une base de* E*. L'application*

$$\begin{aligned}\phi\colon Q(E)&\to S(n,\mathbb{R})\\ q&\mapsto \big(b(e_i,e_j)\big)_{i,j}\end{aligned} \tag{9.255}$$

où b *est forme bilinéaire associée à* q *est une bijection linéaire et continue*[41].

Démonstration. Si $\phi(q)=\phi(q')$; alors

$$q(x)=\sum_{i,j}\phi(q)_{ij}x_ix_j=\sum_{i,j}\phi(q')_{ij}x_ix_j=q'(x). \tag{9.256}$$

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

$$q(x)=\sum_{i,j}B_{ij}x_ix_j \tag{9.257}$$

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

$$\sup_{\|x\|=1}|q_n(x)|\to 0. \tag{9.258}$$

41. Pour les topologies des normes (9.253) et (9.254).

Nous rappelons l'identité de polarisation :

$$b_n(x,y) = \frac{1}{2}\big(q_n(x-y) - q(x) - q(y)\big). \tag{9.259}$$

En ce qui concerne deux des trois termes, il n'y a pas de problèmes :

$$\big|\phi(q_n)_{ij}\big| = \big|b_n(e_i,e_j)\big| \leqslant \frac{1}{2}\big|b_n(e_i-e_j)\big| + \frac{1}{2}\big|q_n(e_i)\big| + \frac{1}{2}\big|q_n(e_j)\big|. \tag{9.260}$$

Si n est assez grand, nous avons tout de suite

$$\big|\phi(q_n)_{ij}\big| \leqslant \frac{1}{2}\big|q_n(e_i-e_j)\big| + \epsilon. \tag{9.261}$$

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

$$q_n(e_i - e_j) = \alpha_{ij}^2 q_n(e_{ij}) \leqslant \alpha^2 q_n(e_{ij}). \tag{9.262}$$

Il suffit maintenant de prendre n assez grand pour avoir $\sup_{\|x\|=1}|q_n(x)| \leqslant \frac{\epsilon}{\alpha^2}$ pour avoir

$$\big|\phi(q_n)_{ij}\big| \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{\alpha^2}. \tag{9.263}$$

□

9.7.3 Isotropie

Définition 9.124 (Isotropie)**.**
*Un vecteur est **isotrope** pour b si 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 :*

$$C(b) = \{x \in E \text{ tel que } b(x,x) = 0\}. \tag{9.264}$$

Nous introduisons quelques notations. D'abord pour $y \in E$ nous notons

$$\begin{aligned} \Phi_y \colon E &\to \mathbb{R} \\ x &\mapsto b(x,y) \end{aligned} \tag{9.265}$$

et ensuite

$$\begin{aligned} \Phi \colon E &\to E^* \\ y &\mapsto \Phi_y. \end{aligned} \tag{9.266}$$

Définition 9.125.
*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*

$$\ker(b) = \{z \in E \text{ tel que } b(z,y) = 0\,\forall y \in E\}. \tag{9.267}$$

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.

9.8 Formes bilinéaires et quadratiques

Plus à propos de formes bilinéaires dans le thème 14.

9.8.1 Généralités

Définition 9.126 ([234]).
Soit un espace vectoriel E et $\mathbb{F}$ un corps de caractéristique différente de 2. Une ***forme quadratique*** *sur E est une application $q\colon E \to \mathbb{F}$ pour laquelle il existe une forme bilinéaire symétrique $b\colon E \times E \to \mathbb{F}$ satisfaisant $q(x) = b(x,x)$ pour tout $x \in E$.*

L'ensemble des formes quadratiques réelles sur E est noté $Q(E)$.

Définition 9.127 (Application bilinéaire définie positive, thème 37).
Si b est une application bilinéaire[42] *sur un espace vectoriel E nous disons qu'elle est*

(1) ***définie positive*** *si $b(x,x) \geqslant 0$ pour tout $x \in E$ et $b(x,x) = 0$ si et seulement si $x = 0$.*

(2) ***semi-définie positive*** *si $b(x,x) \geqslant 0$ pour tout $x \in E$. Nous dirons aussi parfois qu'elle est simplement « positive ».*

Cela est évidemment à lier à la définition 9.215 et à la proposition 9.220 : une application bilinéaire est définie positive si et seulement si sa matrice symétrique associée l'est.

Proposition 9.128.
*Soit une forme bilinéaire b et la forme quadratique associée q. Alors nous avons l'****identité de polarisation*** *:*

$$b(x,y) = \frac{1}{2}\big(q(x) + q(y) - q(x-y)\big). \tag{9.268}$$

Démonstration. Il suffit de substituer dans le membre de droite $q(x) = b(x,y)$ et d'utiliser la bilinéarité :

$$\begin{aligned} q(x) + q(y) - q(x-y) &= b(x,x) + b(y,y) - b(x-y,x-y) && (9.269a)\\ &= b(x,x) + b(y,y) - b(x) + b(x,y) + b(y,x) - b(y,y) && (9.269b)\\ &= 2b(x,y) && (9.269c) \end{aligned}$$

où nous avons utilisé le fait que b est symétrique : $b(x,y) = b(y,x)$. □

Lemme 9.129.
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'identité de polarisation de la proposition 9.128. 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.

Définition 9.130.
Soit une forme quadratique q sur E. Nous disons que $v, w \in E$ sont ***q-orthogonaux*** *si $b(v,w) = 0$ la forme bilinéaire b associée à q par le lemme 9.129.*

9.8.2 Matrice associée à une forme bilinéaire

Définition 9.131.
Soit une forme bilinéaire[43] *$b\colon E \times E \to \mathbb{K}$ et une base quelconque $\{f_\alpha\}$ de E. Nous définissons les nombres*

$$B_{\alpha\beta} = b(f_\alpha, f_\beta), \tag{9.270}$$

qui forment une matrice symétrique dans $\mathbb{M}(n,\mathbb{K})$. Cette matrice est la ***matrice associée*** *à la forme bilinéaire b.*

La matrice d'une forme quadratique est celle associée à sa forme bilinéaire associée.

42. Définition 9.118.
43. Définition 9.118.

Lemme 9.132.
Soit une forme bilinéaire $b\colon E \times E \to \mathbb{K}$ et une base quelconque $\{f_\alpha\}$ de E. Nous notons B la matrice de b (definition 9.131) et q la forme quadratique associée.

Alors nous avons

$$b(x, y) = \sum_{\alpha\beta} B_{\alpha\beta} x_\alpha y_\beta. \tag{9.271}$$

et

$$b(x, y) = x \cdot By. \tag{9.272}$$

où le point est le produit scalaire usuel (composante par composante).

Démonstration. Si $x = \sum_\alpha x_\alpha f_\alpha$ et $y = \sum_\beta y_\beta f_\beta$:

En utilisant la convention (4.83) et les choses autour (voir aussi -2.1),

$$b(x, y) = \sum_\alpha x_\alpha \sum_\beta B_{\alpha\beta} y_\beta = \sum_\alpha x_\alpha (By)_\alpha = x \cdot By. \tag{9.273}$$

□

9.8.3 Changement de base : matrice d'une forme bilinéaire

Proposition 9.133 (Voir la section -2.1)**.**
Soit une forme bilinéaire[44] *$b\colon V \times V \to \mathbb{K}$ dont la matrice*[45] *dans la base $\{e_i\}$ est A et celle dans la base $\{f_\alpha\}$ est B. Nous supposons que les bases sont liées par $f_\alpha = \sum_i Q_{i\alpha} e_i$. Alors*

$$B = Q^t A Q. \tag{9.274}$$

Démonstration. Soit $x, x' \in V$ de coordonnées (x_i) et (x'_i) dans la base $\{e_i\}$ et (y_α), (y'_α) dans la base $\{f_\alpha\}$. Par définition de la matrice associée à une forme bilinéaire,

$$b(x, x') = \sum_{ij} A_{ij} x_i x'_j = \sum_{\alpha\beta} B_{\alpha\beta} y_\alpha y'_\beta. \tag{9.275}$$

En remplaçant les x_i et x'_i par leurs valeurs en fonction de y_α et y'_β données par la proposition 4.109, nous trouvons

$$b(x, x') = \sum_{ij\alpha\beta} A_{ij} Q_{i\alpha} y_\alpha Q_{j\beta} y'_\beta \tag{9.276a}$$

$$= \sum_{\alpha\beta} (Q^t A Q)_{\alpha\beta} y_\alpha y'_\beta \tag{9.276b}$$

où Q^t désigne la transposée de la matrice Q : $Q^t_{ij} = Q_{ji}$. Vu que les nombres y_α et y'_β sont arbitraires nous déduisons[46] que $B = Q^t A Q$. □

Remarque 9.134.
Notons que cette « loi de transformation » n'est pas la même que celle pour une application linéaire[47]. Ici nous avons Q^t alors que pour les applications linéaires nous avions Q^{-1}.

Pour cette raison, tant que nous travaillons avec des bases orthonormées, c'est-à-dire tant que Q est orthogonale[48], nous pouvons confondre une application linéaire avec une application bilinéaire en passant par la matrice. Mais cette identification n'est pas du tout canonique : elle repose sur le fait que les bases soient orthonormées.

Il en découle que la réduction des endomorphismes et la réduction des formes bilinéaires ne sont pas tout à fait les mêmes théories. Par exemple la pseudo-diagonalisation simultanée (corollaire 11.36) est un résultat de réduction de forme bilinéaire et non d'endomorphismes.

44. Définition 9.118
45. Définition 9.270.
46. Lemme 4.64.
47. Proposition 4.112.
48. Définition 4.94.

9.8.4 Orthogonalité

Proposition 9.135 ([235]).
Soient un espace vectoriel $(E, \mathbb{K})$ et une forme quadratique[49] *q. Une base de E est q-orthogonale*[50] *si et seulement si la matrice de q dans cette base est diagonale.*

Démonstration. La matrice de q est donnée par $Q_{ij} = b(e_i, e_j)$. Donc oui, cette matrice est diagonale si et seulement si les e_i sont orthogonaux. □

Proposition 9.136.
Soit une forme quadratique q. Si une base (e_i) de E est q-orthogonale, alors $\mathcal{B} = \{e_i$ tel que $q(e_i) = 0\}$ est une base de $\ker(q)$.

Démonstration. Nous considérons un vecteur de base e_j, et nous montrons que $q(e_j) = 0$ si et seulement si $e_j \in \ker(q)$. Nous savons par la proposition 9.135 que la matrice de q dans la base (e_i) est diagonale et que les éléments diagonaux sont les $q(e_i)$. Soit $K = \{i$ tel que $q_(e_i) = 0\}$.

(i) $\underline{\mathrm{Span}\{e_i\}_{i\in K} \subset \ker(q)}$ Si $x = \sum_{i\in K} x_i e_i$, alors

$$q(x) = b(x, x) = \sum_{i,j\in K} |x_i|^2 b(e_i, e_j) = \sum_{i,j\in K} |x_i|^2 \delta_{ij} q(e_i) = 0 \tag{9.277}$$

parce que $q(e_i) = 0$ dès que $i \in K$.

(ii) $\underline{\ker(q) \subset \mathrm{Span}\{e_i\}_{i\in K}}$ Soit $x \in \ker(q)$ et écrivons-le sous la forme $x = \sum_{i=1}^n x_i e_i$. Nous avons

$$0 = q(x) = \sum_i |x_i|^2 q(e_i). \tag{9.278}$$

Mais $|x_i|^2 \geqslant 0$ et $q(e_i) \geqslant 0$, donc si $q(e_i) \neq 0$, alors $x_i = 0$. Donc les seules composantes non nulles de x sont celles sur lesquelles q s'annule. En d'autres termes $x = \sum_i x_i e_i \in \mathrm{Span}\{e_i\}_{i\in K}$.

□

9.8.5 Isométrie, forme quadratique et bilinéaire

Exemple 9.137.
La forme quadratique $q(x) = x_1^2 + x_2^2$ donne la norme euclidienne. La forme bilinéaire associée est $b(x, y) = x_1 y_1 + x_2 y_2$, qui est le produit scalaire usuel. △

Il ne faudrait pas déduire trop vite que la formule $\|x\|^2 = q(x)$ donne une norme dès que q est non dégénérée. En effet q peut ne pas être définie positive. La forme $q(x) = x_1^2 - x_2^2$ prend des valeurs positives et négatives. A fortiori $d(x, y) = q(x - y)$ ne donne pas toujours une distance.

Définition 9.138.
*Une **isométrie** pour la forme quadratique q est une application bijective $f\colon V \to V$ telle que*

$$q(x - y) = q\big(f(x) - f(y)\big). \tag{9.279}$$

Dans les cas où q donne une distance, alors c'est une isométrie au sens usuel.

Définition 9.139 (Thème 75).
*Soit un espace vectoriel E muni d'une forme bilinéaire b. Une **isométrie** pour b est une bijection $f\colon E \to E$ telle que*

$$b\big(f(x), f(y)\big) = b(x, y) \tag{9.280}$$

pour tout $x, y \in E$.

49. Définition 9.126.
50. Définition 9.130.

Lemme 9.140.
Soient q une forme quadratique et b la forme bilinéaire associée par le lemme 9.129. Une application $f\colon E \to E$ telle que $f(0) = 0$ est une isométrie pour b si et seulement si elle est une isométrie pour q.

Démonstration. Pour une application bijective $f\colon E \to E$ telle que $f(0) = 0$, nous devons prouver l'équivalence des propriétés suivantes :

(1) $b\big(f(x), f(y)\big) = b(x, y)$ pour tout $x, y \in E$;

(2) $q\big(f(x) - f(y)\big) = q(x - y)$ pour tout $x, y \in E$.

Dans le sens direct, en posant $x = y$ nous trouvons tout de suite $q(f(x)) = q(x)$; ensuite en utilisant la distributivité de b,

$$\begin{aligned}
q\big(f(x) - f(y)\big) &= b\big(f(x) - f(y), f(x) - f(y)\big) && (9.281a)\\
&= q\big(f(x)\big) - 2b\big(f(x), f(y)\big) + q\big(f(y)\big) && (9.281b)\\
&= q(x) + q(y) - 2b(x, y) && (9.281c)\\
&= q(x - y). && (9.281d)
\end{aligned}$$

Dans l'autre sens, nous commençons par remarquer que l'hypothèse $f(0) = 0$ donne $q(x) = q\big(f(x)\big)$. Ensuite nous utilisons l'identité de polarisation (9.268) :

$$\begin{aligned}
b\big(f(x), f(y)\big) &= \frac{1}{2}\big[q\big(f(x)\big) + q\big(f(y)\big) - q\big(f(x - y)\big)\big] && (9.282a)\\
&= \frac{1}{2}\big[q(x) + q(y) - q(x - y)\big] && (9.282b)\\
&= b(x, y). && (9.282c)
\end{aligned}$$

□

9.8.6 Isométries

Voici un théorème pas toujours bien énoncé dans les cours de physique qui font de la relativité. Au moment de « prouver » les transformations de Lorentz [51], beaucoup oublient de justifier pourquoi elles devraient être linéaires.

Théorème 9.141 ([236]).
Une isométrie [52] d'une forme bilinéaire non dégénérée est linéaire.

Démonstration. Soient une forme bilinéaire non-dégénérée b sur l'espace vectoriel E ainsi qu'une isométrie f pour icelle. Soit $z \in E$; étant donné que f est bijective nous pouvons considérer l'élément $f^{-1}(z) \in E$ et calculer

$$\begin{aligned}
b\big(f(x + y), z\big) &= b\big(f(x + y), f(f^{-1}(z))\big) && (9.283a)\\
&= b(x + y, f^{-1}(z)) && (9.283b)\\
&= b(x, f^{-1}(z)) + b(y, f^{-1}(z)) && (9.283c)\\
&= b(f(x), z) + b(f(y), z) && (9.283d)\\
&= b\big(f(x) + f(y), z\big), && (9.283e)
\end{aligned}$$

donc $f(x + y) = f(x) + f(y)$ par le lemme 9.122.

De la même façon on trouve $b\big(f(\lambda x), z\big) = b\big(\lambda f(x), z\big)$ qui prouve que $f(\lambda x) = \lambda f(x)$ et donc que f est linéaire. □

51. Théorème 18.192.
52. Définition 9.139.

Exemple 9.142.
Une isométrie peut ne pas être linéaire quand la forme bilinéaire est dégénérée. Par exemple pour la forme bilinéaire sur $\mathbb{R}^2$ donnée par

$$b\big((a,b),(x,y)\big) = ax, \tag{9.284}$$

nous pouvons faire

$$f(x,y) = \begin{pmatrix} x \\ \lambda(x,y) \end{pmatrix} \tag{9.285}$$

où λ est n'importe quoi. △

Définition 9.143.
Soient deux espaces vectoriels E et V. Une application $f\colon E \to V$ est ***affine*** *si il existe une application linéaire $u\colon E \to V$ et un élément $\alpha \in V$ tel que*

$$f(x) = u(x) + \alpha \tag{9.286}$$

pour tout $x \in E$.

Théorème 9.144.
Soit un espace vectoriel E muni d'une forme quadratique q. Soit une isométrie $f\colon E \to E$ pour q. Alors

(1) si $f(0) = 0$, alors f est linéaire ;

(2) si $f(0) \neq 0$ alors f est affine[53].

Démonstration. Nous considérons la forme bilinéaire associée b. Si $f(0) = 0$, nous savons par le lemme 9.140 que $b\big(f(x), f(y)\big) = b(x,y)$. La proposition 9.141 nous dit alors que f est linéaire.

Si $f(0) \neq 0$, alors nous posons $g(x) = f(x) - f(0)$ qui vérifie $g(0) = 0$ et

$$q\big(g(x) - g(y)\big) = q\big(f(x) - f(0) - f(y) + f(0)\big) = q(x-y). \tag{9.287}$$

Nous pouvons donc appliquer le premier point à g, déduire que g est linéaire et donc que f est affine. C'est la caractérisation du lemme 8.59 des fonctions affines. □

Nous pouvons maintenant particulariser tout cela au cas de $\mathbb{R}^n$ muni du produit scalaire usuel et de la norme associée pour voir quel résultat nous avons à peine prouvé.

Lemme 9.145 ([237])**.**
Une isométrie d'un espace vectoriel normé de dimension finie est bijective.

Démonstration. Si $f\colon E \to E$ est une isométrie, elle est linéaire par le théorème 9.141. Elle vérifie également $\|f(x)\| = \|x\|$, et donc $f(x) = 0$ si et seulement si $x = 0$, c'est-à-dire que f est injective. Elle est alors bijective par le corolaire 4.46 du théorème du rang. □

Nous notons ici $T(n)$ le groupe des translations sur $\mathbb{R}^n$. Un élément de $T(n)$ est une translation τ_v donnée par un vecteur v et agissant sur $\mathbb{R}^n$ par

$$\begin{aligned} \tau_v\colon \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto x + v. \end{aligned} \tag{9.288}$$

Ce groupe est isomorphe au groupe abélien $(\mathbb{R}^n, +)$, et nous allons souvent identifier τ_v à v.

Vous savez par culture générale que les isométries de $\mathbb{R}^n$ pour le produit scalaire usuel sont les matrices orthogonales. En voici une petite généralisation (pensez à $\eta = \mathbb{1}$ dans le cas du produit scalaire usuel).

53. Définition 9.143.

Proposition 9.146.
Soit une forme bilinéaire b sur $\mathbb{R}^n$ de matrice symétrique η. Si A est la matrice d'une application linéaire $\mathbb{R}^n \to \mathbb{R}^n$ telle que

$$b(Ax, Ay) = b(x, y) \tag{9.289}$$

pour tout $x, y \in \mathbb{R}^n$, alors

$$A^t \eta A = \eta. \tag{9.290}$$

Démonstration. En suivant la formule générale (9.271),

$$b(Ax, Ay) = \sum_{ij} \eta_{ij} (Ax)_i (Ay)_j = \sum_{ijkl} \eta_{ij} A_{ik} A_{jl} x_k y_l. \tag{9.291}$$

En imposant que ce soit égal à $\sum_{kl} \eta_{kl} \eta_{kl} x_k y_l$ pour tout x, y nous avons la contrainte

$$\sum_{ij} \eta_{ij} A_{ik} A_{jl} = \eta_{kl} \tag{9.292}$$

qui signifie exactement $A^t \eta A = \eta$. □

9.9 Signature, théorème de Sylvester

Définition 9.147 (Signature[238])**.**
Soit une forme quadratique[54] *Q sur un espace vectoriel E de dimension finie n. L'**indice d'inertie** de q est le nombre*

$$q = \max\{\dim(F) \text{ tel que } Q(v) < 0 \, \forall v \in F \backslash \{0\}\}. \tag{9.293}$$

Nous définissons aussi

$$p = \max\{\dim(G) \text{ tel que } Q(v) > 0 \, \forall v \in G \backslash \{0\}\}. \tag{9.294}$$

*Le couple (p, q) est la **signature** de Q.*

Définition 9.148 (Rang d'une forme quadratique)**.**
Si $Q \colon E \times E \to \mathbb{K}$ est une forme quadratique, nous considérons l'application

$$\begin{aligned} f_Q \colon E &\to E^* \\ x &\mapsto \big[y \mapsto B(x, y) \big]. \end{aligned} \tag{9.295}$$

*Le **rang** de Q est le rang de l'application linéaire f_Q.*

Proposition 9.149.
Le rang d'une forme quadratique est le rang de sa matrice dans n'importe quelle base.

Démonstration. Nous considérons une forme quadratique Q sur l'espace vectoriel E. Sa trace est, par définition, la trace de l'application linéaire f_Q de la définition 9.148. Or cette dernière trace ne dépend pas des bases choisies sur E et E^*. Nous la calculons donc maintenant.

Soit une base $\{e_i\}$ de E ainsi que sa base duale $\{e_i^*\}$ de E^*. Si $v = \sum_k v_k e_k \in E$, alors

$$f_Q(e_i) v = \sum_k v_k B(e_i, e_k) = \sum_k Q_{ik} v_k \tag{9.296}$$

où nous avons noté B la forme bilinéaire associée à Q et utilisé la définition 9.131 de la matrice associée à la forme quadratique Q. Nous avons donc $f_Q(ei) = \sum_k Q_{ik} e_k^*$ ou encore

$$f_Q(e_i)_k = Q_{ik}, \tag{9.297}$$

ce qui signifie, par (1) que la matrice associée à f_Q est la matrice Q^t.

Le rang de f_Q est donc celui de Q^t, qui est le même que celui de la matrice Q (ici, nous avons noté Q la matrice de la forme quadratique Q). Le rang de f_Q est celui de sa matrice par la proposition 4.103. □

54. Définition 9.126.

Lemme 9.150 ([238]).
Soient une forme quadratique Q ainsi que deux bases Q-orthogonales $\{e_1, \dots, e_n\}$ et $\{e'_1, \dots, e'_n\}$. Nous posons

$$r = \mathrm{Card}\{e_i \text{ tel que } q(e_i) > 0\} \tag{9.298a}$$
$$r' = \mathrm{Card}\{e_i \text{ tel que } q(e'_i) > 0\} \tag{9.298b}$$
$$s = \mathrm{Card}\{e_i \text{ tel que } q(e_i) < 0\} \tag{9.298c}$$
$$s' = \mathrm{Card}\{e_i \text{ tel que } q(e'_i) < 0\} \tag{9.298d}$$

Alors $r = r'$ et $s = s'$.

Démonstration. Nous posons

$$I = \{i \text{ tel que } Q(e_i) > 0\} \tag{9.299a}$$
$$J = \{j \text{ tel que } Q(e_j) < 0\} \tag{9.299b}$$

Nous commençons par prouver que $\{e_i\}_{i\in I} \cup \{e'_j\}_{j\in J}$ est libre. Supposons pour cela que

$$\sum_{i\in I} x_i e_i + \sum_{j\in J} y_j e'_j = 0, \tag{9.300}$$

et posons $z = \sum_{i\in I} x_i e_i$. Nous avons

$$Q(z) = \sum_{i\in I} x_i^2 Q(e_i) \geqslant 0. \tag{9.301}$$

Mais nous avons aussi $z = -\sum_{j\in J} y_j e'_j$, donc

$$Q(z) = \sum_{j\in J} y_j^2 Q(e'_j) \leqslant 0. \tag{9.302}$$

Donc $Q(z) = 0$. Vu (9.301), et le fait que $Q(e_i) > 0$, avoir $Q(z) = 0$ impose $x_i = 0$ pour tout i. La relation (9.302) nous donne aussi immédiatement que les y_j sont nuls. Donc la partie $\{e_i\}_{i\in I} \cup \{e'_j\}_{j\in J}$ est libre.

Le lemme 4.10 nous indique qu'une partie libre est toujours de cardinal plus petit ou égal à la dimension de l'espace [55]. Tout ça pour dire que

$$\underbrace{\mathrm{Card}(I)}_{=r} + \underbrace{\mathrm{Card}(J)}_{=n-r'} \leqslant n, \tag{9.303}$$

et donc $r \leqslant r'$.

Le même raisonnement, en partant de $I = \{i \text{ tel que } Q(e_i) \leqslant 0\}$ et de $J = \{j \text{ tel que } Q(e'_j) > 0\}$, prouve que $r' \leqslant r$.

La preuve de $s = s'$ est du même tonneau. □

Lemme 9.151 ([239]).
Soit une forme quadratique Q sur E. Si F est un sous-espace de E, alors

$$\dim(F) + \dim(F^\perp) \geqslant n \tag{9.304}$$

où $F^\perp$ est l'orthogonal par rapport à Q.

Démonstration. Nous posons $p = \dim(F)$. Nous considérons une base $\{f_i\}_{i=1,\dots,n}$ de E telle que $\{f_i\}_{i=1,\dots,p}$ est une base de F [56]. Nous posons

$$\begin{aligned} \phi\colon E &\to F \\ x &\mapsto \sum_{i=1}^{p} B(x, f_i) f_i \end{aligned} \tag{9.305}$$

55. Ici nous utilisons l'hypothèse que V est de dimension finie.
56. Théorème de la base incomplète, 4.23.

où B est la forme bilinéaire associée à Q. Ce ϕ est une application linéaire à qui nous appliquons le théorème du rang (4.46) :

$$\dim(E) = \mathrm{rk}(\phi) + \dim\big(\ker(\phi)\big). \tag{9.306}$$

Mais vu que l'image de ϕ est dans F, nous avons $\mathrm{rk}(\phi) \leqslant \dim(F)$. De plus, $\ker(\phi) = F^\perp$. Donc (9.306) devient

$$\dim(E) \leqslant \dim(F) + \dim(F^\perp). \tag{9.307}$$

□

Lemme 9.152 ([239]).
Soit un espace vectoriel E de dimension finie et un sous-espace F sur lequel la forme quadratique Q est strictement définie positive ou négative. Alors

$$E = F \oplus F^\perp. \tag{9.308}$$

Démonstration. D'abord nous montrons que $F \cap F^\perp = \{0\}$. Si $v \neq 0$ est dans F, alors $Q(v) > 0$, et donc v n'est pas dans $F^\perp$. Donc $F \cap F^\perp \subset \{0\}$. L'inclusion inverse est immédiate.

Nous avons vu dans le lemme 9.151 que

$$\dim(E) \leqslant \dim(F) + \dim(F^\perp). \tag{9.309}$$

Vu que F et $F^\perp$ n'ont pas d'intersection autre que $\{0\}$, nous avons

$$\dim(E) \geqslant \dim(F \oplus F^\perp) = \dim(F) + \dim(F^\perp) \geqslant \dim(E). \tag{9.310}$$

Toutes ces inégalités sont donc des égalités et $\dim(E) = \dim(F) + \dim(F^\perp)$. □

9.9.1 Équivalence de formes quadratiques

Définition 9.153 (Équivalence de forme quadratique[234]).
*Deux formes quadratiques Q et Q' sont **équivalentes** si il existe une application linéaire inversible ϕ telle que $Q' = Q \circ \phi$.*

Proposition 9.154 ([238]).
Deux formes quadratiques sont équivalentes[57] *si et seulement si elles ont même signature.*

9.10 Produit scalaire, produit hermitien

Définition 9.155.
*Un **produit scalaire** sur un espace vectoriel réel est une forme bilinéaire*[58] *symétrique strictement définie positive*[59].

La définition suivante est utile pour ceux qui veulent faire de la relativité[60].

Définition 9.156.
*Un **produit pseudo-scalaire** sur un espace vectoriel réel est une forme bilinéaire et symétrique.*

Définition 9.157.
*Nous dirons que deux vecteurs sont **orthogonaux** lorsque leur produit scalaire*[61] *est nul. Nous écrivons que $u \perp v$ lorsque $\langle u, v\rangle = 0$.*

*Si $\{e_i\}_{i=1,\dots,n}$ est une base de E, nous disons qu'elle est **orthonormée** si*

$$\langle e_i, e_j\rangle = \delta_{ij}. \tag{9.311}$$

57. Définition 9.153.
58. Définition 9.118.
59. Définition 9.127.
60. Voir le théorème 18.192 qui établit les transformations de Lorentz.
61. Définition 9.155.

Lemme 9.158.
Un produit scalaire est toujours non dégénéré[62].

Vu que nous allons voir un pâté d'espaces avec des produits scalaires, nous leur donnons un nom.

Définition 9.159.
Un espace vectoriel ***euclidien*** *est un espace vectoriel de dimension finie muni d'un produit scalaire (définition 9.155).*

Avouez que c'est drôle qu'un espace vectoriel est euclidien lorsqu'il possède une *multiplication* alors qu'un anneau est euclidien lorsqu'il possède une *division* (voir la définition 3.105). C'est pas très profond, mais si ça peut vous servir de moyen mnémotechnique...

Proposition-Définition 9.160 (Produit scalaire dans $\mathbb{R}^n$, thème 24)**.**
Si $x, y \in \mathbb{R}^n$, nous définissons

$$x \cdot y = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \tag{9.312}$$

C'est un produit scalaire[63] *qui vérifie*

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* $x = 0$.

Ce produit scalaire est le ***produit scalaire*** *qui sera toujours considéré. C'est de lui qui découle toujours la norme, et la topologie de $\mathbb{R}^n$. Il sera aussi souvent noté $\langle x, y \rangle$.*

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

$$\langle e_i, e_j \rangle = \sum_{k=1}^{m} \delta_{ik} \delta_{jk}. \tag{9.313}$$

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 :

$$\langle e_i, e_j \rangle = \delta_{ii} \delta_{ji} = \delta_{ji}. \tag{9.314}$$

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 9.161.
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.

$$v \cdot e_j = \sum_{i=1}^{m} \langle v_i e_i, e_j \rangle = \sum_{i=1}^{m} v_i \langle e_i, e_j \rangle = \sum_{i=1}^{m} v_i \delta_{ij} \tag{9.315}$$

En effectuant la somme sur i dans le membre de droite de l'équation (9.315), tous les termes sont nuls sauf celui où $i = j$; il reste donc

$$v \cdot e_j = v_j. \tag{9.316}$$

□

62. Définition 9.121.
63. Définition 9.155.

Le produit scalaire ne dépend en réalité pas de la base orthogonale choisie.

Lemme 9.162.
Si $\{e_i\}$ est la base canonique, et si $\{f_i\}$ est une autre base orthonormale[64]*, alors si u et v sont deux vecteurs de $\mathbb{R}^m$, nous avons*

$$\sum_i u_i v_j = \sum_i u'_i v'_j \tag{9.317}$$

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

$$\begin{aligned}
\sum_i u'_i v'_j &= \sum_i \left(\sum_j A_{ij} u_j\right)\left(\sum_k A_{ik} v_k\right) \\
&= \sum_{ijk} A_{ij} A_{ik} u_j v_k \\
&= \sum_{jk} \underbrace{\sum_i (A^t)_{ji} A_{ik}}_{=\delta_{jk}} u_j v_k \\
&= \sum_{jk} \delta_{jk} u_j v_k \\
&= \sum_k u_j v_k.
\end{aligned} \tag{9.318}$$

□

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 exprimons les vecteurs.

9.10.1 Hermitien, unitaire, etc.

Définition 9.163 ([240])**.**
*Soit E est un espace vectoriel sur $\mathbb{C}$. Une application $\langle .,. \rangle \colon E \times E \to \mathbb{C}$ est **sesquilinéaire à droite** si pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{C}$,*

(1) $\langle \lambda x, y \rangle = \lambda \langle x, y, \rangle = \langle x, \bar{\lambda} y \rangle$,

(2) $\langle x + y, z \rangle = \langle x, y \rangle + \langle y, z \rangle$,

(3) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

*Cette forme est **hermitienne** si de plus*

$$\langle x, y \rangle = \overline{\langle y, x \rangle}. \tag{9.319}$$

*Un **produit hermitien** est une forme hermitienne strictement définie positive, c'est-à-dire telle que $\langle x, x \rangle \geqslant 0$ pour tout $x \in E$ et $\langle x, x \rangle = 0$ si et seulement si $x = 0$.*

Proposition-Définition 9.164.
Soit un espace hermitien $(E, \langle .,. \rangle)$. Soit un opérateur linéaire $A \in \mathrm{End}(E)$. Il existe un unique opérateur $B \in \mathrm{End}(E)$ tel que

$$\langle Ax, y \rangle = \langle x, By \rangle \tag{9.320}$$

pour tout $x, y \in E$.

*L'opérateur B ainsi défini est noté $A^\dagger$ et est nommé **hermitien conjugué** de A.*

64. Définition 9.157.

Lemme 9.165.
Au niveau des matrices, nous avons

$$A^{\dagger}_{ij} = A^{*}_{ji} \tag{9.321}$$

où z^* est le conjugué complexe.

Définition 9.166.
Quelque définitions.

(1) *Un opérateur A est **hermitien** si $A^{\dagger} = A$.*
(2) *Un opérateur A est **unitaire** si $A^{\dagger} = A^{-1}$.*

Note : le conjugué hermitien est parfois noté A^* au lieu de $A^{\dagger}$.

9.167.
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. 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 $\mathbb{R}$.

L'ensemble des opérateurs autoadjoints de E est noté $\mathrm{S}(E)$. Cette notation provient du fait que dans $\mathbb{R}^n$ muni du produit scalaire usuel, les opérateurs autoadjoints sont les matrices symétriques.

Remarque 9.168.
Le fait d'être hermitien n'implique en rien le fait d'être inversible.

9.169.
Les normes associées aux produits scalaires font intervenir une racine carré, et donc devront être données plus tard. Voir le thème 24.

Proposition 9.170 ([1])**.**
Nous considérons $\mathbb{C}^n$ vu comme espace vectoriel de dimension n sur $\mathbb{C}$.

(1) *La formule, pour $x, y \in \mathbb{C}^n$,*

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \bar{y}_k \tag{9.322}$$

définit une une forme sesquilinéaire sur $\mathbb{C}^n$.

(2) *L'ensemble $\mathbb{C}^n$ devient un espace vectoriel hermitien.*

9.10.2 Éléments de matrice

Proposition 9.171.
Soit une application linéaire $A\colon \mathbb{R}^n \to \mathbb{R}^n$. Nous considérons le produit scalaire usuel[65] *sur $\mathbb{R}^n$. Alors :*

(1) *Les éléments de matrice de A sont donnés par $A_{ij} = e_i \cdot Ae_j$.*
(2) *Nous avons la formule $x \cdot Ay = \sum_{kl} A_{kl} x_k y_l$.*

Démonstration. Pour Ae_j nous utilisons la formule 4.82 avec des notations plus décontractées : $Ae_j = \sum_k A_{kj} e_k$. Ensuite nous faisons un calcul avec la formule (9.312) :

$$e_i \cdot Ae_j = e_i \cdot \sum_k A_{kj} e_k = \sum_k A_{kj} \delta_{i,k} = A_{ij}. \tag{9.323}$$

La seconde formule à prouver est du même tonneau, en utilisant cette fois la formule (4.83) :

$$x \cdot Ay = \sum_k x_k (Ay)_k = \sum_{kl} x_k A_{kl} y_l = \sum_{kl} A_{kl} x_k y_l. \tag{9.324}$$

□

65. Définition 9.160.

La proposition suivante est une version plus « pragmatique » de la proposition 4.125.

Proposition 9.172 ([121]).
Soient un espace euclidien[66] *de dimension finie V ainsi qu'un sous-espace M. Nous posons*

$$M^\perp = \{x \in V \text{ tel que } x \cdot y = 0 \forall y \in M\}. \tag{9.325}$$

Alors $M \oplus M^\perp = V$.

Démonstration. D'abord si $x \in M \cap M^\perp$, alors $x \cdot x = 0$ et donc $x = 0$. Donc nous avons déjà $M \cap M^\perp = \{0\}$. Nous considérons une base $\{b_1, \dots, b_k\}$ de M, et nous définissons l'application linéaire

$$\begin{aligned} f\colon V &\to \mathbb{R}^k \\ x &\mapsto (x \cdot b_1, \dots, x \cdot b_k). \end{aligned} \tag{9.326}$$

Nous avons que $M^\perp = \ker(f)$. Le théorème du rang 4.44 nous indique que

$$\dim(V) = \dim\big(\ker(f)\big) + \dim\big(\text{Image}(f)\big) \leqslant \dim(M^\perp) + k = \dim(M^\perp) + \dim(M). \tag{9.327}$$

Une justification : vu que f prend ses valeurs dans $\mathbb{R}^k$, la dimension de son image est majorée par k.

Nous en déduisons que

$$\dim(M) + \dim(M^\perp) \geqslant \dim(V), \tag{9.328}$$

et la proposition 4.133 nous permet de conclure que $M \oplus M^\perp = V$. □

9.10.3 Transposée : pas d'approche naïve

Il est légitime, si $t\colon E \to E$ est une application linéaire, de dire que sa transposée soit l'application linéaire $t^t\colon E \to E$ dont la matrice est la matrice transposée de celle de t. Lorsque nous travaillons sur $\mathbb{R}^n$ muni de la base canonique, cela ne pose pas de problème et nous pouvons écrire des égalités du type $\langle x, Ay\rangle = \langle A^t x, y\rangle$.

Proposition 9.173 (Matrice transposée et produit scalaire).
Soit une matrice réelle A. En utilisant l'application linéaire associée[67] *$f_A\colon \mathbb{R}^n \to \mathbb{R}^n$, nous avons*

$$x \cdot f_A(y) = f_{A^t}(x) \cdot y. \tag{9.329}$$

Cette formule est souvent écrite $x \cdot Ay = A^t x \cdot y$ ou $\langle x, Ay\rangle = \langle A^t x, y\rangle$.

Démonstration. Il s'agit d'un calcul utilisant la formule (4.77) et le produit scalaire (9.312) :

$$x \cdot f_A(y) = \sum_i x_i \big(f_A(y)\big)_i = \sum_i x_i \sum_j A_{ij} y_j = \sum_{ij} A^t_{ji} x_i y_j = \sum_j f_{A^t}(x)_j y_j = f_{A^t}(x) \cdot y. \tag{9.330}$$

□

Hélas nous allons voir que cette façon de définir une transposée est mauvaise.

Soit une application linéaire $t\colon E \to E$ de matrice A dans la base $\{e_i\}_{i=1,\dots,n}$ et de matrice B dans la base $\{f_\alpha\}_{\alpha=1,\dots,n}$. Nous notons Q la matrice de passage d'une base à l'autre :

$$e_i = \sum_\alpha Q^{-1}_{\alpha i} f_\alpha. \tag{9.331}$$

Nous nommons t_1 l'application linéaire associée à A^t dans la base $\{e_i\}$ et t_2 l'application linéaire associée à la matrice B^t dans la base $\{f_\alpha\}$. Définir la transposée d'une application linéaire comme étant l'application linéaire associée à la transposée de sa matrice ne sera une bonne définition que si $t_1 = t_2$.

66. Qui possède un produit scalaire, définition 9.159.
67. Définition 4.65. Ici nous considérons la base canonique sur $\mathbb{R}^n$

La première chose facile à voir est

$$t_1(e_i)_j = \sum_k (A^t)_{jk}(e_i)_k = A^t_{ji} = A_{ij}. \tag{9.332}$$

Pour calculer $t_2(e_i)_j$, c'est un peu plus laborieux :

$$t_2(e_i) = \sum_\alpha Q^{-1}_{\alpha i} t_2(f_\alpha) = \sum_{\beta\gamma\alpha} Q^{-1}_{\alpha i} B^t_{\gamma\beta} \underbrace{(t_\alpha)_\beta}_{\delta_{\alpha,\beta}} f_\gamma = \sum_{\beta\gamma} Q^{-1}_{\beta i} B^t_{\gamma\beta} f_\gamma \tag{9.333a}$$

$$= (B^t Q^{-1})_{\gamma i} Q_{j\gamma} e_j \tag{9.333b}$$

$$= \sum_j (QB^t Q^{-1})_{ji} e_j. \tag{9.333c}$$

Donc $t_2(e_i)_j = (QB^tQ^{-1})_{ji}$. En tenant compte du fait que $B = Q^{-1}AQ$ nous avons

$$t_2(e_i)_j = (QQ^tA^t(Q^{-1})^tQ^{-1})_{ji}. \tag{9.334}$$

Ceci est égal à l'expression (9.332) lorsque $Q^t = Q^{-1}$. Nous voyons que confondre transposée d'une application linéaire avec transposée de la matrice associée n'est valable que si nous sommes certain de ne considérer que des changements de base par des matrices orthogonales.

C'est la situation typique dans laquelle nous nous trouvons lorsque nous considérons des applications linéaires sur $\mathbb{R}^n$ muni de la base canonique, et que nous n'avons aucune intention de changer de base, et encore moins de chercher une base non orthonormale. Cette situation est clairement la situation la plus courante.

Exemple 9.174 ([78]).
Soit la base canonique $\{e_1, e_2\}$ de $\mathbb{R}^2$. Nous considérons l'application linéaire $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$t(e_1) = e_1 \tag{9.335a}$$

$$t(e_2) = 0. \tag{9.335b}$$

La matrice de t dans cette base est

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{9.336}$$

Elle est symétrique : elle vérifie $A^t = A$. Si nous comptions sur la transposée de matrice pour définir la transposée de t, nous aurions $t^t = t$.

Soit maintenant la base $f_1 = e_1$, $f_2 = e_1 + e_2$. Nous avons $t(f_1) = f_1$ et

$$t(f_2) = t(e_1) + t(e_2) = e_1 = f_1. \tag{9.337}$$

Donc la matrice de t dans cette base est

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}. \tag{9.338}$$

Et là, nous avons $B^t \neq B$. Donc en comptant sur cette base pour définir la transposée de t nous aurions $t^t \neq t$. △

9.175.
Autrement dit, la façon « usuelle » de voir la transposée d'une application linéaire, ne fonctionne dans les livres pour enfants uniquement parce qu'on y considère toujours $\mathbb{R}^n$ muni de la base canonique ou de bases orthonormées.

Notons que nous avons tout de même les notions d'opérateur adjoint et autoadjoint pour parler d'application orthogonale sans passer par la transposée, voir 9.35.

9.10.4 Transposée : la bonne approche

Définition 9.176.
Si $f\colon E \to F$ est une application linéaire entre deux espaces vectoriels, la ***transposée*** *est l'application $f^t\colon F^* \to E^*$ donnée par*

$$f^t(\omega)(x) = \omega\big(f(x)\big). \tag{9.339}$$

pour tout $\omega \in F^*$ et $x \in E$.

Lemme 9.177.
Soit E muni de la base $\{e_i\}$ et F muni de la base $\{g_i\}$ et une application $f\colon E \to F$. 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^*.

Autrement dit, en utilisant l'application $\psi\colon \mathbb{M} \to \mathcal{L}(F^*, E^*)$ de la proposition 4.68,

$$\psi(A^t) = f^t. \tag{9.340}$$

Démonstration. Attention aux indices, ça va chauffer[68].

Nous allons montrer que $f^t = \psi(A^t)$ sur la base $\{g_i^*\}$, et pour cela nous appliquons $f^t(g_i^*)$ à $x \in E$:

$$\begin{aligned}
f^t(g_i^*)x &= g_i^*\big(f(x)\big) && \text{Définition 9.176} && (9.341a)\\
&= g_i^*\big(\sum_k x_k f(e_k)\big) && && (9.341b)\\
&= g_i^*\big(\sum_{kl} x_k A_{lk} g_l\big) && \text{eq. (4.82)} && (9.341c)\\
&= \sum_{kl} x_k A_{lk} \underbrace{g_i^*(g_l)}_{=\delta_{i,l}} && && (9.341d)\\
&= \sum_k x_k A_{ik} && && (9.341e)\\
&= \sum_k (A^t)_{ki} x_k && && (9.341f)\\
&= \sum_k (A^t)_{ki} e_k^*(x). && && (9.341g)
\end{aligned}$$

Voila. Donc nous avons

$$f^t(g_i^*) = \sum_k (A^t)_{ki} e_k^* = \psi(A^t) g_i^*. \tag{9.342}$$

□

9.178.
Intuitivement, les rangs de f et de f^t sont égaux parce que le rang est donné par la plus grande matrice carrée de déterminant non nul.

Nous donnons maintenant une vraie preuve de ce résultat.

Lemme 9.179 ([241])**.**
Si $f\colon E \to F$ est une application linéaire, alors

$$\mathrm{rk}(f) = \mathrm{rk}(f^t). \tag{9.343}$$

Démonstration. Soient $n = \dim(E)$ et $r = \dim(F)$. Nous posons $\dim\ker(f) = p$ et donc $\mathrm{rk}(f) = n - p$. 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

$$g_i = f(e_{p+i}) \tag{9.344}$$

68. Et merci à Alain Vigne pour m'avoir fait remarquer qu'il fallait mettre de l'ordre dans les indices.

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

$$\sum_{k=1}^{n-p} a_k f(e_{p+k}) = 0, \tag{9.345}$$

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 pouvons les compléter en une base de F :

$$\{\underbrace{g_1, \dots, g_{n-p}}_{\text{images}}, \underbrace{g_{n-p+1}, \dots, g_r}_{\text{complétion}}\}. \tag{9.346}$$

Nous prouvons maintenant que $\mathrm{rk}(f^t) \geqslant n-p$ en montrant que les formes $\{g_i^*\}_{i=1,\dots,n-p}$ forment une partie libre (et donc l'espace image de f^t est au moins de dimension $n-p$). Pour cela nous prouvons que $f^t(g_i^*) = e_{i+p}^*$. En effet

$$f^t(g_i^*)(e_k) = g_i^*(f(e_k)), \tag{9.347}$$

Si $k = 1, \dots, p$, alors $f(e_k) = 0$ et donc $g_i^*(f(e_k)) = 0$; si $k = p + l$ alors

$$f^t(g_i^*)(e_k) = g_i^*(f(e_{k+l})) = g_i^*(g_l) = \delta_{i,l} = \delta_{i,k-p} = \delta_{k,i+p}. \tag{9.348}$$

Donc $f^t(g_i^*) = e_{i+p}^*$. Cela prouve que les formes $f^t(g_i^*)$ sont libres et donc que

$$\mathrm{rk}(f^t) \geqslant n - p = \mathrm{rk}(f). \tag{9.349}$$

En appliquant le même raisonnement à f^t au lieu de f, nous trouvons

$$\mathrm{rk}\left((f^t)^t\right) \geqslant \mathrm{rk}(f^t) \tag{9.350}$$

et donc, sachant que $(f^t)^t = f$, nous obtenons $\mathrm{rk}(f) = \mathrm{rk}(f^t)$.

□

Proposition 9.180 ([242]).
Si f est une application linéaire entre les espaces vectoriels E et F, alors nous avons

$$\mathrm{Image}(f^t) = \ker(f)^\perp. \tag{9.351}$$

Démonstration. Soient donc l'application $f\colon E \to F$ et sa transposée $f^t\colon F^* \to E^*$. Nous commençons par prouver que $\mathrm{Image}(f^t) \subset (\ker f)^\perp$. Pour cela nous prenons $\omega \in \mathrm{Image}(f^t)$, 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

$$\begin{aligned}
\dim\big(\mathrm{Image}(f^t)\big) &= \mathrm{rk}(f^t) & & (9.352a)\\
&= \mathrm{rk}(f) & \text{lemme 9.179} & (9.352b)\\
&= \dim(E) - \dim\ker(f) & \text{théorème 4.44} & (9.352c)\\
&= \dim\big((\ker f)^\perp\big) & \text{proposition 4.125.} & (9.352d)
\end{aligned}$$

□

Lemme 9.181 ([113]).
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[69] *f^t est surjective.*

69. Définition 9.176.

Démonstration. Supposons que f soit injective. Alors par le lemme 4.53, 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 faisant jouer les identifications $(f^t)^t = f$ et $(E^*)^* = E$ nous savons qu'il existe $s\colon E^* \to F^*$ tel que $f^t \circ s = \mathrm{Id}\,|_{E^*}$. En passant à la transposée,

$$s^t \circ f = \mathrm{Id}\,|_E, \tag{9.353}$$

qui implique que f est injective. □

9.10.5 Polynômes de Lagrange

Lemme-Définition 9.182.
Soit $E = \mathcal{P}_n(\mathbb{R})$ l'ensemble des polynômes à coefficients réels de degré au plus n. Soient $n+1$ réels distincts $a_0, \dots, a_n$. Nous considérons les formes linéaires associées $f_i \in \mathcal{P}_n(\mathbb{R})^*$,

$$f_i(P) = P(a_i). \tag{9.354}$$

La partie $\{f_1, \dots, f_n\}$ est une base de $\mathcal{P}_n(\mathbb{R})^*$.

*Les **polynômes de Lagrange** aux points (a_i) sont les polynômes de la base préduale de la base $\{f_i\}$.*

Démonstration. Nous prouvons que l'orthogonal est réduit au singleton nul :

$$\mathrm{Span}\{f_0, \dots, f_n\}^\perp = \{0\}. \tag{9.355}$$

La proposition 4.125 conclura. Si $P \in \mathrm{Span}\{f_i\}^\perp$, 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. □

Proposition 9.183.
Les polynômes de Lagrange aux points $(a_i)_{i=1,\dots,n}$ sont donnés par

$$P_i = \prod_{k \neq i} \frac{X - a_k}{a_i - a_k}. \tag{9.356}$$

Démonstration. Il suffit de vérifier que $f_j(P_i) = \delta_{i,j}$. Nous avons

$$f_j(P_i) = P_i(a_j) = \prod_{k \neq i} \frac{a_j - a_k}{a_i - a_k}. \tag{9.357}$$

Si $j \neq i$ alors un des termes est nul. Si au contraire $i = j$, tous les termes valent 1. □

9.10.6 Dual de $\mathbb{M}(n, \mathbb{K})$

Proposition 9.184 ([89]).
Soit $\mathbb{K}$, un corps. Les formes linéaires sur $\mathbb{M}(n, \mathbb{K})$ sont les applications de la forme

$$\begin{aligned} f_A\colon \mathbb{M}(n, \mathbb{K}) &\to \mathbb{K} \\ M &\mapsto \mathrm{Tr}(AM). \end{aligned} \tag{9.358}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} f\colon \mathbb{M}(n, \mathbb{K}) &\to \mathbb{M}(n, \mathbb{K})^* \\ A &\mapsto f_A \end{aligned} \tag{9.359}$$

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 $(E_{ij})_{kl} = \delta_{i,k}\delta_{j,l}$:

$$0 = f_A(E_{ij}) = \sum_k (AE_{ij})_{kk} = \sum_{kl} A_{kl}(E_{ij})_{lk} = \sum_{kl} A_{kl}\delta_{i,l}\delta_{j,k} = A_{ij}. \tag{9.360}$$

Donc $A = 0$. □

Corolaire 9.185 ([89]).
Soient un corps $\mathbb{K}$ ainsi qu'une application $\phi \in \mathbb{M}(n,\mathbb{K})^*$ telle que pour tout $M, N \in \mathbb{M}(n,\mathbb{K})$ on ait

$$\phi(MN) = \phi(NM). \tag{9.361}$$

Alors il existe $\lambda \in \mathbb{K}$ tel que $\phi = \lambda \operatorname{Tr}$.

Démonstration. La proposition 9.184 nous donne une matrice $A \in \mathbb{M}(n,\mathbb{K})$ telle que $\phi = f_A$. L'hypothèse nous dit que $f_A(MN) = f_A(NM)$, c'est-à-dire

$$\operatorname{Tr}(AMN) = \operatorname{Tr}(ANM) \tag{9.362}$$

pour toutes matrices $M, N \in \mathbb{M}(n,\mathbb{K})$. L'invariance cyclique de la trace [70] appliqué au membre de droite nous donne $\operatorname{Tr}(AMN) = \operatorname{Tr}(MAN)$, ce qui signifie que

$$\operatorname{Tr}\big((AM - MA)N\big) = 0 \tag{9.363}$$

ou encore que $f_{AM-MA} = 0$, et ce, pour toute matrice M. La fonction f étant injective nous en déduisons que la matrice A doit satisfaire

$$AM = MA \tag{9.364}$$

pour tout $M \in \mathbb{M}(n,\mathbb{K})$. En particulier, en prenant pour M les fameuses matrices E_{ij} et en calculant un peu,

$$A_{li}\delta_{j,m} = \delta_{i,l}A_{jm} \tag{9.365}$$

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,

$$\phi(X) = f_{\lambda\mathbb{1}}(X) = \lambda \operatorname{Tr}(X). \tag{9.366}$$

□

Corolaire 9.186 ([89]).
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 $\mathbb{M}(n,\mathbb{K})$. Il existe une forme linéaire ϕ sur $\mathbb{M}(n,\mathbb{K})$ telle que $\mathcal{H} = \ker(\phi)$. Encore une fois la proposition 9.184 nous donne $A \in \mathbb{M}(n,\mathbb{K})$ telle que $\phi = f_A$; nous notons r le rang de A. Par le lemme 4.105 nous avons $A = PJ_rQ$ avec $P, Q \in \mathrm{GL}(n,\mathbb{K})$ et

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{9.367}$$

Pour tout $M \in \mathbb{M}(n,\mathbb{K})$ nous avons

$$\phi(M) = \operatorname{Tr}(AM) = \operatorname{Tr}(PJ_rQM) = \operatorname{Tr}(J_rQMP), \tag{9.368}$$

70. Lemme 4.62.

la dernière égalité découlant de l'invariance cyclique de la trace[71]. Ce que nous cherchons est $M \in \mathrm{GL}(n, \mathbb{K})$ telle que $\phi(M) = 0$. Nous commençons par trouver $N \in \mathrm{GL}(n, \mathbb{K})$ telle que $\mathrm{Tr}(J_r N) = 0$. Celle-là est facile : c'est

$$N = \begin{pmatrix} 0 & 1 \\ \mathbb{1}_{n-1} & 0 \end{pmatrix}. \tag{9.369}$$

Les éléments diagonaux de $J_r N$ sont tous nuls. Par conséquent en posant $M = Q^{-1} N P^{-1}$ nous avons notre matrice inversible dans le noyau de ϕ. □

9.11 Diagonalisation et trigonalisation

Ici encore $\mathbb{K}$ est un corps commutatif.

9.11.1 Matrices semblables

Proposition-Définition 9.187 (matrices semblables[1])**.**
Nous définissons, sur l'ensemble $\mathbb{M}(n, \mathbb{K})$ des matrices $n \times n$ à coefficients dans $\mathbb{K}$, la relation $A \sim B$ si et seulement si il existe une matrice $P \in \mathrm{GL}(n, \mathbb{K})$ telle que $B = P^{-1}AP$.

Cette relation est une relation d'équivalence.

*Deux matrices équivalentes en ce sens sont dites **semblables**.*

Proposition-Définition 9.188 (Matrices équivalentes)**.**
Nous définissons, sur l'ensemble $\mathbb{M}(n, \mathbb{K})$ des matrices $n \times n$ à coefficients dans $\mathbb{K}$, la relation $A \sim B$ si et seulement si il existe des matrice inversibles P et Q telles que $A = PBQ^{-1}$.

Cette relation est une relation d'équivalence.

*Deux telles matrice sont dite **équivalentes**.*

Proposition-Définition 9.189.
Soit un espace vectoriel E. Nous définissons sur $\mathrm{End}(E)$ la relation $u \sim v$ si et seulement si il existe une application inversible $A \colon E \to E$ telle que $v = A^{-1} \circ u \circ A$.

Cette relation est une relation d'équivalence.

*Deux endomorphismes équivalents en ce sens sont dits **semblables**.*

Proposition 9.190.
Deux applications linéaires sont semblables si et seulement si leurs matrices sont semblables dans toute base.

Lemme 9.191.
Le polynôme caractéristique[72] est un invariant sous les similitudes.

Démonstration. En effet si P est une matrice inversible,

$$\begin{aligned} \chi_{P^{-1}AP} &= \det(P^{-1}AP - \lambda X) && (9.370a) \\ &= \det\left(P(P^{-1}AP - \lambda X)P^{-1}\right) && (9.370b) \\ &= \det(A - \lambda X) \qquad = \chi_A. && (9.370c) \end{aligned}$$

□

La permutation de lignes ou de colonnes ne sont pas des similitudes, comme le montrent les exemples suivants :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}. \tag{9.371}$$

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.

71. Lemme 4.62.
72. Définition 9.106.

9.11.2 Trace de matrices semblables

Proposition 9.192 ([1]).
Invariance de la trace par similarité.

(1) *Soit une application linéaire f. Si la matrice de f dans une base est A et est B dans une autre base, alors*
$$\mathrm{Tr}(A) = \mathrm{Tr}(B). \tag{9.372}$$

(2) *Deux opérateurs semblables*[73] *ont même trace.*

Démonstration. Les matrices A et B sont liées par la proposition 4.112 : $B = Q^{-1}AQ$ où Q est la matrice qui lie les vecteurs des deux bases. L'invariance cyclique de la trace donnée en le lemme 4.62 implique que
$$\mathrm{Tr}(B) = \mathrm{Tr}(Q^{-1}AQ) = \mathrm{Tr}(QQ^{-1}A) = \mathrm{Tr}(A). \tag{9.373}$$
□

Lemme 9.193.
Soit une matrice A. Nous avons
$$\mathrm{Tr}(A^\dagger A) = \sum_{ij} |A_{ij}|^2. \tag{9.374}$$

Démonstration. Utilisant la proposition 9.165 pour les éléments de la matrice $A^\dagger$, nous avons
$$\mathrm{Tr}(A^\dagger A) = \sum_k (A^\dagger A)_{kk} = \sum_{kl} A^*_{lk} A_{lk} = \sum_{kl} |A_{lk}|^2. \tag{9.375}$$
□

Lemme 9.194 ([243]).
Si les matrice A et B sont unitairement semblables[74]*, alors*
$$\sum_{ij} |A_{ij}|^2 = \sum_{ij} |B_{ij}|^2. \tag{9.376}$$

Démonstration. Soit une matrice unitaire U telle que $B = UAU^\dagger$. Nous savons par le proposition 9.192 que des matrice semblables ont même trace. Or $B^\dagger B$ est unitairement semblable à $A^\dagger A$ parce que
$$B^\dagger B = UA^\dagger U^\dagger UAU^\dagger = UA^\dagger AU^\dagger. \tag{9.377}$$
Donc $\mathrm{Tr}(A^\dagger A) = \mathrm{Tr}(B^\dagger B)$. Le lemme 9.193 conclu. □

9.11.3 Endomorphismes nilpotents

La **trace** d'une matrice $A \in \mathrm{M}(n, \mathbb{K})$ est la somme de ses éléments diagonaux :
$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}. \tag{9.378}$$

Une propriété importante est son invariance cyclique.

Lemme 9.195.
Quelques propriétés de la trace.

(1) *Si A et B sont des matrices carrées, alors* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

(2) *La trace est un invariant de similitude.*

73. Opérateurs semblables, définition 9.187
74. Définition 9.187.

Démonstration. C'est un simple calcul :

$$\mathrm{Tr}(AB) = \sum_{ik} A_{ik}B_{ki} = \sum_{ik} A_{ki}B_{ik} = \sum_{ik} B_{ik}A_{ki} = \sum_{i}(BA)_{ii} = \mathrm{Tr}(BA) \tag{9.379}$$

où nous avons simplement renommé les indices $i \leftrightarrow k$.

En particulier, la trace est un invariant de similitude parce que $\mathrm{Tr}(ABA^{-1}) = \mathrm{Tr}(A^{-1}AB) = \mathrm{Tr}(B)$ par l'invariance cyclique démontrée en 4.62(2). □

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 9.196 ([101]).
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 (9.225) est

$$\chi_u = (-1)^n X^\alpha (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_r)^{\alpha_r}. \tag{9.380}$$

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 (9.380), nous voyons que le coefficient du terme X^{n-1} dans le polynôme caractéristique est

$$0 = \mathrm{Tr}(u^p) = \alpha_1\lambda_1^p + \dots + \alpha_r\lambda_r^p. \tag{9.381}$$

Donc les nombres $(\alpha_1, \dots, \alpha_r)$ sont une solution non triviale[75] du système

$$\begin{cases} \lambda_1 X_1 + \dots + \lambda_r X_r = 0 & (9.382a) \\ \quad\vdots & (9.382b) \\ \lambda_1^r X_1 + \dots + \lambda_r^r X_r = 0. & (9.382c) \end{cases}$$

Ce sont les équations (9.381) écrites pour $p = 1, \dots, r$. Le déterminant de ce système est

$$\lambda_1 \dots \lambda_r \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^{r-1} & \dots & \lambda_r^{r-1} \end{pmatrix} \neq 0, \tag{9.383}$$

qui est un déterminant de Vandermonde (proposition 9.11) valant

$$0 = \lambda_1 \dots \lambda_r \prod_{1 \leqslant i \leqslant j \leqslant r} (\lambda_i - \lambda_j). \tag{9.384}$$

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 (9.382). □

Proposition 9.197 ([244]).
Soit un $\mathbb{K}$-espace vectoriel E. Un endomorphisme $u \in \mathrm{End}(E)$ est nilpotent si et seulement si il existe une base de E dans laquelle la matrice de u est strictement triangulaire supérieure.

Démonstration. (i) $\Rightarrow$ Nous faisons la démonstration par récurrence sur la dimension de E. Lorsque $n = 1$ nous avons $u = (a)$ avec $a \in \mathbb{K}$. Puisque $a^k = 0$ pour un certain k nous avons $a = 0$ parce qu'un corps est toujours un anneau intègre[76].

75. Si $\alpha_1 = \dots = \alpha_r = 0$, alors les valeurs propres sont toutes nulles et la matrice est en réalité nulle dès le départ.
76. Lemme 1.242.

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

$$\{x, e_2, \dots, e_n\} \tag{9.385}$$

donnée par le théorème de la base incomplète 4.12. La matrice de u dans cette base s'écrit

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A & \\ 0 & & & \end{array}\right). \tag{9.386}$$

Un tout petit peu de calcul de produit de matrice montre que la matrice de u^k est de la forme

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A^k & \\ 0 & & & \end{array}\right). \tag{9.387}$$

Étant donné que l'endomorphisme u est nilpotent, 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).

(ii) $\Leftarrow$ Soit une base $\{e_1, \dots, e_n\}$ dans laquelle la matrice de u est strictement triangulaire supérieure.

Alors $u(e_1) = 0$ et plus généralement, $u(e_k) \in \operatorname{Span}\{e_1, \dots, e_{k-1}\}$. Voyez par récurrence que $u^l(e_k) \in \operatorname{Span}\{e_1, \dots, e_{k-l}\}$. Donc $u^l(e_k) = 0$ dès que $l \geqslant k$.

□

Lemme 9.198 ([1]).
Si $N \in \mathbb{M}(n, \mathbb{C})$ est une matrice nilpotente d'ordre de nilpotence r, alors $\{N^k\}_{k=0,\dots,r-1}$ est libre dans $\mathbb{M}(n, \mathbb{C})$.

Proposition 9.199 (Thème 39).
Soit E un espace de Banach (espace vectoriel normé complet[77]). Si $A \in \mathcal{L}(E, E)$ est nilpotente, alors $(\mathbb{1} - A)$ est inversible et son inverse est donné par

$$(\mathbb{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k, \tag{9.388}$$

où l'infini peut évidemment être remplacé par l'ordre de nilpotence de A.

Démonstration. En ce qui concerne la convergence de la somme, elle ne fait pas de doute 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 :

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n} (A^k - A^{k+1}) = \mathbb{1} - A^{n+1}. \tag{9.389}$$

Par conséquent

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \to 0. \tag{9.390}$$

La dernière limite est en réalité une égalité pour n assez grand. □

77. Définition 7.230.

9.11.4 Endomorphismes diagonalisables

Définition 9.200.
Une matrice est ***diagonalisable*** *si elle est semblable*[78] *à une matrice diagonale.*

Une application linéaire est diagonalisable si elle est semblable[79] *à une application linéaire diagonale.*

Autrement dit, A est diagonalisable si il existe un opérateur diagonal D et un opérateur inversible P tels que $PAP^{-1} = D$.

La proposition 9.190 nous assure que la notion de diagonalisabilité pour les matrices et pour les applications sont les mêmes.

Proposition 9.201.
Si A est un opérateur diagonalisable dont les valeurs propres sont λ_i, alors il existe un opérateur inversible Q tel que

$$A = Q^{-1}DQ \tag{9.391}$$

où D est l'opérateur diagonal contenant les λ_i sur sa diagonale.

Lemme 9.202.
Une matrice triangulaire supérieure avec des 1 sur la diagonale n'est diagonalisable que si elle est diagonale (c'est-à-dire si c'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 9.203.
Soit F un sous-espace stable par u. Soit une décomposition du polynôme minimal

$$\mu_u = P_1^{n_1} \dots P_r^{n_r} \tag{9.392}$$

où les P_i sont des polynômes irréductibles unitaires distincts. Si nous posons $E_i = \ker P_i^{n_i}$, alors

$$F = (F \cap E_1) \oplus \dots \oplus (F \cap E_r). \tag{9.393}$$

Théorème 9.204.
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[80].

(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.

78. Définition 9.187.
79. Définition 9.189.
80. Le polynôme *caractéristique*, lui, n'a pas spécialement ses racines simples ; il peut encore être de la forme

$$\chi_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}, \tag{9.394}$$

mais alors $\dim(E_{\lambda_i}) = \alpha_i$.

(i) **(2) implique (3)** Étant donné que $P(u) = 0$, il est dans l'idéal des polynômes annulateurs de u, et le polynôme minimal μ_u le divise parce que l'idéal des polynômes annulateurs est généré par μ_u par le théorème 6.40.

(ii) **(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

$$\mu_u(X) = (X - \lambda_1) \dots (X - \lambda_r) \tag{9.395}$$

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 9.85) nous enseigne que

$$E = \ker(u - \lambda_1) \oplus \dots \oplus \ker(u - \lambda_r). \tag{9.396}$$

Mais $\ker(u - \lambda_i)$ est l'espace propre $E_{\lambda_i}(u)$. Donc u est diagonalisable.

(iii) **(1) implique (4)** Soit $\{e_1, \dots, e_n\}$ une base qui diagonalise u, soit F un sous-espace de E et $\{f_1, \dots, f_r\}$ une base de F. Par le théorème 4.16(2), nous pouvons compléter la base de F par des éléments de la base $\{e_i\}$. Le complément ainsi construit est stable par u.

(iv) **(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.

(v) **(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

$$P(x) = (X - \lambda_1) \dots (X - \lambda_r). \tag{9.397}$$

Alors $P(u) = 0$. En effet si e_i est un vecteur propre pour la valeur propre λ_i,

$$P(u)e_i = \prod_{j \neq i} (u - \lambda_j) \circ (u - \lambda_i)e_i = 0 \tag{9.398}$$

par le lemme 9.84. Par conséquent $P(u)$ s'annule sur la base $\{e_i\}$.

(vi) **(5) implique (2)** Si la matrice A est diagonale alors le polynôme $P = \prod_{i=1}^n (X - A_{ii})$ est annulateur de A. En effet,

$$P(A)e_k = \prod_{i=1}^n (A - A_{ii})x = \prod_{i=1}^n \big(u(e_k) - A_{ii}e_k\big) = \prod_{i=1}^n \big(A_{kk}e_k - A_{ii}e_k\big) = 0 \tag{9.399}$$

parce que le facteur $i = k$ est nul.

(vii) **(3) implique (5)** le polynôme minimal de u s'écrit

$$\mu = (X - \lambda_1) \dots (X - \lambda_r), \tag{9.400}$$

et les espaces E_i du lemme 9.203 sont les espaces propres $E_i = \ker(u - \lambda_i)$. Nous avons donc une somme directe

$$E = E_1 \oplus \dots \oplus E_r. \tag{9.401}$$

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.

□

Corolaire 9.205.
Si u est diagonalisable et si F est un sous-espace stable par u, alors

$$F = \bigoplus_\lambda E_\lambda(u) \cap F \tag{9.402}$$

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 9.204, le polynôme μ_u est scindé et ne possède que des racines simples. Notons le

$$\mu_u(X) = (X - \lambda_1) \dots (X - \lambda_r). \tag{9.403}$$

Les espaces E_i du lemme 9.203 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 9.206.
Soient 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 9.81). Par conséquent $\mathrm{Span}\{E_{\lambda_i}(u)\}$ est de dimension $n = \dim(E)$ et u est diagonalisable. □

Voici un résultat de diagonalisation simultanée. Nous donnerons un résultat de trigonalisation simultanée dans le lemme 12.101.

Proposition 9.207 (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 $u_j\big(E_\lambda(u_i)\big) \subset E_\lambda(u_i)$.

(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 $u_j(x) \in E_\lambda(u_i)$. Nous avons

$$u_i\big(u_j(x)\big) = u_j\big(u_i(x)\big) = \lambda u_j(x). \tag{9.404}$$

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

$$E_{\lambda_k}(u_0) \neq E, \tag{9.405}$$

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 :

$$E = \bigoplus_k E_{\lambda_k}(u_0). \tag{9.406}$$

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

$$\Big(u_i|_{E_{\lambda_k}(u_0)}\Big)_{i\in I}. \tag{9.407}$$

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 9.208.
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^2-1 est donc scindé à racines simples et les involutions sont diagonalisables (9.204). 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

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{9.408a}$$

$$A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{9.408b}$$

Cette matrice A représente donc une involution, mais n'est pas diagonalisable. △

9.11.5 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 9.204(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 9.209.
La matrice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{9.409}$$

a pour polynôme caractéristique $\chi_A(X) = X^2$. C'est également son polynôme minimal, qui n'est pas à racine simple.

Il est par ailleurs facile de voir que le seul espace propre de A est $\mathrm{Span}\{(1,0)\}$ (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.

9.210.
Lorsque la diagonalisation n'est pas possible, il est souvent possible de trigonaliser. Les matrices triangulaires ne sont pas aussi faciles à manipuler que les matrices diagonales, mais c'est toujours ça de pris.

Nous étudierons ça plus tard, en 12.11.1 parce que ça va nécéssiter le théorème de d'Alembert.

9.11.6 Diagonalisation : cas réel

Lemme 9.211 (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

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & * & * & * & * \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \lambda_r & * & * \\ 0 & 0 & 0 & \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} & * \\ 0 & 0 & 0 & 0 & \begin{pmatrix} a_s & b_s \\ c_s & d_s \end{pmatrix} \end{pmatrix}. \tag{9.410}$$

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

$$Au + iAv = A(u + iv) = (\alpha + i\beta)(u + iv) = \alpha u - \beta v + i(\alpha v + \beta v), \tag{9.411}$$

et en égalisant les parties réelles et imaginaires,

$$Au = \alpha u - \beta v \tag{9.412a}$$
$$Av = \alpha v + \beta u. \tag{9.412b}$$

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 $\mathrm{Span}\{u, v\}$. 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 $\mathrm{Span}\{e_1, e_2\}$ est stable par $Q^{-1}AQ$, en effet nous avons

$$Q^{-1}AQe_1 = Q^{-1}Aq_1 = Q^{-1}(aq_1 + bq_2) = ae_1 + be_2. \tag{9.413}$$

La matrice $Q^{-1}AQ$ est donc de la forme

$$Q^{-1}AQ = \begin{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} & C_1 \\ 0 & A_1 \end{pmatrix} \tag{9.414}$$

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 conjugués.

En ce qui concerne les valeurs propres, il est facile de voir en regardant (9.410) que les valeurs propres sont celles des blocs diagonaux. Étant donné que $Q^{-1}AQ$ et A ont même polynôme caractéristique, ce sont les valeurs propres de A. □

Théorème 9.212 (Théorème spectral, matrice symétrique[89])**.**
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. Elle agit sur l'espace $\mathbb{C}^n$ par la définition 4.65, et en particulier la formule 4.78. Nous munissons de plus $\mathbb{C}^n$ de la forme sesquilinéaire définie en la proposition 9.170.

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 $\lambda = \bar{\lambda}$, et λ est réelle.

Le lemme de Schur réel 9.211 donne une matrice orthogonale Q qui trigonalise A. Les valeurs propres étant toutes réelles, la matrice $Q^{-1}AQ$ est même triangulaire (il n'y a pas de blocs dans la forme (9.410)). Prouvons que $Q^{-1}AQ$ est symétrique :

$$(Q^{-1}AQ)^t = Q^tA^t(Q^{-1})^t = Q^{-1}A^tQ = Q^{-1}AQ \tag{9.415}$$

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 telle que $A = Q^t D Q$ avec D diagonale. Nous posons $f_i = Q^t e_i$ et en tenant compte du fait que $Q^t = Q^{-1}$ nous avons $Af_i = Q^t D Q Q^t e_i = Q^t \lambda_i e_i = \lambda_i f_i$. Donc les f_i sont des vecteurs propres de A. De plus ils sont orthonormés parce qu'en utilisant la proposition 9.173,

$$\langle f_i, f_j \rangle = \langle Q^t e_i, Q^t e_j \rangle = \langle e_i, Q^t Q e_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}. \tag{9.416}$$

□

Le théorème spectral pour les opérateurs autoadjoints sera traité plus bas parce qu'il a besoin de notions sur les formes bilinéaires, théorème 11.6.

Remarque 9.213.
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 $P^t D P = A$, alors en notant $*$ les quantités qui ne dépendent pas de a, b ou c,

$$\begin{aligned}\begin{pmatrix} a & * & * \\ b & * & * \\ c & * & * \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}\begin{pmatrix} a & b & c \\ * & * & * \\ * & * & * \end{pmatrix} &= \begin{pmatrix} a & * & * \\ b & * & * \\ c & * & * \end{pmatrix}\begin{pmatrix} \lambda_1 a & \lambda_1 b & \lambda_1 c \\ * & * & * \\ * & * & * \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 a^2 + * & \lambda_1 ab + * & \lambda_1 ac + * \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}.\end{aligned} \tag{9.417}$$

Nous voyons donc que si nous changeons les signes de a, b et c en même temps, le résultat ne change pas.

Proposition 9.214.
Une forme bilinéaire est non-dénénérée[81] *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 9.212). 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 $(e_k)_j = \delta_{kj}$:

$$b(x, e_k) = \sum_{ij} x_i (e_k)_j = \sum_i b_{ik} x_i = b_{kk} x_k. \tag{9.418}$$

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. □

9.11.7 Matrice définie positive

Définition 9.215 (Matrice définie positive, opérateur défini positif)**.**
Un opérateur sur un espace vectoriel sur $\mathbb{C}$ ou $\mathbb{R}$ est ***défini positif*** *si toutes ses valeurs propres sont réelles et strictement positives. Il est* ***semi-défini positif*** *si ses valeurs propres sont réelles positives ou nulles.*

Mêmes définitions pour une matrice.

81. Définition 9.121.

Afin d'éviter l'une ou l'autre confusion, nous disons souvent *strictement* défini positif pour positif.

9.216.
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 9.217.
Nous ne définissons pas la notion de matrice définie positive pour une matrice non symétrique.

Proposition 9.218.
Soit M, une matrice symétrique. Nous avons

(1) $\det(M)>0$ *et* $\mathrm{Tr}(M)>0$ *implique M définie positive*[82],

(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.*

Lorsqu'un énoncé parle d'une matrice symétrique, le premier réflexe est de la diagonaliser : considérer une matrice orthogonale Q telle que $Q^tMQ=D$ avec D diagonale. Et les valeurs propres sur la diagonale : $D_{kk}=\lambda_k$. Les matrices symétriques définies positives ont cependant des propriétés même en dehors de leur base de diagonalisation.

Pour rappel, $\langle x,y\rangle$ est le produit scalaire dans $\mathbb{R}^n$ défini par la proposition 9.160.

Lemme 9.219.
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.

(i) **(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 9.212. Nous nommons x_i les coordonnées de x dans cette base. Alors,

$$\langle x,Mx\rangle=\sum_{i,j}x_i\langle e_i,x_jMe_j\rangle=\sum_{i,j}x_ix_j\langle e_i,\lambda_je_j\rangle=\sum_{ij}x_ix_j\lambda_j\delta_{ij}=\sum_i\lambda_ix_i^2 \tag{9.419}$$

où les λ_i sont les valeurs propres de M. Le produit$\langle x,Mx\rangle$ est strictement positif pour tout x si et seulement si tous les λ_i sont strictement positifs.

(ii) **(2)** Nous avons encore

$$\langle x,Mx\rangle=\sum_i\lambda_ix_i^2 \tag{9.420}$$

qui est plus grand ou égal à zéro si et seulement si tous les λ_i sont plus grands ou égaux à zéro.

(iii) **(3)** Soit une matrice orthogonale T diagonalisant M, c'est-à-dire telle que $T^tMT=D$ avec D diagonale. Nous allons vérifier que si $\lambda\leqslant\min\{\lambda_i\}$, alors

$$\langle Tx,MTx\rangle\geqslant\lambda\|Tx\|^2 \tag{9.421}$$

82. Définition 9.215.

pour tout x. Si nous considérons la base de diagonalisation $\{e_k\}$ pour les valeurs propres λ_k, nous avons le calcul

$$\begin{aligned}
\langle Tx, MTx\rangle &= \langle x, T^tMTx\rangle && (9.422a)\\
&= \langle x, Dx\rangle && (9.422b)\\
&= \sum_k \langle x, x_k De_k\rangle && (9.422c)\\
&= \sum_k \lambda_k x_k \underbrace{\langle x, e_k\rangle}_{=x_k} && (9.422d)\\
&\geqslant \sum_k \lambda |x_k|^2 \qquad \text{en posant} \lambda = \min\{\lambda_i\} && (9.422e)
\end{aligned}$$

Nous avons donc

$$\langle Tx, MTx\rangle \geqslant \sum_k \lambda |x_k|^2 = \lambda \|x\|^2 = \lambda \|Tx\|^2. \tag{9.423}$$

Au dernier passage nous avons utilisé le fait que T est une isométrie (proposition 9.38). L'inéquation (9.421) est démontrée.

Comme T est une bijection[83], cela implique le résultat pour tout x.

□

Les personnes qui aiment les vecteurs lignes et colonnes écriront des inégalités comme

$$x^tMx \geqslant x^tx. \tag{9.424}$$

Tout à l'autre bout du spectre des personnes névrosées des notations, on trouvera des inégalités comme

$$M(x \otimes x) \geqslant x \cdot x. \tag{9.425}$$

Le penchant personnel de l'auteur de ces lignes est la notation avec le produit tensoriel. Si vous aimez ça, vous pouvez lire la section 11.11.5 et en particulier ce qui suit (11.444).

La notation adoptée ici avec le produit scalaire $\langle x, Mx\rangle$, qui peut aussi être écrite $x \cdot Mx$ 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.

Proposition 9.220.
Une application bilinéaire est définie positive[84] *si et seulement si sa matrice symétrique associée l'est.*

Démonstration. La définition 9.127 dit que b est strictement définie positive lorsque $b(x,x) \geqslant 0$ et $b(x,x) = 0$ si et seulement si $x = 0$.

D'autre part, le lemme 9.219 dit que la matrice B est strictement définie positive lorsque $x \cdot Bx \geqslant 0$ et $x \cdot Bx = 0$ si et seulement si $x = 0$.

Le lien entre les deux est que le lemme 9.132 nous enseigne que pour tout x et y,

$$b(x,y) = x \cdot By \tag{9.426}$$

où B est la matrice de b. □

83. Une matrice orthogonale a un déterminant qui vaut ± 1.
84. Définition 9.127.

Proposition 9.221.
Soit une forme quadratique $q\colon E \to \mathbb{K}$ et sa matrice[85] $(q_{ij}) \in \mathrm{M}(n, \mathbb{K})$. *Nous avons*

$$q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j \tag{9.427a}$$

$$= \sum_{i=1}^{n} q_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} q_{ij} x_i x_j. \tag{9.427b}$$

9.222.
De nombreux auteurs préfèrent écrire des choses comme $x^t B y$ ou $x B^t y$ ou $x B y^t$ et se poser de longues questions sur qui est un « vecteur colonne » et qui est un « vecteur ligne », et si la matrice B soit être transposée ou non. Toutes ces notations servent(?) à cacher un bête produit scalaire.

9.223.
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.

Corolaire 9.224.
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 9.219. □

Lemme 9.225.
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 9.219. Si A est définie positive et que $\{e_i\}$ est une base, alors

$$A_{ii} = \langle A e_i, e_i \rangle \geqslant \lambda \| e_i \|^2 = \lambda \geqslant 0. \tag{9.428}$$

Si A est strictement définie positive, alors λ peut être choisi strictement positif. □

9.11.8 Réduction de Gauss

Théorème 9.226 (Réduction de Gauss[245, 235])**.**
Soit une forme quadratique non nulle q sur l'espace vectoriel E sur le corps $\mathbb{K}$. Il existe une base $\{l_i\}_{i=1,\dots,n}$ de E^* et des coefficients $\alpha_i \in \mathbb{K}$ tels que

$$q(x) = \sum_{i=1}^{n} \alpha_i l_i(x)^2. \tag{9.429}$$

Démonstration. Notre point de départ sont les formules (9.427) pour la forme quadratique. Nous allons faire la preuve par récurrence sur la dimension de l'espace. Si $n = 1$, alors nous avons seulement

$$q(x) = \alpha x^2 \tag{9.430}$$

et donc le théorème est fait avec $l(x) = x$.

Nous supposons que le théorème est prouvé pour tout espace de dimension n. Une forme quadratique pour un espace de dimension $n + 1$ s'écrit

$$q(x) = \sum_{i=1}^{n+1} m_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n+1} m_{ij} x_i x_j. \tag{9.431}$$

85. Matrice associée à une forme quadratique, définition 9.131.

Vu que q est non nulle, un des m_{ij} est non nul. Nous allons diviser en plusieurs cas.

— $m_{11} \neq 0$
— $m_{kk} \neq 0$ avec $k \neq 1$
— $m_{12} \neq 0$ et $m_{ii} = 0$ pour tout i.
— $m_{kl} \neq 0$ avec $(k,l) \neq (1,2)$ et $m_{ii} = 0$ pour tout i.

Ces cas ne sont pas exclusifs, mais ils couvrent toutes les possibilités.

(i) **Si $m_{11} \neq 0$** Nous écrivons q sous la forme

$$q(x) = m_{11}x_1^2 + \sum_{i=2}^{n+1} m_{ii}x_i^2 + 2\sum_{i=1}^{n}\big(\sum_{j=i+1}^{n+1} m_{ij}x_ix_j\big) \tag{9.432a}$$

$$= m_{11}x_1^2 + \sum_{i=2}^{n+1} m_{ii}x_i^2 + 2\sum_{j=2}^{n+1} m_{1j}x_1x_j + 2\sum_{i=2}^{n}\sum_{j=i+1}^{n+1}(m_{ij}x_ix_j) \tag{9.432b}$$

$$= m_{11}x_1^2 + 2x_1\sum_{j=2}^{n+1} m_{1j}x_k + R(x_2,\dots,x_{n+1}) \tag{9.432c}$$

$$= m_{11}\left(x_1^2 + 2x_1\sum_{j=2}^{n+1}\frac{m_{1j}}{m_{11}}x_j\right) + R(x_2,\dots,x_{n+1}) \tag{9.432d}$$

$$= m_{11}\big(x_1^2 + 2x_1 f(x_2,\dots,x_{n+1})\big) + R(x_2,\dots,x_{n+1}) \tag{9.432e}$$

$$= m_{11}\big(x_1 + f(x_2,\dots,x_{n+1})\big)^2 - f(x_2,\dots,x_{n+1}) + R(x_2,\dots,x_{n+1}) \tag{9.432f}$$

où

— R est une forme quadratique de $n-1$ variables ;
— nous avons noté $f(x_2,\dots,x_{n+1}) = \sum_{j=2}^{n+1}\frac{m_{1j}}{m_{11}}x_j$.

Maintenant, toute la partie $-f(x_2,\dots,x_{n+1})^2 + R(x_2,\dots,x_{n+1})$ est une forme quadratique de n variables. Par hypothèse de récurrence, il existe des coefficients α_i et des formes linéairement indépendantes sur $\mathbb{K}^n$ $l'_i(x_2,\dots,x_{n+1})$ telles que

$$-f(x_2,\dots,x_{n+1})^2 + R(x_2,\dots,x_{n+1}) = \sum_{i=2}^{n+1}\alpha_i l'_i(x_2,\dots,x_{n+1})^2. \tag{9.433}$$

En posant ensuite $l_j(x_1,\dots,x_{n+1}) = l'_j(x_2,\dots,x_{n+1})$, ainsi que $l_1(x_1,\dots,x_{n+1}) = x_1 + f(x_2,\dots,x_{n+1})$, nous avons

$$q(x) = m_{11}l_1(x)^2 + \sum_{j=2}^{n+1}\alpha_j l_j(x)^2. \tag{9.434}$$

(ii) **Si $m_{kk} \neq 0$ avec $k \neq 1$** Nous nommons k le plus petit entier pour lequel $m_{kk} \neq 0$, et nous supposons que $k \neq 1$, parce que nous avons déjà couvert ce cas. Dans ce cas, nous avons

$$q(x) = m_{kk}x_k^2 + \sum_{j=k+1}^{n+1} m_{jj}x_j^2 + 2\sum_{i=1}^{n}\big(\sum_{j=i+1}^{n+1} m_{ij}x_ix_j\big), \tag{9.435}$$

et tout tourne comme dans le premier cas.

(iii) **$m_{ii} = 0$ pour tout i et $m_{12} \neq 0$** Nous écrivons q en séparant les termes m_{1k} :

$$q(x) = 2\sum_{1\leqslant i<j\leqslant n+1} m_{ij}x_ix_j \tag{9.436a}$$

$$= 2m_{12}x_1x_2 + 2\sum_{2\leqslant j\leqslant n+1} m_{1j}x_1x_j + 2\sum_{2\leqslant i<j\leqslant n+1} m_{ij}x_ix_j \tag{9.436b}$$

$$= 2m_{12}x_1x_2 + 2x_1\sum_{2\leqslant j\leqslant n+1} m_{1j}x_j + 2\sum_{3\leqslant j\leqslant n+1} m_{2j}x_2x_j + 2\sum_{3\leqslant i<j\leqslant n+1} m_{ij}x_ix_j \tag{9.436c}$$

$$= 2m_{12}x_1x_2 + x_1 f(x_2,\dots,x_{n+1}) + x_2 g(x_3,\dots,x_{n+1}) + T(x_3,\dots,x_{n+1}) \tag{9.436d}$$

où f et g sont linéaires et T est multilinéaire.
À ce moment, nous tentons de factoriser toute la partie concernant x_1 et x_2. L'idée est d'utiliser ceci :

$$(x_1 + g)(x_2 + f) = x_1x_2 + x_1f + x_2g + fg, \tag{9.437}$$

mais en mettant les bons coefficients pour reproduire ce que nous avons dans (9.436d) :

$$(2m_{12} + 2g)(x_1 + \frac{f}{m_{12}}) - \frac{2fg}{m_{12}} = 2m_{12}x_1x_2 + 2x_1f + 2x_2g. \tag{9.438}$$

Cela pour dire que

$$q(x) = 2(m_{12}x_1 + g)(x_2 + \frac{f}{m_{12}}) - \frac{2fg}{m_{12}} + T \tag{9.439}$$

où $-2fg/m_{12} + T$ est une forme quadratique de $x_3, \dots, x_{n+1}$, c'est-à-dire de $n-1$ variables. L'hypothèse de récurrence nous donne des formes linéaires $(l_i)_{i=3,\dots,n+1}$ telles que

$$\frac{2fg}{m_{12}} + T = \sum_{i=3}^{n+1} \alpha_i l_i(x)^2. \tag{9.440}$$

Nous pouvons donc déjà écrire

$$q(x) = 2l_1'(x)l_2'(x) + \sum_{i=3}^{n+1} \alpha_i l_i(x)^2 \tag{9.441}$$

où

— Les forme l_i avec $i \geqslant 3$ ne dépendent pas de x_1 et x_2, et sont donc indépendantes de l_1 et l_2.
— La forme l_1' ne dépend pas de x_2,
— La forme l_2' ne dépend pas de x_1.

Ce sont donc $n+1$ formes linéaires indépendantes. Le seul problème résiduel est que les formes l_1' et l_2' arrivent en produit l'une de l'autre. Nous en définissons donc deux de plus :

$$\begin{aligned} l_1(x) &= \frac{1}{2}(l_1' + l_2') \\ l_2(x) &= \frac{1}{2}(l_1' - l_2'), \end{aligned} \tag{9.442}$$

qui sont linéairement indépendantes l'une de l'autre et indépendantes des l_i ($i \geqslant 3$). Au final,

$$q(x) = l_1(x)^2 + l_2(x)^2 + \sum_{i=3}^{n+1} \alpha_i l_i(x)^2. \tag{9.443}$$

(iv) **Si $m_{ii} = 0$ et $m_{12} = 0$ et $m_{kl} \neq 0$ avec $k < l$** Nous considérons la permutation

$$\begin{aligned} \sigma\colon \{1, \dots, n+1\} &\to \{1, \dots, n+1\} \\ i &\mapsto \begin{cases} 1 & \text{si } i = k \\ 2 & \text{si } i = l \\ k & \text{si } i = 1 \\ l & \text{si } i = 2 \\ i & \text{sinon,} \end{cases} \end{aligned} \tag{9.444}$$

c'est-à-dire que σ permute 1 et k ainsi que 2 et l. Ensuite nous posons

$$\begin{aligned} s\colon \mathbb{R}^{n+1} &\to \mathbb{R}^{n+1} \\ e_i &\mapsto e_{\sigma(i)}. \end{aligned} \tag{9.445}$$

Nous allons un peu considérer $q \circ s$, pour changer :

$$(q \circ s)(x) = \sum_{i,j} m_{ij} s(x)_i s(x)_j = \sum_{ij} x_{\sigma(i)} x_{\sigma(j)}. \tag{9.446}$$

parce que $s(x)_i = x_{\sigma(i)}$.
Utilisons un petit abus de notation pour considérer

$$\begin{aligned} \sigma\colon \{1,\dots,n+1\} \times \{1,\dots,n+1\} &\to \{1,\dots,n+1\} \times \{1,\dots,n+1\} \\ (i,j) &\mapsto \big(\sigma(i),\sigma(j)\big). \end{aligned} \tag{9.447}$$

Cela est une bijection ; nous pouvons utiliser le lemme 1.254 pour permuter les termes dans (9.446) :

$$(q \circ s)(x) = \sum_{ij} m_{\sigma(i)\sigma(j)} x_{\sigma\sigma(i)} x_{\sigma\sigma(j)} \tag{9.448a}$$

$$= \sum_{ij} a_{ij} x_i x_j \tag{9.448b}$$

où nous avons posé $a_{ij} = m_{\sigma(i)\sigma(j)}$ et utilisé le fait que $\sigma = \sigma^{-1}$. Le point intéressant de l'histoire est que dans (9.448b), $a_{12} = m_{kl} \neq 0$. La forme $q \circ s$ est donc dans le cas déjà traité et il existe des formes linéaires l'_i telles que

$$(q \circ s)(x) = \sum_{i=1}^{n+1} \alpha_i l'_i(x)^2. \tag{9.449}$$

En évaluant cela en $s(x)$, et en tenant compte de $s = s^{-1}$, nous trouvons

$$q(x) = \sum_i \alpha_i (l_i \circ s)(x)^2, \tag{9.450}$$

de telle sorte que $l_i = l'_i \circ s$ soit la réponse à notre théorème.

□

Théorème 9.227 ([235, 1]).
Toute forme quadratique sur un espace vectoriel de dimension finie admet une base formée de vecteurs 2 à 2 orthogonaux (pour la forme considérée).

Démonstration. Nous considérons la base $\{l_i\}$ de E^* donnée par la réduction de Gauss (théorème 9.226). La forme quadratique q s'écrit

$$q(x) = \sum_{i=1}^{n} \alpha_i l_i(x)^2. \tag{9.451}$$

La base préduale[86] $\{e_i\}$ de $\{l_i\}$ répond aux conditions. Pour le vérifier, nous considérons la forme bilinéaire associée à q par l'identité de polarisation 9.128 :

$$b(e_i, e_j) = \frac{1}{2}\big(q(e_i) + q(e_j) - q(e_i - e_j)\big). \tag{9.452}$$

Vu que $l_k(e_i) = \delta_{ki}$, nous avons

$$q(e_i) = \sum_{k=1}^{n} \alpha_k l_k(e_i)^2 = \alpha_i. \tag{9.453}$$

86. Définition, existence, unicité dans la proposition 4.123.

En utilisant la linéarité,

$$q(e_i - e_j) = \sum_k \alpha_k l_k(e_i - e_j)^2 \tag{9.454a}$$
$$= \sum_k \alpha_k (\delta_{ki} - \delta_{kj})^2 \tag{9.454b}$$
$$= \sum_k \alpha_k (\delta_{ki} + \delta_{kj} - 2\delta_{ki}\delta_{ki}) \tag{9.454c}$$
$$= \alpha_i + \alpha_j - 2\delta_{ij}\alpha_i. \tag{9.454d}$$

Donc

$$b(e_i, e_j) = \delta_{ij}\alpha_i. \tag{9.455}$$

Les vecteurs $\{e_i\}$ sont donc bien deux à deux q-orthogonaux. □

Notons qu'en l'absence de notion de racine carrée sur $\mathbb{K}$, il n'est pas possible de considérer $\sqrt{\alpha_i}$ et donc de base q-orthonormée.

Proposition 9.228.
Si $A \in \mathbb{M}(n, \mathbb{K})$ est telle que $\det(A) = 0$, alors il existe des matrices de manipulation de lignes et de colonnes $G_1, \dots, G_N$ telles que $G_1 \dots G_N A$ ait une colonne de zéros.

Démonstration. Si la matrice A elle-même n'a pas de colonnes de zéros, alors nous pouvons faire un pas de réduction de Gauss [87] et obtenir des matrices $G_1, \dots, G_{N_1}$ telles que

$$G_1 \dots G_{N_1} A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A^{(1)} & \\ 0 & & & \end{pmatrix}. \tag{9.456}$$

Si $A^{(1)}$ ne possède pas de colonnes de zéros, nous pouvons continuer.

Si nous parvenons à faire n pas de la sorte, alors nous aurions

$$G_1 \dots G_N A = \delta, \tag{9.457}$$

et donc $\det(G_1 \dots G_N)\det(A) = 1$, ce qui est impossible lorsque $\det(A) = 0$. Nous en concluons que le processus doit s'arrêter et qu'une des matrices $A^{(k)}$ doit avoir une colonne de zéros [88]. □

Proposition 9.229.
Une matrice dont le déterminant est nul n'est pas inversible.

Démonstration. Par la proposition 9.228, il existe des matrices de manipulation de lignes et de colonnes $G_1, \dots, G_N$ telles que la matrice $G_1 \dots G_N A$ ait une colonne de zéros. De là, la proposition 4.92 implique que la matrice

$$G_1 \dots G_N A \tag{9.458}$$

n'est pas inversible. Vu les déterminants des matrices G_i, la proposition 4.91 implique que $G_1 \dots G_N$ est inversible. Si A était inversible, nous aurions

$$G_1 \dots G_N A A^{-1} (G_1 \dots G_N)^{-1} = \delta, \tag{9.459}$$

c'est-à-dire que $A^{-1}(G_1 \dots G_N)^{-1}$ serait un inverse de la matrice (9.458). Cette dernière n'ayant pas d'inverse, nous concluons que A n'en a pas non plus. □

87. La réduction de Gauss est le théorème 9.226, mais le lien avec ce que nous disons ici n'est peut-être pas directement clair.

88. En réalité, le processus tel que nous l'avons décrit ne s'arrête que lorsque la première colonne est remplie de zéros.

Théorème 9.230.
Une matrice sur un corps commutatif est inversible si et seulement si son déterminant est non nul.

Démonstration. Dans un sens c'est la proposition 4.91 et dans l'autre sens c'est la proposition 9.229. □

Proposition 9.231.
Soient des matrices A et B sur un corps commutatif. Alors

$$\det(AB) = \det(A)\det(B). \tag{9.460}$$

Démonstration. Les propositions 4.95 et 4.96 ont déjà fait une grosse partie du travail. Il ne reste que le cas où $\det(A) = \det(B) = 0$.

Dans ce cas, les matrices A et B ne sont pas inversibles (proposition 9.230). Le produit AB n'est alors pas inversible non plus [89]. La proposition 9.230, utilisée dans le sens inverse, nous dit alors que $\det(AB) = 0$.

Au final dans le cas $\det(A) = \det(B) = 0$ nous avons $0 = \det(AB) = \det(A)\det(B) = 0$. □

Faisons maintenant le cas général des manipulations de lignes et colonnes.

Proposition 9.232.
Soit une matrice carrée $A \in \mathbb{M}(n, \mathbb{K})$. La matrice B obtenue par la substitution simultanée

$$C_j \to \sum_k a_{kj} C_k \tag{9.461}$$

a pour déterminant

$$\det(B) = \det(a)\det(A). \tag{9.462}$$

Démonstration. L'élément B_{ij} de la matrice B est une combinaison linéaire de tous les éléments de sa ligne :

$$B_{ij} = \sum_k a_{kj} A_{ik} = (Aa)_{ij}. \tag{9.463}$$

Donc $B = Aa$. La proposition 9.231 nous dit alors que $\det(B) = \det(a)\det(A)$. □

Théorème 9.233 (de Sylvester[238])**.**
Soit Q une forme quadratique réelle de signature [90] (p, q). Alors pour toute base Q-orthogonale $\{e_i\}$ de $\mathbb{R}^{p+q}$ nous avons les propriétés suivantes.

(1) Les nombres p et q sont donnée par

$$p = \mathrm{Card}\{i \text{ tel que } Q(e_i) > 0\} \tag{9.464a}$$
$$q = \mathrm{Card}\{i \text{ tel que } Q(e_i) < 0\}. \tag{9.464b}$$

(2) Si A est la matrice de Q dans une base, alors il existe une matrice inversible P telle que

$$P^t A P = \begin{pmatrix} -\mathbb{1}_q & & \\ & \mathbb{1}_p & \\ & & 0 \end{pmatrix}. \tag{9.465}$$

(3) Le rang de Q est $p + q$.

89. Citez le lemme 4.88 si vous voulez justifier ça.
90. Définition 9.147.

Démonstration. Soit F un sous-espace de dimension maximale q sur lequel Q est définie négative. Le fait que la dimension de F soit q est la définition 9.147 de la signature. Nous notons $F^\perp$ sont Q-orthogonal, c'est-à-dire que

$$F^\perp = \{v \in E \text{ tel que } B(v,x) = 0 \,\forall x \in F\}. \tag{9.466}$$

Le lemme 9.152 nous assure que $E = F \oplus F^\perp$.

Le théorème 9.227 sur l'existence de bases Q-orthogonales nous permet de considérer une base Q-orthogonale de F et une de $F^\perp$. En réunissant les deux, nous avons une base de E. Nous la notons $\{f_1, \dots, f_n\}$ avec

— La partie $\{f_1, \dots, f_q\}$ est une base de F,

— La partie $\{f_{q+1}, \dots, f_n\}$ est une base de $F^\perp$,

— Remarquez cependant qu'il n'est pas dit que $n = q + p$.

Notons que pour $i > q$, nous avons $Q(f_i) \geqslant 0$, sinon la maximalité de F serait contredite par $\mathrm{Span}\{f_1, \dots, f_q, f_i\}$.

Cela prouve que

$$\mathrm{Card}\{i \text{ tel que } Q(f_i) > 0\} = p. \tag{9.467}$$

Le lemme 9.150 nous dit alors que

$$\mathrm{Card}\{i \text{ tel que } Q(e_i) > 0\} = \mathrm{Card}\{i \text{ tel que } Q(f_i) > 0\} = p. \tag{9.468}$$

C'est l'égalité (9.464b). L'égalité (9.464a) se prouve de la même façon, en prenant F maximal pour la propriété que Q y est strictement définie positive.

Le point (1) est prouvé.

Dans une base Q-orthogonale, la matrice de Q est diagonale, et contient sur la diagonale les valeurs de $Q(e_i)$. Parmi celles-ci, on en a p strictement positives et q strictement négatives. Les $n - p - q$ autres sont nulles. Vu que Q est à valeur réelle, nous avons une notion de racine carré, et nous pouvons considérer $e_i/\sqrt{|Q(e_i)|}$ au lieu de e_i. De cette façon, $Q(e_i)$ est normalisé. Avec ça, la matrice de Q est

$$D = \begin{pmatrix} \mathbb{1}_p & & \\ & -\mathbb{1}_q & \\ & & 0 \end{pmatrix}. \tag{9.469}$$

Nous venons de prouver qu'il existe une base $\{e_i\}$ dans laquelle la matrice de Q est (9.469). Si A est la matrice de Q dans une base quelconque $\{f_i\}$ et si P est la matrice de changement de base $f_j = \sum_i P_{ij} e_i$, la proposition 9.133 donne $D = P^t A P$.

Le point (2) est prouvé.

Pour (3), la proposition 9.149 nous permet de calculer le rang de Q par le rang de sa matrice dans n'importe quelle base. Nous choisissons la base qui donne la matrice (9.469). Le rang est alors bien $p + q$. □

9.11.9 Diagonalisation

Lemme-Définition 9.234.
Soit une forme quadratique[91] *q sur l'espace vectoriel V sur $\mathbb{K}$. Soit A la matrice de q dans la base $\{e_i\}$ et B sa matrice dans la base $\{f_\alpha\}$. Nous supposons que le changement de base est orthogonal.*

Alors les valeurs propres de A et B sont les mêmes.

Ces valeurs sont les ***valeurs propres*** *de q.*

Démonstration. Nous nous rappelons de la définition 9.131 de la matrice associée à Q, et à la proposition 9.133 qui parle de changement de base : $B = Q^t A Q$ où Q est orthogonale.

91. Définition 9.126.

Soit un vecteur propre v de A, de valeur propre λ. Alors nous prouvons que $Q^t v$ est un vecteur propre pour B, de même valeur propre λ. En effet,

$$BQ^t v = Q^t AQQ^t v = Q^t Av = \lambda Q^t v \tag{9.470}$$

où nous avons utilisé $QQ^t = \mathbb{1}$ et $Av = \lambda v$. □

Proposition 9.235.
Dans la base de diagonalisation de sa matrice associée, une forme quadratique a la forme

$$q(x) = \sum_i \lambda_i x_i^2 \tag{9.471}$$

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[92] de la matrice de b alors dans cette base nous avons

$$q(x) = b(x,x) = \sum_{ij} x_i x_j b(f_i, f_j) = \sum_i \lambda_i x_i^2 \tag{9.472}$$

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 justifie la définition pour dire que nous nous permettrons de parler des **valeurs propres** d'une forme quadratique comme étant les valeurs propres de la matrice associée.

9.11.10 Diagonalisation

Le théorème 9.227 a déjà donné une base orthogonale pour toute forme quadratique sur un espace vectoriel $(E, \mathbb{K})$ de dimension finie. Dans le cas de $\mathbb{R}^n$, nous pouvons en donner une preuve basée sur le théorème spectral, c'est la proposition 9.236.

Proposition 9.236.
Soit une forme bilinéaire symétrique b sur un $\mathbb{R}^n$. Il existe une matrice orthogonale Q telle que

(1) $D = Q^t bQ$ est diagonale

(2) $D(x,y) = b(Qx, Qy)$ pour tout $x, y \in E$.

Il existe une base $(f_i)_{i=1,\dots,n}$ qui est b-orthogonale.

Dans cet énoncé, nous mélangeons sans vergogne les formes et les matrices, en supposant qu'une base soit fixée[93]*. Par exemple*

$$D(x,y) = \sum_{ij} D_{ij} x_i y_j. \tag{9.473}$$

Démonstration. Pour la matrice diagonale, c'est le théorème spectral 9.212(2) qui joue parce que la matrice d'une forme bilinéaire symétrique est symétrique (c'est vu de la définition (9.270)).

Pour le reste c'est un calcul :

$$\begin{aligned}
D(x,y) &= \sum_{ijkl} Q^t_{ik} b_{kl} Q_{lj} x_i y_j && (9.474a)\\
&= \sum_{ijkl} b_{kl} (Q_{ki} x_i)(Q_{lj} y_j) && (9.474b)\\
&= \sum_{kl} b_{kl} (Qx)_k (Qy)_l && (9.474c)\\
&= b(Qx, Qy). && (9.474d)
\end{aligned}$$

92. Qui existe parce que la matrice est symétrique, théorème 9.212.

93. Autrement dit, si vous avez en tête d'utiliser cette proposition pour $\mathbb{R}^n$ c'est bon ; mais sinon vous devez choisir une base et considérer toutes les matrices dans cette base.

Nous avons utilisé le produit matrice fois vecteur donné par (4.78).

En ce qui concerne l'existence d'une base b-orthogonale, vu que D est diagonale, nous avons, pour $i \neq j$ que $D(e_i, e_j) = 0$. Donc en posant $f_i = Qe_i$, nous trouvons

$$0 = D(e_i, e_j) = b(Qe_i, Qe_j) = b(f_i, f_j). \tag{9.475}$$

La base $(Qe_i)_{i=1,\dots,n}$ est donc b-orthogonale. □

9.12 Fonctions

Soient $(E, \|.\|_E)$ et $(F, \|.\|_F)$ deux espaces vectoriels normés, et une fonction f de E dans F. 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 $\mathbb{R}$ et en changeant simplement les valeurs absolues par les normes sur E et F.

La proposition suivante explicite la définition 7.92 dans le cas où la topologie est donnée par des boules.

Proposition 9.237 (Caractérisation de la limite)**.**
Soient des espaces vectoriels normés. Soit $f\colon E \to F$ une fonction de domaine $\mathrm{Dom}(f) \subset E$ et soit a un point d'accumulation de $\mathrm{Dom}(f)$.

(1) Si F est séparé[94] *et si f admet une limite en a, alors cette limite est unique.*

(2) La fonction f admet une limite en $a \in E$ si et seulement si il existe un élément $\ell \in F$ tel que pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $x \in D = \mathrm{Dom}(f)$,

$$0 < \|x - a\|_E < \delta \Rightarrow \|f(x) - \ell\|_F < \varepsilon. \tag{9.476}$$

*Si la limite existe et est unique, nous écrivons $\lim_{x\to a} f(x) = \ell$ et nous disons que ℓ est la **limite** de f lorsque x tend vers a.*

Démonstration. L'unicité est la proposition 7.95.

(i) $\Rightarrow$ La définition 7.92 nous assure de l'existence d'un élément ℓ tel que pour tout voisinage S de ℓ, il existe un ouvert U autour de a tel que $f\big(U \cap D\backslash\{a\}\big) \subset S$.

Soit $\epsilon > 0$. Nous posons $S = B(\ell, \epsilon)$. Il existe un voisinage U de a tel que $f\big(U \cap D\backslash\{a\}\big) \subset B(\ell, \epsilon)$. Puisque S est un voisinage de a, il contient une boule centrée en a (c'est dans la définition 7.98 de la topologie métrique). Soit donc $\delta > 0$ tel que $B(a, \delta) \subset U$.

Un élément de D qui est dans $B(0, \delta)\backslash\{a\}$ est un élément de D qui vérifie $0 < \|x - a\| < \delta$. Nous avons donc, pour $x \in D$ que

$$0 < \|x - a\| < \delta \Rightarrow \|f(x) - \ell\| < \epsilon. \tag{9.477}$$

(ii) $\Leftarrow$ C'est le même raisonnement.

□

Remarque 9.238.
Le fait que nous limitions la formule (9.476) 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

$$\lim_{x\to 2} \sqrt{x^2 - 4} = 0. \tag{9.478}$$

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 $[2, \infty[$.

94. C'est le cas en dimension finie et en particulier pour $\mathbb{R}^n$. En dimension infinie, il faut être très prudent.

9.13 Sous espaces caractéristiques

Lorsqu'un opérateur n'est pas diagonalisable, les valeurs propres jouent quand même un rôle important.

Définition 9.239.
Soient E un $\mathbb{K}$-espace vectoriel et $f \in \mathrm{End}(E)$. Pour $\lambda \in \mathbb{K}$ nous définissons

$$F_\lambda(f) = \{v \in E \text{ tel que } (f - \lambda\mathbb{1})^n v = 0, n \in \mathbb{N}\} \tag{9.479}$$

et nous appelons cet ensemble 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 9.240.
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. Réciproquement, si v est un vecteur 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

$$(f - \lambda\mathbb{1})^n f(x) = f\big((f - \lambda\mathbb{1})^n x\big) = 0, \tag{9.480}$$

par conséquent $f(x) \in F_\lambda(f)$. □

Contrairement à ce que l'on pourrait 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 9.241.
Le théorème 9.204 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

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{9.481}$$

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

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.482}$$

qui revient au système

$$\begin{cases} x + y = x & \text{(9.483a)} \\ y = y. & \text{(9.483b)} \end{cases}$$

La première équation donne directement $y = 0$. Le seul espace propre est de dimension 1 et est engendré par $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. △

Nous donnons maintenant un exemple un peu plus avancé de matrice réelle non diagonalisable, qui montre la multiplicité algébrique et géométrique d'une racine d'un polynôme caractéristique.

Remarque 9.242.
Considérons l'endomorphisme $f \in \mathrm{End}(\mathbb{C}^3)$ donné par la matrice

$$\begin{pmatrix} a & \alpha & \beta \\ 0 & a & \gamma \\ 0 & 0 & b \end{pmatrix} \tag{9.484}$$

avec $a \neq b$, $a \neq 0$, $b \neq 0$, $\alpha \neq 0$, β et γ sont des nombres complexes quelconques. Son polynôme caractéristique est

$$\chi_f(X) = (a - X)^2(b - X), \tag{9.485}$$

et les valeurs propres sont donc a et b. Nous trouvons les vecteurs propres pour la valeur a en résolvant

$$\begin{pmatrix} a & \alpha & \beta \\ 0 & a & \gamma \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax \\ ay \\ az \end{pmatrix}. \tag{9.486}$$

La troisième équation est $bz = az$ qui oblige $z = 0$ parce que $a \neq b$ et $0 \neq a$. La première est $ax + \alpha y = ax$ qui implique $y = 0$ parce que $\alpha \neq 0$. Enfin la première équation se réduit à $ax = ax$ qui ne donne pas de contraintes sur x. En résumé : 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 le système

$$\begin{pmatrix} a & \alpha & \beta \\ & a & \gamma \\ & & b \end{pmatrix} \begin{pmatrix} w \\ y \\ z \end{pmatrix} = \begin{pmatrix} bx \\ by \\ bz \end{pmatrix} \tag{9.487}$$

La seconde équation donne $ay + \gamma z = by$, et donc

$$y = \frac{\gamma}{b - a} z. \tag{9.488}$$

La première équation est $ax + \alpha y + \beta z = bx$ qui donne

$$x = \frac{1}{b - a}(\alpha y + \beta z). \tag{9.489}$$

En y remettant la valeur déjà trouvée de y, nous trouvons que l'espace propre pour la valeur propre b est engendré par le vecteur

$$\begin{pmatrix} \frac{1}{b-a}\left(\beta + \frac{\alpha\gamma}{b-a}\right) \\ \frac{\gamma}{b-a} \\ 1 \end{pmatrix}. \tag{9.490}$$

Vu que nous savons que a et b sont les seules valeurs propres et que nous venons de voir que leurs espaces propres sont de dimension 1, 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 $(f - a\mathbb{1})^2 e_2 = 0$. En effet :

$$(f - a\mathbb{1})^2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & b- \end{pmatrix} \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{9.491}$$

de telle sorte que le vecteur $(0, 1, 0)$ est 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.

9.13.1 Théorèmes de décomposition

Théorème 9.243 (Théorème spectral, décomposition primaire)**.**
Soit E un espace vectoriel de dimension finie sur le corps algébriquement clos $\mathbb{K}$ et $f \in \mathrm{End}(E)$. Alors

$$E = F_{\lambda_1}(f) \oplus \ldots \oplus F_{\lambda_k}(f) \tag{9.492}$$

où la somme est sur les espaces caractéristiques engendrés par les valeurs propres distinctes de f.

Les projecteurs sur les espaces caractéristiques forment un système complet et orthogonal.

Démonstration. Soit P le polynôme caractéristique de f et une décomposition

$$P = (f - \lambda_1)^{\alpha_1} \ldots (f - \lambda_r)^{\alpha_r} \tag{9.493}$$

en facteurs irréductibles. Par le théorème des noyaux (9.85) nous avons

$$E = \ker(f - \lambda_1)^{\alpha_1} \oplus \ldots \oplus \ker(f - \lambda_r)^{\alpha_r}. \tag{9.494}$$

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

$$\ker(f - \lambda_i)^{\alpha_i} \subset F_{\lambda_i}(f) \tag{9.495}$$

est évidente. Nous devons montrer l'inclusion inverse.

(i) $\underline{F_{\lambda_i}(f) \cap F_{\lambda_j}(f) = 0}$ Soit $v \in F_{\lambda_i}(f) \cap F_{\lambda_j}(f)$. Le fait que $v \in F_{\lambda_i}(f)$ implique qu'il existe $n \in \mathbb{N}$ tel que $(f - \lambda_i)^n v \neq 0$ et $(f - \lambda_i)^{n+1} v = 0$ (éventuellement $n = 0$ si v est un vecteur propre). Posons $v_1 = (f - \lambda_i)^n v$.
Étant donné que $(f - \lambda_i)$ commute avec $(f - \lambda_j)$, ce v_1 est encore dans $F_{\lambda_j}(f)$. En effet, si k est tel que $(f - \lambda_j)^k v = 0$, alors

$$(f - \lambda_j)^k v_1 = (f - \lambda_j)^k (f - \lambda_i)^n v = (f - \lambda_i)^n (f - \lambda_j)^k v = 0. \tag{9.496}$$

Il existe donc $m \in \mathbb{N}$ tel que $(f - \lambda_j)^m v_1 \neq 0$ et $(f - \lambda_j)^{m+1} v_1 = 0$. En posant $w = (f - \lambda_j)^m v_1$, nous avons

$$\begin{cases} (f - \lambda_i) w = (f - \lambda_j)^m (f - \lambda_i)^{n+1} v = 0 & \text{(9.497a)} \\ (f - \lambda_j) w = (f - \lambda_j)^{m+1} v_1 = 0. & \text{(9.497b)} \end{cases}$$

Ce w serait donc un vecteur propre simultané pour les valeurs propres λ_i et λ_j. Vu que les espaces propres sont linéairement indépendants, les seules possibilités sont $i = j$ ou $v = 0$.

(ii) **Questions de dimension** Étant donné que espaces F_{λ_i} sont en somme directe, la somme de leurs dimensions est au maximum la dimension de E :

$$\sum_i \dim F_{\lambda_i}(f) \leqslant \dim E. \tag{9.498}$$

En tenant compte de l'inclusion (9.495) nous avons même

$$\dim E = \sum_i \dim \ker(f - \lambda_i)^{\alpha_i} \leqslant \sum_i \dim F_{\lambda_i}(f) \leqslant \dim E. \tag{9.499}$$

Vu qu'il y a $\dim(E)$ des deux côtés des inégalités, toutes les inégalités sont des égalités et nous avons

$$\sum_i \dim \ker(f - \lambda_i)^{\alpha_i} = \sum_i \dim F_{\lambda_i}(f). \tag{9.500}$$

L'inclusion (9.495) nous dit qu'il y a une inégalité terme à terme dans les sommes de (9.500). Vu qu'il y a égalité des sommes, il y a en réalité égalité de chacun des termes : $\dim \ker(f - \lambda_i)^{\alpha_i} = \dim F_{\lambda_i}(f)$ et l'égalité des deux espaces de (9.495) :

$$\ker(f - \lambda_i)^{\alpha_i} = F_{\lambda_i}(f). \tag{9.501}$$

□

¡¡ Avertissement/question au lecteur !! 9.244
Dans le cas où le corps n'est pas algébriquement clos, il paraît qu'il faut remplacer « diagonalisable » par « semi-simple ».

Si vous connaissez un énoncé précis et une démonstration, écrivez-moi.

Il y a peut-être une réponse dans [246].

Si l'espace vectoriel est sur un corps algébriquement clos, alors les endomorphismes semi-simples[95] sont les endomorphismes diagonaux.

Théorème 9.245 (Décomposition de Dunford)**.**
Soit E un espace vectoriel sur le corps algébriquement clos[96] $\mathbb{K}$ *et* $u \in \mathrm{End}(E)$ *un endomorphisme de E.*

(1) L'endomorphisme u se décompose de façon unique sous la forme

$$u = s + n \tag{9.502}$$

où s est diagonalisable, n est nilpotent et $[s, n] = 0$[97]*.*

(2) Les endomorphismes s et n sont des polynômes en u et commutent avec u.

(3) Si notons $\{\lambda_i\}$ les valeurs propres distinctes de u, et $F_{\lambda_i}(u)$ les espaces caractéristiques correspondants, alors les parties s et n sont données par

$$s = \sum_i \lambda_i p_i \tag{9.503a}$$

$$n = \sum_i (s - \lambda_i \mathbb{1}) p_i \tag{9.503b}$$

$p_i \colon E \to F_{\lambda_i}(u)$ est la projection de E sur $F_{\lambda_i}(u)$.

Démonstration. Le théorème spectral 9.243 nous indique que

$$E = \bigoplus_i F_{\lambda_i}(f). \tag{9.504}$$

Nous considérons l'endomorphisme s de E qui consiste à dilater d'un facteur λ_i l'espace caractéristique $F_{\lambda_i}(f)$:

$$s = \sum_i \lambda_i p_i \tag{9.505}$$

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

$$f - s \colon F_\lambda(f) \to F_\lambda(f). \tag{9.506}$$

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

$$s'u = s'^2 + s'n' = s'^2 + n's' = (s' + n')s' = us', \tag{9.507}$$

95. Définition 9.101.

96. Je crois quo'n peut remplacer l'ypothèse de corps alébriquement clos par une hypothèse de polynôme caractéristique scindé. Écrivez-moi si vous avez une idée à ce propos.

97. Lorsque a et b sont des opérateurs, la notation $[a, b]$ signifie le commutateur entre a et b, c'est à dire $a \circ b - b \circ a$. Dire que $[a, b] = 0$ signifie que $ab = ba$.

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 et s' sont alors deux endomorphismes diagonalisables qui commutent. Par la proposition 9.207, ils sont simultanément diagonalisables. Dans la base de diagonalisation simultanée, 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,

$$(A + B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}. \tag{9.508}$$

Si n est assez grand, au moins un parmi A^k ou B^{n-k} est nul.

Nous savons que $n - n'$ est diagonal et nilpotent. Le seul opérateur diagonal à être nilpotent est l'opérateur nul[98]. Nous en déduisons que $n = n'$. Nous avons alors immédiatement aussi $s = s'$. □

9.246.
Le théorème 12.417 montrera que $A^n x \to 0$ pour tout x si et seulement si $\rho(A) < 1$, mais ça demande un résultat de vitesse comparée entre l'exponentielle et la puissance.

Une application de la décomposition de Jordan est l'existence d'un logarithme pour les matrices. La proposition suivante 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 15.143 une formule pour le logarithme sous forme d'une série ; ce logarithme sera réel.

9.13.2 Valeurs singulières

Définition 9.247.
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 si il existent des vecteurs unitaires $u \in \mathbb{K}^m$, $v \in \mathbb{K}^n$ tels que*

$$Mv = \sigma u \tag{9.509a}$$
$$M^* u = \sigma v. \tag{9.509b}$$

Théorème 9.248 (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

$$M = ADB \tag{9.510}$$

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[99] *de M.*

Corolaire 9.249.
Soit $M \in \mathbb{M}(n, \mathbb{C})$. Il existe un isomorphisme $f \colon \mathbb{C}^n \to \mathbb{C}^n$ 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. □

98. Parce qu'une puissance d'un opérateur diagonal est diagonal.
99. Définition 4.43.

9.14 Extension du corps de base

Nous avons discuté dans la section 6.4 de ce qui arrive au corps lorsqu'on l'étend. Dans cette section 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 $\mathbb{L}$.

Soient $\mathbb{K}$ un corps (commutatif) et une extension $\mathbb{L}$ de $\mathbb{K}$. Soient E et F, des $\mathbb{K}$-espaces vectoriels de dimension finie, et entrons dans le vif du sujet [100].

9.14.1 Extension des applications linéaires

Définition 9.250 ([247]).
L'espace vectoriel obtenu par ***extension du corps de base*** *de E est l'espace vectoriel*

$$E_{\mathbb{L}} = \mathbb{L} \otimes_{\mathbb{K}} E. \tag{9.511}$$

Ce dernier est le quotient $\mathbb{L} \otimes_{\mathbb{K}} E = (\mathbb{L} \times E)/\sim$ par la relation d'équivalence

$$(\lambda, v) \sim \big(a\lambda, \frac{1}{a}v\big) \tag{9.512}$$

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) pour la relation d'équivalence $\sim$:

$$[\lambda, v] = \{(\mu, w) \in \mathbb{L} \times E \text{ tel que } (\mu, w) \sim (\lambda, v)\}. \tag{9.513}$$

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 $f\colon E \to F$ est une applications linéaire, nous définissons

$$\begin{aligned} f_{\mathbb{L}}\colon E_{\mathbb{L}} &\to F_{\mathbb{L}} \\ [\lambda, v] &\mapsto [\lambda, f(v)]. \end{aligned} \tag{9.514}$$

Remarque 9.251.
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.

Lemme 9.252.
Soient trois $\mathbb{K}$-espaces vectoriels E, F et G ainsi que deux applications linéaires $f\colon E \to F$ et $g\colon G \to E$. Si $\mathbb{L}$ est une extension de $\mathbb{K}$, alors

$$f_{\mathbb{L}} \circ g_{\mathbb{L}} = (f \circ g)_{\mathbb{L}}. \tag{9.515}$$

Démonstration. Il suffit de composer la définition (9.514) :

$$\begin{aligned} (f_{\mathbb{L}} \circ g_{\mathbb{L}})\big([\lambda, v]\big) &= f_{\mathbb{L}}\big([\lambda, g(v)]\big) && (9.516a) \\ &= [\lambda, (f \circ g)(v)] && (9.516b) \\ &= (f \circ g)_{\mathbb{L}}\big([\lambda, v]\big). && (9.516c) \end{aligned}$$

□

Nous définissons aussi l'injection canonique

$$\begin{aligned} \tau\colon E &\to E_{\mathbb{L}} \\ v &\mapsto [1, v]. \end{aligned} \tag{9.517}$$

Proposition 9.253 ([113]).
Injectivité et surjectivité respectées.

100. Le sujet étant le corps étendu.

(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 $f_{\mathbb{L}}$ soit injective. Le diagramme

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\tau"'] & F \arrow[d, "\tau"] \\ E_{\mathbb{L}} \arrow[r, "f_{\mathbb{L}}"'] & F_{\mathbb{L}} \end{tikzcd} \tag{9.518}$$

est un diagramme commutatif. En effet

$$(\tau \circ f)(v) = [1, f(v)] \tag{9.519}$$

tandis que

$$(f_{\mathbb{L}} \circ \tau)(v) = f_{\mathbb{L}}[1, v] = [1, f(v)]. \tag{9.520}$$

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 4.53(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

$$(f_{\mathbb{L}} \circ g_{\mathbb{L}})[\lambda, v] = f_{\mathbb{L}}[\lambda, g(v)] = [\lambda, (f \circ g)(v)] = [\lambda, v]. \tag{9.521}$$

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 9.254 ([1, 248]).
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

$$\lambda \otimes ke_i = \frac{\lambda}{k} \otimes e_i = \frac{\lambda}{k}(1 \otimes e_i). \tag{9.522}$$

Cela pour introduire l'application

$$\begin{aligned} \psi\colon \mathbb{L} \otimes_{\mathbb{K}} E &\to \bigoplus_i \mathbb{L}(1 \otimes e_i) \\ \sum_k \lambda_k \otimes v_k &\mapsto \oplus_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) \end{aligned} \tag{9.523}$$

où $v_k = \sum_i v_{ik} e_i$ avec $v_{ik} \in \mathbb{K}$, qui représente un isomorphisme de $\mathbb{L}$-espaces vectoriels. La surjectivité est facile. En ce qui concerne l'injectivité, si

$$\sum_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) = 0 \tag{9.524}$$

alors les quantités suivantes sont nulles également :

$$\sum_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) = \sum_{ik} (\lambda_k \otimes v_{ik} e_i) = \sum_k (\lambda_k \otimes \sum_i v_{ik} e_i) = \sum_k (\lambda_k \otimes v_k). \tag{9.525}$$

La dernière est l'argument de ψ. Le fait qu'il soit nul implique que ψ est injective. □

Remarque 9.255.
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, puisqu'il y a plusieurs façons de noter un élément de $E_{\mathbb{L}}$ sous la forme de tels termes.

Corolaire 9.256.
La $\mathbb{L}$-dimension de $E_{\mathbb{L}}$ est égale à la $\mathbb{K}$-dimension de E.

9.14.2 Projections

¡¡ Avertissement/question à la lectrice !! 9.257
Nous allons définir `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 quantités définies ne dépendent pas des bases choisies. Il y a surement 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 6.56). 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

$$\mathbb{L} = \mathbb{K} \oplus V. \tag{9.526}$$

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 $\tau(e_i)$ et $\tau(e_a)$ correspondants. Grâce à la proposition 9.254, 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 ces bases comme

$$f(e_i) = \sum_a f_{ai} e_a \tag{9.527}$$

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

$$(\texttt{proj}\, f)e_i = \sum_a \texttt{proj}(f_{ai}) e_a. \tag{9.528}$$

Proposition 9.258 ([1]).
L'application `proj` *définie en* (9.528) *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 quantités qui se notent `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 à priori des bases choisies.

Nous choisissons de nouvelles bases qui sont liées aux anciennes bases par

$$\begin{cases} e'_b = \sum_a B_{ab} e_a & (9.529a) \\ e'_j = \sum_i A_{ji} e_i. & (9.529b) \end{cases}$$

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 :

$$\begin{aligned} f(e'_j) &= \sum_i A_{ji} f(e_i) && (9.530a) \\ &= \sum_a \sum_b \sum_i A_{ji} f_{ai} (B^{-1})_{ba} e'b && (9.530b) \\ &= \sum_b \Big(\sum_{ai} A_{ji} f_{ai} (B^{-1})_{ba} \Big) e'b, && (9.530c) \end{aligned}$$

ce qui donne

$$(\texttt{proj}'\, f)e'_j = \sum_b \Big(\texttt{proj}\, \big(A_{ji} f_{ai} (B^{-1})_{ba} \big) \Big) e'_b. \tag{9.531}$$

Nous calculons maintenant $(\mathtt{proj}'\, f)e_j$ en substituant $e_j = \sum_l (A^{-1})_{lj} e'_l$ et en utilisant (9.531), la linéarité de $\mathtt{proj}'$ et la $\mathbb{K}$-linéarité de $\mathtt{proj}\colon \mathbb{L} \to \mathbb{K}$:

$$(\mathtt{proj}'\, f)\Big(\sum_l (A^{-1})_{lj} e'_l\Big) = \sum_l (A^{-1})_{lj} \sum_b \sum_{ai} \mathtt{proj}\,\big(A_{li} f_{ai} (B^{-1})_{ba}\big) e_b \tag{9.532a}$$

$$= \sum_a \mathtt{proj}(f_{aj}) e_a \tag{9.532b}$$

$$= (\mathtt{proj}\, f) e_j. \tag{9.532c}$$

Donc $\mathtt{proj} = \mathtt{proj}'$. □

Note au passage : 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 9.259 ([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 $\mathtt{proj}\colon \mathbb{L} \to \mathbb{K}$ 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 suite nous « fixons » un choix de V générique, et lorsque nous rencontrerons l'égalité $\mathtt{proj}(\lambda) = 0$ nous en déduirons $\lambda = 0$.

Proposition 9.260.
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) $(f_{\mathbb{L}})_{aj} = f_{aj} \in \mathbb{K}$.

Démonstration. Nous avons

$$f_{\mathbb{L}}(e_i) = \sum_a f_{ai}(1 \otimes e_a) = \sum_a f_{ai} \tau(e_a), \tag{9.533}$$

donc

$$(\mathtt{proj}\, f_{\mathbb{L}}) e_i = \sum_a \mathtt{proj}(f_{ai}) e_a = \sum_a f_{ai} e_a = f(e_i). \tag{9.534}$$

Cela prouve que $\mathtt{proj}\, f_{\mathbb{L}} = f$.

Par ailleurs,

$$f_{\mathbb{L}}(\tau(e_i)) = f_{\mathbb{L}}(1 \otimes e_i) = 1 \otimes f(e_i) = \tau\big(f(e_i)\big) = \sum_a f_{ai} \tau(e_a) \tag{9.535}$$

alors que par définition,

$$f_{\mathbb{L}}(\tau(e_i)) = \sum_a (f_{\mathbb{L}})_{ai} \tau(e_a). \tag{9.536}$$

Les éléments $\tau(e_a)$ formant une base[101], la comparaison de (9.535) avec (9.536) donne $(f_{\mathbb{L}})_{ai} = f_{ai} \in \mathbb{K}$. □

Lemme 9.261.
Soient des espaces vectoriels E, F et G ainsi que des applications linéaires

(1) $f\colon E \to F$,

(2) $\tilde{h}\colon G_{\mathbb{L}} \to E_{\mathbb{L}}$.

101. Encore la proposition 9.254.

Alors nous avons

$$\mathtt{proj}(f_{\mathbb{L}} \circ \tilde{h}) = \mathtt{proj}(f_{\mathbb{L}}) \circ \mathtt{proj}(\tilde{h}). \tag{9.537}$$

De même[102] *si* $f\colon E \to F$ *et* $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$, *alors*

$$\mathtt{proj}(\tilde{h} \circ f_{\mathbb{L}}) = \mathtt{proj}(\tilde{h}) \circ \mathtt{proj}(f_{\mathbb{L}}). \tag{9.538}$$

Démonstration. Nous considérons une base $\{e_i\}$ de E, une base $\{e_a\}$ de F, et une base $\{e_\alpha\}$ de G. Pour écrire $\mathtt{proj}(f_{\mathbb{L}} \circ \tilde{h})$ à partir de la définition (9.528) nous commençons par écrire

$$(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a (f_{\mathbb{L}} \circ \tilde{h})_{a\alpha} e_a = \sum_{ai} (f_{\mathbb{L}})_{ai} (\tilde{h})_{i\alpha} e_a = \sum_a \Big(\sum_i f_{ai} (\tilde{h})_{i\alpha} \Big) e_a \tag{9.539}$$

où nous avons utilisé le fait que $(f_{\mathbb{L}})_{ai} = f_{ai}$. Donc, en utilisant la $\mathbb{K}$-linéarité de $\mathtt{proj}$,

$$\mathtt{proj}(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a \sum_i \mathtt{proj}\Big(f_{ai} (\tilde{h})_{i\alpha} \Big) e_a = \sum_a \sum_i f_{ai}\, \mathtt{proj}\Big((\tilde{h})_{i\alpha} \Big) e_a. \tag{9.540}$$

D'autre part,

$$\begin{aligned} \mathtt{proj}(f_{\mathbb{L}}) \circ \mathtt{proj}(\tilde{h}) e_\alpha &= \mathtt{proj}(f_{\mathbb{L}}) \sum_i \mathtt{proj}\Big((\tilde{h})_{i\alpha} \Big) e_i \\ &= \sum_i \mathtt{proj}\Big((\tilde{h})_{i\alpha} \Big) \sum_a f_{ai} e_a \\ &= \sum_{ai} \mathtt{proj}\Big((\tilde{h})_{i\alpha} \Big) f_{ai} e_a, \end{aligned} \tag{9.541}$$

et c'est égal à (9.540). □

Remarque 9.262.
Nous n'avons en général pas $\mathtt{proj}(xy) = \mathtt{proj}(x)\,\mathtt{proj}(y)$ pour tout $x, y \in \mathbb{L}$. Par exemple si $\mathbb{K} = \mathbb{R}$ et $\mathbb{L} = \mathbb{C}$ avec la projection canonique,

$$\mathtt{proj}(i \cdot i) = \mathtt{proj}(-1) = -1 \tag{9.542}$$

alors que $\mathtt{proj}(i) = 0$.

Proposition 9.263.
Soient trois $\mathbb{K}$*-espaces vectoriels* E, F *et* G. *Nous considérons deux applications linéaires* $f\colon E \to F$ *et* $g\colon E \to G$.

Il existe une application linéaire $h\colon F \to G$ *telle que* $g = h \circ f$ *si et seulement si il existe une application linéaire* $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$ *telle que* $g_{\mathbb{L}} = \tilde{h} \circ f_{\mathbb{L}}$.

Démonstration. Dans le sens direct, il suffit de poser $\tilde{h} = h_{\mathbb{L}}$ et d'invoquer le lemme 9.252.

Dans le sens inverse, si nous avons $\tilde{h}\colon F_{\mathbb{L}} \to G_{\mathbb{L}}$ tel que $g_{\mathbb{L}} = \tilde{h} \circ f_{\mathbb{L}}$ alors en appliquant $\mathtt{proj}$ des deux côtés et en utilisant le lemme 9.261,

$$\mathtt{proj}(g_{\mathbb{L}}) = \mathtt{proj}(\tilde{h}) \circ \mathtt{proj}(f_{\mathbb{L}}) \tag{9.543}$$

c'est-à-dire

$$g = \mathtt{proj}(\tilde{h}) \circ f, \tag{9.544}$$

c'est-à-dire que l'application $\mathtt{proj}(\tilde{h})\colon F \to G$ est la réponse à la proposition. □

102. Je n'ai pas vérifié cette deuxième partie. Soyez prudente.

9.14.3 Rang, polynôme minimal, polynôme caractéristique

Proposition 9.264 (Stabilité du rang par extension des scalaires[113]).
Si $f\colon E \to F$ est linéaire alors nous avons

$$\mathrm{rk}(f) = \mathrm{rk}(f_{\mathbb{L}}). \tag{9.545}$$

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 $\dim(V) = \mathrm{rk}(f)$. Nous pouvons factoriser f en

$$f = f_2 \circ f_1 \tag{9.546}$$

avec $f_1\colon E \to V$ est la projection parallèle à $\ker(f)$ et est surjective (vers V) parce que $\dim(V) = \mathrm{rk}(f) = \dim\big(\mathrm{Image}(f)\big)$. 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

$$f(v) = (f_2 \circ f_1)(v) = f_2(v) = 0. \tag{9.547}$$

Cela donne $v \in \ker(f) \cap V = \{0\}$. Par la proposition 9.253, 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

$$(f_2)_{\mathbb{L}}\colon V_{\mathbb{L}} \to \mathrm{Image}(f_{\mathbb{L}}) \tag{9.548}$$

est un isomorphisme de $\mathbb{L}$-espaces vectoriels. Nous avons alors les égalités

$$\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{L}}\big(\mathrm{Image}(f_{\mathbb{L}})\big) = \mathrm{rk}(f_{\mathbb{L}}). \tag{9.549}$$

Mais aussi, par les définitions posées plus haut,

$$\dim(V) = \mathrm{rk}(f) = \dim\big(\mathrm{Image}(f)\big). \tag{9.550}$$

Mais le corolaire 9.256 nous dit que $\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{K}}(V)$. Donc il y a égalité des deux lignes (9.549) et (9.550), ce qui donne $\mathrm{rk}(f) = \mathrm{rk}(f_{\mathbb{L}})$. □

Proposition 9.265.
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 9.260(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 9.106) : les coefficients de ce polynôme sont des polynômes 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 [103]. 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 polynômes en les coefficients de la matrice). La proposition suivante montre que le polynôme minimal est conservé par extension de corps, mais que pour le savoir, il faut travailler plus.

Proposition 9.266 ([113, 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 $\mu_f = \mu_{f_{\mathbb{L}}}$.

103. Définition 6.60.

Démonstration. Nous allons montrer que l'application

$$\begin{aligned}\tilde{g}\colon \frac{\mathbb{L}[X]}{(\mu)} &\to \operatorname{End}(E_{\mathbb{L}})\\ [P] &\mapsto P(f_{\mathbb{L}})\end{aligned} \tag{9.551}$$

est bien définie et injective. La proposition 9.96 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 $P(f_{\mathbb{L}}) = P(f)_{\mathbb{L}}$:

$$P(f_{\mathbb{L}})\lambda\otimes v = \sum_k a_k f_{\mathbb{L}}^k \lambda\otimes v \tag{9.552a}$$

$$= \lambda\otimes\sum_k a_k f^k(v) \tag{9.552b}$$

$$= \lambda\otimes P(f)v \tag{9.552c}$$

$$= P(f)_{\mathbb{L}}\lambda\otimes v. \tag{9.552d}$$

Par conséquent $\mu(f_{\mathbb{L}}) = 0$ et l'application est bien définie.

Sur $\mathbb{L}[X]/(\mu)$ nous considérons la base $\{1, [X], \dots, [X^{\deg(\mu)-1}]\}$, et sur $\operatorname{End}(E_{\mathbb{L}})$ nous considérons une base qui commence[104] 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

$$\sum_k \mathtt{proj}\left(\lambda_k f_{\mathbb{L}}^k\right) = 0. \tag{9.553}$$

Pour détailler ce que cela implique, nous calculons ceci :

$$(\lambda f_{\mathbb{L}})(\tau(e_i)) = \lambda f_{\mathbb{L}}(\tau(e_i)) = \sum_a \lambda f_{ai} e_a, \tag{9.554}$$

par conséquent $\mathtt{proj}(\lambda f_{\mathbb{L}})e_i = \sum_a \mathtt{proj}(\lambda f_{ai})e_a$, et comme $\mathtt{proj}$ est $\mathbb{K}$-linéaire et que $f_{ai}\in\mathbb{K}$,

$$\mathtt{proj}(\lambda f_{\mathbb{L}})e_i = \mathtt{proj}(\lambda)\sum_a f_{ai}e_a = \mathtt{proj}(\lambda)\,\mathtt{proj}(f_{\mathbb{L}})e_i = \mathtt{proj}(\lambda)f(e_i). \tag{9.555}$$

Appliquer la projection $\mathtt{proj}$ à l'équation (9.553) donne alors $\sum_k \mathtt{proj}(\lambda)_k f^k = 0$. Mais comme les f^k sont linéairement indépendants sur $\mathbb{K}$ nous avons pour tout k : $\mathtt{proj}(\lambda_k) = 0$ (égalité dans $\mathbb{K}$). En nous souvenant de la remarque 9.259 nous en déduisons $\lambda_k = 0$ dans $\mathbb{L}$.

Dans les choix de bases effectués, l'application $\tilde{g}$ a la forme

$$\tilde{g} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}, \tag{9.556}$$

qui est injective.

Puisque $\tilde{g}$ est injective, μ est le polynôme minimal de $f_{\mathbb{L}}$ et donc $\mu = \mu_{\mathbb{L}}$. □

9.15 Frobenius et Jordan

9.15.1 Matrice compagnon

Définition 9.267.
*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*

104. Théorème de la base incomplète 4.12(2).

P est la matrice donnée par

$$C(P) = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & a_{n-1} \end{pmatrix} \tag{9.557}$$

si $n \geqslant 2$ et par (a_0) si $n = 1$.

Une matrice est dite compagnon si elle a cette forme.

Proposition 9.268.
Si f est l'endomorphisme associé à la matrice $C(P)$ nous avons

$$f(e_i) = \begin{cases} e_{i+1} & \text{si } 1 \leqslant i < n \\ (a_0, \dots, a_{n-1}) & \text{si } i = n. \end{cases} \tag{9.558}$$

De plus l'endomorphisme f vérifie $P(f)e_1 = 0$.

Lemme 9.269 ([249])**.**
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, $\big(f^k(e_1)\big)_{k=1,\dots,n-1}$ est libre, donc le polynôme minimal ponctuel en e_1 est de degré n au minimum. En reprenant les notations du théorème 6.40, 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 9.270.
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 égal au polynôme minimal sont denses dans l'ensemble des 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.

9.15.2 Réduction de Frobenius

Lemme 9.271.
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. Si f est cyclique, alors $\mu = \chi$.

Le théorème 9.284 donnera une version plus complète de ce lemme.

Démonstration. Soit v un vecteur cyclique de f, c'est-à-dire que $\{f^k(v)\}_{k=0,\dots,n-1}$ 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 9.114), donc il est le polynôme minimal. □

Théorème 9.272 (Réduction de Frobenius [250, 251, 252])**.**
Soit E, un $\mathbb{K}$-espace vectoriel, et $f \in \mathrm{End}(E)$. Alors il existe une suite de sous-espaces $E_1, \dots, E_r$ 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

$$\chi_f = \prod_{i=1}^{r} \mu_i \tag{9.559a}$$

$$\mu_f = \mu_1 \tag{9.559b}$$

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 caractéristiques des $f|_{E_i}$, mais ces derniers endomorphismes étant cycliques[105], leurs polynômes caractéristiques sont égaux à leurs polynômes minimaux (lemme 9.271). Cela prouve l'égalité (9.559a). 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 9.97). Le plus petit espace stable sous f contenant y est

$$E_y = \operatorname{Span}\{y, f(y), \dots, f^{d-1}(y)\}. \tag{9.560}$$

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

$$\big(\sum_k a_k X^k\big)(f)y = 0, \tag{9.561}$$

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[106] $\{e_1, \dots, e_d\}$ en une base $\{e_1, \dots, e_n\}$ de E. Ensuite nous allons montrer que

$$E = E_y \oplus F \tag{9.562}$$

avec

$$F = \{x \in E \text{ tel que } e_d^*\big(f^k(x)\big) = 0, \forall k \in \mathbb{N}\}. \tag{9.563}$$

qui fait appel aux définitions d'espace dual (définition 4.119) et de base duale (définition 4.120).

Par construction, F est invariant sous f. Montrons pour commencer que $E_y \cap F = \{0\}$. Un élément de E_y s'écrit

$$z = a_1 e_1 + \cdots + a_k e_k \tag{9.564}$$

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 $a_1 = \dots = a_d = 0$, c'est-à-dire que $E_y \cap F = \{0\}$.

Nous montrons maintenant que $\dim F = n - d$. Pour cela nous considérons l'application

$$\begin{aligned} T\colon \mathbb{K}[f] &\to E^* \\ g &\mapsto e_d^* \circ g. \end{aligned} \tag{9.565}$$

105. Définition 9.98.

106. Pour autant que j'aie compris, cette extension manque dans [250]. Corrigez-moi si je me trompe.

Cette application est injective. En effet un élément général de $\mathbb{K}[f]$ est

$$g = a_1 \operatorname{Id} + a_2 f + \cdots + a_p f^{p-1} \tag{9.566}$$

avec $p \leqslant d$. Si $T(g) = 0$, alors nous avons en particulier

$$0 = T(g)e_{d-p+1} = e_d^*(a_1 e_{d-p+1} + a_2 e_{d-p+2} + \cdots + a_p e_d) = a_p. \tag{9.567}$$

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, $\dim \operatorname{Image}(T) = d$. Étudions l'orthogonal[107] de l'image de T :

$$\begin{aligned}
(\operatorname{Image}(T))^\perp &= \{x \in E \text{ tel que } T(g)x = 0, \forall g \in \mathbb{K}[f]\} && (9.568a)\\
&= \{x \in E \text{ tel que } e_d^*\big(g(x)\big) = 0, \forall g \in \mathbb{K}[f]\} && (9.568b)\\
&= F. && (9.568c)
\end{aligned}$$

Par conséquent $F^\perp = \operatorname{Image}(T)$. Puisque $\dim \operatorname{Image}(T) = d$, 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 $f|_{E_y}$, alors par le lemme 9.99 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 $F_1, \dots, F_r$ et $G_1, \dots, G_s$ 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 (9.559). Nous supposons que $P_i = Q_i$ pour $1 \leqslant i \leqslant j-1$ et nous tentons de montrer que $P_j = Q_j$.

Nous avons

$$P_j(f) = P_j(f)|_{F_1} \oplus \ldots \oplus P_j(f)|_{F_{j-1}}. \tag{9.569}$$

En effet étant donné que P_{j+k} divise P_j, nous avons[108] $P_j(f) = A(f) \circ P_{j+k}(f)$, mais $P_{j+k}(f)F_{j+k} = 0$, donc $P_j(f)F_{j+k} = 0$. Les espaces G_i n'ayant à priori aucun rapport avec les polynômes P_i, nous écrivons

$$P_j(f) = P_j(f)|_{G_1} \oplus \ldots \oplus P_j(f)|_{G_{j-1}} \oplus P_j(f)|_{G_j} \oplus \ldots \oplus P_j(f)|_{G_s}. \tag{9.570}$$

Pour $1 \leqslant i \leqslant j-1$, nous avons supposé $P_i = Q_i$. Étant donné que la matrice de $f|_{F_i}$ est semblable à C_{P_i} et celle de $f|_{G_i}$ est semblable à C_{Q_i}, la matrice de $f|_{E_i}$ est semblable à la matrice de $f|_{G_i}$. En particulier,

$$\dim P_j(f)F_i = \dim P_j(f)G_i. \tag{9.571}$$

En prenant les dimensions des images dans les égalités (9.569) et (9.570), nous trouvons que

$$P_j(f)|_{G_j} = \ldots = P_j(f)|_{G_s} = 0. \tag{9.572}$$

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.

□

107. Définition 4.124.

108. En vertu du lemme 9.84.

Remarque 9.273.
Sous forme matricielle, ce théorème dit que toute matrice est semblable à une matrice de la forme bloc-diagonale

$$f = \begin{pmatrix} C_{\mu_1} & & \\ & \ddots & \\ & & C_{\mu_r} \end{pmatrix} \tag{9.573}$$

où les C_{μ_i} sont des matrices compagnons (définition 9.267).

En particulier, et ceci est très important, deux applications sont semblables si et seulement si elles ont même suite d'invariants de similitude.

Remarque 9.274.
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.

9.15.3 Forme normale de Jordan

Il existe une preuve directe de la réduction de Jordan ne nécessitant pas la réduction de Frobenius[253]. Cette dernière passe par les espaces caractéristiques [109] 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 réduction de Frobenius.

Théorème 9.275 (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[110]*. Il existe une base de E dans laquelle la matrice de f s'écrit sous la forme*

$$M = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \\ 0 & \ddots & \\ & & J_{n_k}(\lambda_k) \end{pmatrix} \tag{9.574}$$

où les λ_i sont les valeurs propres de f (avec éventuelles répétitions) et $J_{n_i}(\lambda)$ représente le bloc $n_i \times n_i$

$$J_{n_i}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & & \\ & \lambda & 1 & 0 & \\ 0 & & \lambda & & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}. \tag{9.575}$$

En d'autres termes, $J_{n_i}(\lambda)_{ii} = \lambda$ 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 9.269 que (la matrice de) f est semblable à sa matrice compagnon. En l'occurrence pour f nous avons

$$C_{X^m} = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix}. \tag{9.576}$$

109. Aussi appelés « espaces propres généralisés ».
110. C'est à cause de cette hypothèse que $\mathbb{K} = \mathbb{R}$ n'est pas le bon exemple de contexte de travail.

Ensuite le changement de base (qui est une similitude) $(e_1, e_2, \dots, e_n) \mapsto (e_n, \dots, e_2, e_1)$ 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

$$\chi_f = (X - \lambda_1)^{l_1} \dots (X - \lambda_m)^{l_m}. \tag{9.577}$$

Le lemme des noyaux (théorème 9.85) nous enseigne que

$$E = \bigoplus_{i=1}^{m} \underbrace{\ker(f - \lambda_i \mathbb{1})^{l_i}}_{F_i}. \tag{9.578}$$

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

$$f|_{F_i} = (f - \lambda_i \mathbb{1})|_{F_i} + \lambda_i \mathbb{1}_{F_i}, \tag{9.579}$$

nous voyons que $f|_{F_i}$ s'écrit comme un bloc de Jordan avec λ_i sur la diagonale. □

Remarque 9.276.
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 la réduction de Frobenius, elle, est disponible sur tout corps, y compris $\mathbb{R}$.

9.16 Commutant et endomorphismes cycliques

9.16.1 Endomorphisme cyclique

Lemme 9.277.
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

$$E_{f,x} = \mathrm{Span}\{f^k(x) \text{ tel que } k \in \mathbb{N}\}. \tag{9.580}$$

Définition 9.278 (Commutant $\mathcal{C}(f)$ de f)**.**
Soit E un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et un endomorphisme $f\colon E \to E$. Le ***commutant*** *de f est l'ensemble des endomorphismes de E qui commutent avec f :*

$$\mathcal{C}(f) = \{g \in \mathcal{L}(E, E) \text{ tel que } g \circ f = f \circ g\}. \tag{9.581}$$

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.

9.16.2 Commutant : cas diagonalisable

Proposition 9.279 ([254]).
Si f est diagonalisable, alors

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim(E_\lambda)^2. \tag{9.582}$$

où les E_λ sont les espaces propres de f.

Démonstration. D'abord, soit $g \in \mathcal{C}(f)$ 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

$$\begin{aligned}\psi\colon \mathcal{C}(f) &\to \mathrm{End}(E_1)\times\ldots\times\mathrm{End}(E_r)\\ g &\mapsto g|_{E_1}\times\ldots\times g|_{E_r}\end{aligned} \tag{9.583}$$

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.

(i) <u>ψ **est injective**</u> Supposons que $g, h \in \mathcal{C}(f)$ tels que $\psi(g) = \psi(h)$. Puisque f est diagonalisable nous pouvons décomposer $x \in E$ en ses composantes sur les espaces propres [111] :

$$x = \sum_{\lambda\in\mathrm{Spec}(f)} x_\lambda \tag{9.584}$$

avec $x_\lambda \in E_\lambda$. Nous avons alors

$$g(x) = \sum_\lambda g(x_\lambda) = \sum_\lambda \psi(g)_\lambda(x_\lambda). \tag{9.585}$$

Par hypothèse nous avons $\psi(g)_\lambda = \psi(h)_\lambda$, et donc aussi

$$g(x) = \sum_\lambda \psi(g)_\lambda(x_\lambda) = \sum_\lambda \psi(h)_\lambda(x_\lambda) = \sum_\lambda h(x_\lambda) = h(x). \tag{9.586}$$

Cela prouve $g = h$ et donc que ψ est injective.

(ii) <u>ψ **est surjective**</u> Si nous avons pour chaque $\lambda \in \mathrm{Spec}(f)$ un endomorphisme g_λ de E_λ alors en posant

$$g(x) = \sum_{\lambda\in\mathrm{Spec}(f)} g_\lambda(x_\lambda) \tag{9.587}$$

nous avons bien

$$\psi(g) = \big(g_{\lambda_1},\ldots,g_{\lambda_r}\big). \tag{9.588}$$

Nous pouvons donc conclure en écrivant

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim\big(\mathrm{End}(E_\lambda)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim(E_\lambda)^2. \tag{9.589}$$

□

Remarque 9.280.
Nous avons alors immédiatement

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{9.590}$$

lorsque f est diagonalisable.

111. Théorème 9.204(5).

En suivant la notation (9.192), un endomorphisme est cyclique lorsqu'il existe $x \in E$ tel que $E_x = E$.

Proposition 9.281 ([254]).
Si f est un endomorphisme diagonalisable d'un espace vectoriel E de dimension n. Nous avons équivalence des points suivants.

(1) Le polynôme minimal est égal au polynôme caractéristique : $\mu_f = \chi_f$
(2) L'endomorphisme f est cyclique.
(3) $\mathcal{C}(f) = \mathbb{K}[f]$ [112].
(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 réside dans les équivalences (1)-(3). Les autres sont des résultats intermédiaires. En particulier, dans le cas diagonalisable, nous allons voir que le point (5) est essentiellement une reformulation de (1).

(i) **(4) implique (5)** Par la formule (9.582), les espaces propres de f ont tous une dimension de 1. Par conséquent f possède n valeurs propres distinctes.

(ii) **(5) implique (6)** Le théorème 9.204 nous dit que le polynôme minimal est scindé à racines simples. Puisque f possède n valeurs propres distinctes, μ est de degré n. Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 9.105 nous avons $\dim\big(\mathbb{K}[f]\big) = \deg(\mu) = n$ par la proposition 6.41.

(iii) **(6) implique (1)** Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 9.105 et la proposition 6.41 nous avons $n = \dim\big(\mathbb{K}[f]\big) = \deg(\mu)$. Comme χ est un polynôme annulateur (Caley-Hamilton 9.114), il est divisé par μ. Maintenant μ et χ sont des polynômes unitaires de degré n et μ divise χ. Ils sont donc égaux.

(iv) **(1) implique (2)** Le fait que f soit diagonalisable permet d'utiliser le théorème 9.204 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

$$f^k(v) = \sum_{i=1}^{n} \lambda_i^k e_i. \tag{9.591}$$

Pour prouver que cette famille (avec $k = 0, \dots, n-1$) est libre [113] nous en prenons une combinaison linéaire nulle et nous prouvons que les coefficients sont tous nuls. Soit donc

$$0 = \sum_{l=0}^{n-1} a_l f^l(v) = \sum_{l=0}^{n-1} a_l \sum_{i=1}^{n} \lambda_i^l e_i = \sum_{i=1}^{n} \Big(\sum_{l=0}^{n-1} a_l \lambda_i^l \Big) e_i. \tag{9.592}$$

Par hypothèse, la double somme est nulle, et nous avons pour tout i :

$$\sum_{l=0}^{n-1} a_l \lambda_i^l = 0. \tag{9.593}$$

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 qui est $a_j = 0$ pour tout j. Nous devons donc calculer le déterminant

$$\det \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{pmatrix}. \tag{9.594}$$

112. Rappel : $\mathcal{C}(f)$ est le commutant de f, définition 9.278.
113. Ce sera alors une base parce que n vecteurs libres dans un espace de dimension n est toujours une base, proposition 4.17.

Il s'agit du déterminant de Vandermonde déjà étudié par la proposition 9.11. Nous avons $\det(A) = \prod_{1\leqslant i<j\leqslant n}(\lambda_j - \lambda_i)$. Ce déterminant est bien non nul parce que toutes les valeurs propres sont distinctes.

(v) **(2) implique (3)** Soit v un vecteur cyclique[114] de f. Si g est un endomorphisme, nous définissons des coefficients $(a_k(g))_{k=0,\dots,n-1}$ par

$$g(v) = \sum_{k=0}^{n-1} a_k(g) f^k(v). \tag{9.595}$$

C'est une bonne définition parce que $\{f^k(v)\}$ est une base.

Nous définissons alors l'application

$$\begin{aligned} \psi\colon \mathcal{C}(f) &\to \mathbb{K}[f] \\ g &\mapsto \sum_{k=0}^{n-1} a_k(g) f^k. \end{aligned} \tag{9.596}$$

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, il est nul.

L'application ψ est surjective. En effet si un polynôme $P = \sum_{k=0}^{n-1} a_k f^k$ est donné, il suffit de poser

$$g(x) = \sum_k a_k f^k(x) \tag{9.597}$$

pour avoir $\psi(g) = P$.

(vi) **(3) implique (4)** Si $n_1, \dots, n_r$ sont les dimensions des différents espaces propres de f, nous avons les inégalités

$$\dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant n = n_1 + \dots + n_r \leqslant n_1^2 + \dots + n_r^2 = \dim\big(\mathcal{C}(f)\big). \tag{9.598}$$

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 9.282.
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 exemple 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. △

9.16.3 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 son polynôme minimal et μ_x le polynôme minimal ponctuel en x.

Lemme 9.283 ([255, 250, 113])**.**
Nous avons

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{9.599}$$

Démonstration. Si f est donnée, l'espace $\mathcal{C}(f)$ est l'espace des solutions de $fg = gf$. Supposons avoir choisi une base de E et notons A la matrice de f et X celle de g. L'équation est $AX - XA = 0$.

114. Définition 9.278.

(i) **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. Si 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

$$(AX - XA)_{kk} = \sum_i \big(A_{ki}X_{ik} - X_{ki}A_{ik}\big) = \sum_{k\leqslant i\leqslant k} (A_{ki}X_{ik} - X_{ki}A_{ik}) = 0. \tag{9.600}$$

Nous avons donc au maximum

$$\frac{n(n+1)}{2} - n \tag{9.601}$$

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.

(ii) **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 9.272. 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

$$\sum_{i=1}^r \deg(\mu_i) = \deg(\chi_f) = n. \tag{9.602}$$

Par conséquent $\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E)$.

□

Théorème 9.284 ([229]).
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 points suivants :

(1) $\mu = \chi$,

(2) f *est cyclique,*

(3) $\mathcal{C}(f) = \mathbb{K}[f]$.

Démonstration. Plusieurs implications. Notons que (2) implique (1) a déjà été démontré par le lemme 9.271.

(i) **(1) implique (2)** Conformément à ce que nous permet le lemme 9.97 nous choisissons[115] $a \in E$ de telle sorte à avoir $\mu_a = \mu$. De plus pour $x \in E$ nous considérons l'application

$$\begin{aligned}\varphi_x\colon \mathbb{K}[X] &\to E\\ P &\mapsto P(f)x.\end{aligned} \tag{9.603}$$

115. Dans toute la suite, nous devrions écrire μ_f et $\mu_{f,a}$ mais nous omettons d'indiquer explicitement la dépendance en f.

Nous avons $\varphi_a(P) = P(f)a$. Étant donné 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

$$[P] = \{P + Q\mu_a\}_{Q\in\mathbb{K}[X]}. \tag{9.604}$$

Nous considérons donc l'application quotient

$$\begin{aligned} \psi\colon \frac{\mathbb{K}[X]}{(\mu_a)} &\to E_a \\ [P] &\mapsto \varphi_a(P), \end{aligned} \tag{9.605}$$

et nous prouvons que c'est un isomorphisme d'espace vectoriel.

(i) **Linéaire** Parce que $(\lambda P + Q)(f) = (\lambda P)(f) + Q(f)$.

(ii) **Injectif** Si $\psi([P]) = 0$, alors $\varphi_a(P) = 0$, c'est à dire que $P \in \ker(\varphi_a)$. Mais, par définition 9.93 du polynôme minimal ponctuel, μ_a est générateur de $\ker(\varphi_a)$; donc il existe $Q \in \mathbb{K}[X]$ tel que $P = Q\mu_a$. Par conséquent $[P] = 0$, ce que nous voulions.

(iii) **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 $x = P(f)a = \varphi_a(P) = \psi([P])$.

Mais par hypothèse et par choix de a nous avons $\mu_a = \mu = \chi$, donc en fait $E_a = \mathbb{K}[X]/(\chi)$. Nous savons aussi que $\deg(\chi) = \dim(E)$ et que $\dim\big(\mathbb{K}[X]/P\big) = \deg(P)$ par la proposition 9.105. Au final nous avons $\dim(E_a) = \deg(\chi) = \dim(E)$. Et par conséquent $E_a = E$. Cela prouve que a est un vecteur cyclique pour f.

(ii) **(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 $g = P(f)$. Soient $y \in E$ et Q un polynôme tels que $y = Q(f)v$; en notant que g commute avec $P(f)$ nous avons

$$g(y) = g\big(Q(f)v\big) = Q(f)\big(g(v)\big) = Q(f)\big(P(f)v\big) = P(f)Q(f)v = P(f)y. \tag{9.606}$$

Donc $g = P(f)$.

(iii) **(3) implique (1)** Nous avons les inégalités :

$$n \leqslant \dim\big(\mathcal{C}(f)\big) = \dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant \deg(\chi) = n. \tag{9.607}$$

La première inégalité est le lemme 9.283. Ensuite, toutes les inégalités se trouvent être 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.

□

Corolaire 9.285 ([113]).
En suivant les notations sur les extensions de corps de base de la section 9.14, 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 9.284 qu'un endomorphisme est cyclique si et seulement si son polynôme minimal est égal à son polynôme caractéristique. Or par les propositions 9.265 et 9.266, nous savons que ces polynômes sont identiques pour f et pour $f_{\mathbb{L}}$. □

Théorème 9.286 (Similitude et extension de corps[113]).
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, si il existe $m \in \mathrm{GL}(E)$ tel que $f = mgm^{-1}$ alors il suffit d'appliquer le lemme 9.261 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 9.272. Il existe une unique suite de polynômes unitaires μ_i $(i = 1, \dots, s)$ tels que $\mu_{i+1} \mid \mu_i$ et pour laquelle nous avons une

décomposition $E = E_1 \oplus \ldots \oplus E_s$ pour laquelle $f|_{E_i} \colon E \to E_i$ est cyclique et de polynôme minimal μ_i.

Nous avons aussi $E_{\mathbb{L}} = (E_1)_{\mathbb{L}} \oplus \ldots \oplus (E_s)_{\mathbb{L}}$ et les $(E_i)_{\mathbb{L}}$ sont stables sous $f_{\mathbb{L}}$ qui y sera également cyclique (corolaire 9.285). 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 9.273 nous permet de conclure que f et g sont semblables si et seulement si $f_{\mathbb{L}}$ et $g_{\mathbb{L}}$ le sont. □

9.17 Hyperplans et formes linéaires

Définition 9.287.
Si E est un espace vectoriel de dimension n, un ***hyperplan*** *de E est un sous-espace vectoriel de dimension $n-1$.*

Proposition 9.288 ([256]).
À propos d'hyperplans et de formes linéaires sur un espace vectoriel E sur le corps $\mathbb{K}$.

(1) Si φ est une forme linéaire non nulle, alors $\ker(\varphi)$ est un hyperplan.

(2) Si H est un hyperplan de E, il existe une forme linéaire dont H est le noyau :

$$H = \ker(\varphi). \tag{9.608}$$

Démonstration. En deux parties.

(1) Soit un supplémentaire A de H. Nous considérons la restriction $\varphi_A \colon A \to \mathbb{K}$. Vu que les éléments non nuls de A sont hors de H, nous avons $\varphi(x) \neq 0$ dès que x est non nul dans A. Cela implique que φ_A est surjective.

D'autre part, φ_A est également injective : si $\varphi_A(x) = \varphi_A(y)$, alors $\varphi_A(x-y) = 0$, ce qui signifie que $x - y = 0$ ou encore que $x = y$.

Donc φ_A est un isomorphisme de $\mathbb{K}$-espaces vectoriels ; nous en déduisons par le corolaire 4.42 que A est de dimension 1 sur $\mathbb{K}$, parce que $\mathbb{K}$ est de dimension 1.

(2) Nous utilisons le théorème de la base incomplète 4.12(4) pour considérer une base $\{e_i\}_{i=1,\ldots,n}$ de E telle que $\mathrm{Span}\{e_1, \ldots, e_{n-1}\} = H$. Nous pouvons alors considérer la forme linéaire définie par

$$\varphi(e_i) = \begin{cases} 0 & \text{si } i = 1, \ldots, n-1 \\ 1 & \text{si } i = n. \end{cases} \tag{9.609}$$

Cette forme vérifie $\ker(\varphi) = H$.

□

Proposition 9.289 ([112]).
Soit un espace vectoriel E de dimension finie $n \geqslant 2$. Soit un sous-espace vectoriel V de E de dimension s. Alors V est une intersection de $n-s$ hyperplans de E.

Démonstration. Nous considérons une base de V que nous complétons[116] en une base de E : si $x = \sum_{i=1}^n x_i e_i$, nous avons $x \in V$ si et seulement si $x_{s+1} = \ldots = x_n = 0$. Nous considérons les formes linéaires

$$\begin{aligned} \varphi_i \colon E &\to \mathbb{R} \\ x &\mapsto x_i, \end{aligned} \tag{9.610}$$

et nous considérons les parties $H_i = \ker(\varphi_i)$ qui sont de hyperplans par la proposition 9.288. Les H_i avec $s+1 \leqslant i \leqslant n$ sont une famille de $n-s$ hyperplans qui vérifient

$$V = \bigcap_{i=s+1}^{n} \ker(\varphi_i) \tag{9.611}$$

116. Théorème de la base incomplète, 4.12(4).

parce que $x \in \ker(\varphi_i)$ si et seulement si $x_i = 0$.

Donc V peut être écrit comme intersection de $n - s$ hyperplans de E. □

Proposition 9.290 ([112]).
Soit un $\mathbb{K}$-espace vectoriel E de dimension finie $n \geqslant 2$. Si H_i sont des hyperplans de E, alors

$$\dim\Big(\bigcup_{i=1}^{m} H_i\Big) \geqslant n - m. \tag{9.612}$$

Démonstration. N'oubliez pas de prouver que $\bigcap_{i=1}^{m} H_i$ est un espace vectoriel. À part ça, nous faisons une petite récurrence.

(i) **Pour $m = 2$** Nous savons déjà par la proposition 4.45 que

$$\dim(H_1 \cap H_2) = \dim(H_1) + \dim(H_2) - \dim(H_1 \cap H_2). \tag{9.613}$$

De plus $\dim(H_1 + H_2) \leqslant n$. En remplaçant, par les valeurs,

$$\begin{aligned}
\dim(H_1 \cap H_2) &= \dim(H_1) + \dim(H_2) - \dim(H_1 \cap H_2) && (9.614a)\\
&= n - 1 + n - 1 - \dim(H_1 + H_2) && (9.614b)\\
&\geqslant 2n - 2 - n && (9.614c)\\
&= n - 2. && (9.614d)
\end{aligned}$$

Donc $\dim(H_1 \cap H_2) \geqslant n - 2$.

(ii) **La récurrence** Nous calculons $\dim(H_1 \cap \ldots \cap H_m \cap H_{m+1})$ en commençant encore par la proposition 4.45 :

$$\begin{aligned}
\dim(H_1 \cap \ldots \cap H_m \cap H_{m+1}) &= \underbrace{\dim(H_1 \cap \ldots \cap H_m)}_{\leqslant n-m} + \dim(H_{m+1}) && (9.615a)\\
&\quad - \underbrace{\dim\big((H_1 \cap \ldots H_m) + H_{m+1}\big)}_{\leqslant n} && (9.615b)\\
&\geqslant n - m + (n-1) - n && (9.615c)\\
&= n - m - 1. && (9.615d)
\end{aligned}$$

C'est bon pour la récurrence.

□

9.17.1 Trouver la matrice d'une symétrie donnée

Les notions de déterminants, produit scalaire et vectoriels[117] donnent une bonne intuition géométrique des matrices. Nous pouvons alors chercher les matrices de quelques symétriques dans $\mathbb{R}^2$ ou $\mathbb{R}^3$.

9.17.1.1 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,

117. Définitions 9.8, 9.155 et 11.24.

(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

$$A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}. \tag{9.616}$$

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 [118], 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

$$f_1=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad f_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}. \tag{9.617}$$

Nous avons

$$\begin{aligned}Af_1&=f_1\\Af_2&=f_2.\end{aligned} \tag{9.618}$$

Afin de trouver un vecteur perpendiculaire au plan, calculons le produit vectoriel :

$$f_3=f_1\times f_2=\begin{vmatrix}e_1&e_2&e_3\\0&1&0\\1&0&-1\end{vmatrix}=-e_1-e_3=\begin{pmatrix}-1\\0\\-1\end{pmatrix}. \tag{9.619}$$

Nous avons

$$Af_3=-f_3. \tag{9.620}$$

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 équations (9.617) dit que

$$f_1=e_2. \tag{9.621}$$

Ensuite, nous avons

$$\begin{aligned}f_2&=e_1-e_3\\f_3&=-e_1-e_3.\end{aligned} \tag{9.622}$$

La somme de ces deux équations donne $-2e_3=f_2+f_3$, c'est-à-dire

$$e_3=-\frac{f_2+f_3}{2} \tag{9.623}$$

Et enfin, nous avons

$$e_1=\frac{f_2-f_3}{2}. \tag{9.624}$$

118. Pour le trouver, penser au produit vectoriel.

Maintenant nous pouvons calculer les images de e_1, e_2 et e_3 en faisant

$$\begin{aligned} Ae_1 &= \frac{Af_2 - Af_3}{2} = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \\ Ae_2 &= Af_1 = f_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\ Ae_3 &= -\frac{f_2 - f_3}{2} = -\frac{1}{2}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{9.625}$$

La matrice A s'écrit maintenant en mettant les trois images trouvées en colonnes :

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{9.626}$$

9.17.1.2 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

$$f_2 = \begin{pmatrix} 1 \\ 0 \\ x \end{pmatrix}, \tag{9.627}$$

et puis fixer le x pour que le produit scalaire de f_2 avec f_1 soit nul. Si 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.

9.17.1.3 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.

9.18 Théorème de Burnside

Lemme 9.291.
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 9.292.
Toute représentation[119] *d'un groupe abélien d'exposant fini sur $\mathbb{C}^n$ a une image finie.*

Démonstration. 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 9.291 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 9.204). 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

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \tag{9.628}$$

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 9.293 (Burnside[101, 257])**.**
Un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$ *est fini si et seulement si il est d'exposant*[120] *fini.*

Démonstration. Soit G un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$. Si G est fini, l'ordre de ses éléments divise $|G|$ (corolaire 2.14 au théorème de Lagrange) et l'exposant est le PPCM qui est donc fini également. Le théorème est déjà démontré dans un sens.

Dans l'autre sens, nous notons $e < \infty$ l'exposant de G, et nous allons prouver que l'ensemble G est fini. Nous commençons par remarquer que tous les éléments de G sont des racines du polynôme $X^e - 1$, et ensuite nous nous lançons dans le travail.

(i) **Générateurs** Le groupe G est une partie de $\mathbb{M}(n, \mathbb{C})$ dont nous considérons l'algèbre engendrée[121] $\mathcal{G}$. Soit $C_1, \dots, C_r$ une famille génératrice de $\mathcal{G}$ constituée d'éléments de G et la fonction

$$\begin{aligned} \tau \colon G &\to \mathbb{C}^r \\ A &\mapsto \big(\mathrm{Tr}(AC_1), \dots, \mathrm{Tr}(AC_r)\big). \end{aligned} \tag{9.629}$$

(ii) **τ est injective** Soient $A, B \in G$ tels que $\tau(A) = \tau(B)$. Si C_i est un générateur de G, nous avons $\mathrm{Tr}(AC_i) = \mathrm{Tr}(BC_i)$ et par la linéarité de la trace, nous avons

$$\mathrm{Tr}(AM) = \mathrm{Tr}(BM) \tag{9.630}$$

pour tout $M \in G$. Notons par ailleurs

$$N = AB^{-1} - \mathbb{1}, \tag{9.631}$$

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

$$\begin{aligned} \mathrm{Tr}\left((AB^{-1})^p\right) &= \mathrm{Tr}\left(AB^{-1}(AB^{-1})^{p-1}\right) && (9.632a) \\ &= \mathrm{Tr}\left(BB^{-1}(AB^{-1})^{p-1}\right) \quad (9.630) && (9.632b) \\ &= \mathrm{Tr}\left((AB^{-1})^{p-1}\right). && (9.632c) \end{aligned}$$

119. Définition 4.126.
120. Définition 1.176.
121. Définition 1.293.

En continuant nous obtenons

$$\mathrm{Tr}\left((AB^{-1})^p\right) = \mathrm{Tr}(\mathbb{1}) = n. \tag{9.633}$$

D'autre part,

$$N^k = (AB^{-1} - \mathbb{1})^k = \sum_{p=0}^{k} \binom{p}{k} (-1)^{k-p} (AB^{-1})^p \tag{9.634}$$

En prenant la trace, et en tenant compte du fait que $\mathrm{Tr}\left((AB^{-1})^p\right) = n$,

$$\mathrm{Tr}(N^k) = \sum_{p=0}^{k} \binom{p}{k} (-1)^{k-p} n = n(1-1)^k = 0. \tag{9.635}$$

Donc la trace de N^k est nulle et le lemme 9.196 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.

(iii) **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 9.204 nous assure alors que ces éléments sont diagonalisables. Du coup les valeurs 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

$$\mathrm{Image}(\tau) = \{\mathrm{Tr}(A) \text{ tel que } A \in G\}^r, \tag{9.636}$$

qui est un ensemble fini. Par conséquent G est fini parce que τ est injective.

□

9.19 Ellipsoïde

Lemme 9.294.
Toute matrice peut être décomposée de façon unique en une partie symétrique et une partie antisymétrique. Cette décomposition est donnée par

$$S = \frac{M + M^t}{2}, \quad A = \frac{M - M^t}{2} \tag{9.637}$$

Démonstration. L'existence est une vérification immédiate de $S + A = M$ en utilisant (9.637). 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 9.295.
*Un **ellipsoïde** dans $\mathbb{R}^n$ centré en v est le lieu des points x vérifiant l'équation*

$$\langle x - v, M(x - v)\rangle = 1 \tag{9.638}$$

où M est une matrice symétrique strictement définie positive[122].

Lorsque nous parlons d'ellipsoïde plein, *il suffit de changer l'égalité en une inégalité.*

Remarque 9.296.
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

122. Définition 9.215.

positives, nous nommons S la partie symétrique de M et A la partie antisymétrique (lemme 9.294). Alors pour tout $x \in \mathbb{R}^n$ nous avons

$$x^t Ax = \langle x, Ax \rangle = \langle A^t x, x \rangle = -\langle Ax, x \rangle = -\langle x, Ax \rangle, \tag{9.639}$$

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

$$0 < x^t Mx = x^t Sx. \tag{9.640}$$

Proposition 9.297.
Si $\mathcal{E}$ est un ellipsoïde centrée à l'origine, il existe une base de $\mathbb{R}^n$ dans laquelle son équation est :

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} = 1. \tag{9.641}$$

Démonstration. Nous avons une matrice symétrique strictement définie positive S telle que l'équation soit $\langle x, Sx \rangle = 1$. Le théorème spectral 9.212 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

$$\sum_i \lambda_i x_i^2 = 1. \tag{9.642}$$

En posant $a_i = \frac{1}{\sqrt{\lambda_i}}$, nous trouvons le résultat. Cette définition des a_i est toujours possible parce que $\lambda_i > 0$. □

Corolaire 9.298.
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.

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

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} \leqslant 1. \tag{9.643}$$

Nous considérons la forme quadratique

$$\begin{aligned} q\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \sum_{i=1}^{n} \frac{\langle x, e_i \rangle^2}{a_i^2}. \end{aligned} \tag{9.644}$$

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 $q(x) > 0$. □

9.20 Système d'équations linéaires : méthode de Gauss

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 :

(2a) Remplacer une ligne par elle-même + un multiple d'une autre ;

$$\text{p.ex.} \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right) \overset{L_1 - 2.L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right)$$

(2b) Remplacer une ligne par un multiple d'elle-même ;

$$\text{p.ex.} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right) \overset{-L_1 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 2 & 4 \end{array}\right)$$

(2c) 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 9.299. — 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 a moins d'équations que d'inconnues, alors il y a une infinité de solutions qui dépendent d'un ou plusieurs paramètres :

$$\text{p.ex.} \left(\begin{array}{ccc|c} 1 & 0 & -2 & 2 \\ 0 & 1 & 3 & 0 \end{array}\right) \Leftrightarrow \begin{cases} x - 2z = 2 \\ y + 3z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 2 + 2\lambda \\ y = -3\lambda \\ z = \lambda \end{cases}$$