

Jean-Pierre Escofier

Toute l'algèbre de la Licence

5^e édition

DUNOD

Du même auteur :

Théorie de Galois, 2^e éd., 2004

Toute l'analyse de la Licence, 2020

Exercices d'analyse, 2015

Petite histoire des mathématiques, 2018

Illustration de couverture : rawf8 – shutterstock.com

© Dunod, 2002, 2006, 2011, 2016, 2020

11 rue Paul Bert 92240 Malakoff

www.dunod.com

ISBN 978-2-10-081248-6

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2^e et 3^e a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

TABLE DES MATIÈRES

Avant-propos

XI

PREMIÈRE ANNÉE

Chapitre 1 • Équations différentielles linéaires

1.1	Sommes et produits de fonctions	3
1.2	Équations différentielles linéaires sans second membre	6
1.3	Résolution des équations différentielles linéaires à coefficients constants	6
1.4	Combinaisons linéaires et espace engendré	9
1.5	Solutions des équations différentielles linéaires du second ordre à coefficients constants sans second membre	9
1.6	Résultats pour les équations différentielles linéaires du second ordre à coefficients constants avec second membre	10
1.7	Le logiciel <i>Sage</i>	12
	Exercices	13

Chapitre 2 • Suites récurrentes linéaires

2.1	Sommes et produits de suites	17
2.2	Suites satisfaisant une relation de récurrence linéaire	18
2.3	Suites satisfaisant $u_n + au_{n-1} + bu_{n-2} = 0$	19
2.4	Un peu d'histoire	21
2.5	Étude de la suite de Fibonacci	22
	Exercices	23

Chapitre 3 • L'espace vectoriel $\mathbb{R}^n$

3.1	Introduction de la géométrie à n dimensions	27
3.2	Famille d'éléments, suites finies, n -uplets	30
3.3	Définition de $\mathbb{R}^n$	31
3.4	Combinaisons linéaires et espace engendré	32
3.5	Base canonique de $\mathbb{R}^n$	34
3.6	Familles triangulaires et échelonnées	35
3.7	La droite vectorielle $\mathbb{R}$	37

Table des matières

3.8	Espaces engendrés dans $\mathbb{R}^2$	38
3.9	Espaces engendrés dans $\mathbb{R}^3$	40
3.10	Algorithme du pivot de Gauss dans $\mathbb{R}^n$	42
	Exercices	46

Chapitre 4 • Systèmes linéaires

4.1	Histoire ancienne	53
4.2	Leibniz, Cramer, Gauss	55
4.3	Systèmes linéaires	56
4.4	Exemples de résolution	56
4.5	Systèmes équivalents	58
4.6	Systèmes triangulaires et échelonnés	59
4.7	Méthode du pivot de Gauss	60
4.8	Exemples	64
4.9	Systèmes avec paramètres	66
4.10	Problèmes actuels	67
	Exercices	69

Chapitre 5 • Généralités sur les espaces vectoriels

5.1	Introduction	73
5.2	Un peu d'histoire	74
5.3	Structure de $\mathbb{R}$ -espace vectoriel	75
5.4	Exemples fondamentaux	77
5.5	Précisions sur les corps	78
5.6	Sous-espaces vectoriels	79
5.7	Exemples de sous-espaces vectoriels	80
5.8	Combinaisons linéaires et espace engendré	81
5.9	Somme de sous-espaces	83
	Exercices	84

Chapitre 6 • Bases et dimension

6.1	Introduction	89
6.2	Famille génératrice	89
6.3	Famille libre	90
6.4	Base d'un espace vectoriel	92
6.5	Dimension	94
6.6	Exemples de bases	96
6.7	Retour au rang	98
	Exercices	99

Chapitre 7 • Applications linéaires

7.1	Naissance du concept	109
7.2	Applications linéaires	110
7.3	Exemples	111
7.4	Propriété universelle	113
7.5	Noyau d'une application linéaire	115
7.6	Image d'une application linéaire	116
7.7	Le théorème du rang ou des dimensions	117
7.8	Résolution d'une équation linéaire	118
7.9	Résolution d'un système linéaire	119
7.10	Isomorphismes	121
	Exercices	123

Chapitre 8 • Matrices

8.1	Matrice d'une application linéaire	131
8.2	Matrices et applications linéaires	134
8.3	Un peu d'histoire	135
8.4	Matrices particulières	137
8.5	Exemples	139
8.6	Matrice de la composée	140
8.7	Propriétés du produit	143
8.8	Calcul de l'inverse d'une matrice	144
8.9	Changement de base	147
8.10	Rang et trace	152
	Exercices	153

Chapitre 9 • Sommes directes, produits, quotients

9.1	Exemples	163
9.2	Décomposition en somme directe	164
9.3	Sommes directes finies	165
9.4	Produit de deux espaces vectoriels	166
9.5	Projecteurs	169
9.6	Espaces vectoriels quotients	170
	Exercices	173

Chapitre 10 • Dualité

10.1	Introduction	177
10.2	Formes linéaires et hyperplans	178
10.3	Base duale	180
10.4	Orthogonal d'un sous-espace	181

Table des matières

10.5 Transposée d'une application linéaire	183
Exercices	185

DEUXIÈME ANNÉE

Chapitre 11 • Groupes

11.1 Introduction	193
11.2 Généralités	194
11.3 Exemples	196
11.4 Sous-groupes	197
11.5 Homomorphismes de groupes	199
11.6 Étude des groupes de permutation	201
11.7 Signature d'une permutation	204
11.8 Groupe linéaire	206
11.9 Centre du groupe linéaire	207
11.10 Générateurs du groupe linéaire	208
Exercices	209

Chapitre 12 • Arithmétique, anneaux

12.1 Introduction	215
12.2 Division euclidienne dans $\mathbb{Z}$	215
12.3 Congruence modulo n , définition de $\frac{\mathbb{Z}}{n\mathbb{Z}}$	216
12.4 Addition et multiplication dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$	218
12.5 Structures d'anneau commutatif unitaire et de corps	219
12.6 Homomorphismes d'anneaux	221
12.7 Utilisations des congruences	222
12.8 Éléments inversibles	223
12.9 Idéal	223
12.10 Sous-groupes, idéaux de $\mathbb{Z}$	224
12.11 Divisibilité, nombres premiers	225
12.12 Pgcd, ppcm, nombres premiers entre eux	226
12.13 Les corps $\frac{\mathbb{Z}}{p\mathbb{Z}}$	231
Exercices	232

Chapitre 13 • Polynômes

13.1 Introduction	245
13.2 Polynômes sur un corps K	246
13.3 Degré, division euclidienne	248

13.4	Pgcd de polynômes	250
13.5	Racines d'un polynôme	252
13.6	Dérivation	254
13.7	Éléments irréductibles	257
13.8	La structure de K -algèbre de $K[X]$	258
	Exercices	261

Chapitre 14 • Déterminants

14.1	Introduction historique	267
14.2	Calcul des déterminants : méthode de Bézout	272
14.3	Le caractère alterné	274
14.4	Multilinéarité	276
14.5	Formules et calculs	279
14.6	Déterminant d'un endomorphisme	282
14.7	Déterminant d'une matrice carrée	284
14.8	Retour sur le rang	286
14.9	Déterminant et volume	287
14.10	Déterminant et orientation	289
	Exercices	291

Chapitre 15 • Autour de la diagonalisation

15.1	Introduction	299
15.2	Étude du problème	300
15.3	Définitions	301
15.4	Exemple	302
15.5	Condition suffisante de diagonalisabilité	303
15.6	Condition nécessaire et suffisante de diagonalisabilité	304
15.7	Changement de corps de base	308
15.8	Seconde condition nécessaire et suffisante de diagonalisabilité	309
15.9	Triangularisation	311
15.10	Théorème de Hamilton-Cayley	313
15.11	Quelques applications	314
	Exercices	319

Chapitre 16 • Orthogonalité

16.1	Introduction	329
16.2	Orthogonalité dans le plan et l'espace ordinaires	329
16.3	Produit scalaire	332
16.4	Expression du produit scalaire	333
16.5	Norme et angle	336

Table des matières

16.6	Bases orthogonales et orthonormées	339
16.7	Orthogonalité de sous-espaces	342
16.8	Projection orthogonale	344
16.9	Transformations orthogonales	348
16.10	Groupe orthogonal de $\mathbb{R}^2$	351
16.11	Groupe orthogonal de $\mathbb{R}^3$	353
16.12	Endomorphisme adjoint et autoadjoint	356
16.13	Polynômes orthogonaux : exemple des polynômes de Legendre	359
	Exercices	367

Chapitre 17 • Carl Friedrich Gauss (1777-1855)	379

TROISIÈME ANNÉE

Chapitre 18 • Ouvertures sur les groupes

18.1	Relation d'équivalence sur un ensemble	400
18.2	Notion de sous-groupe distingué	403
18.3	Groupe quotient	406
18.4	Correspondance entre sous-groupes d'un groupe et sous-groupes d'un de ses quotients	409
18.5	Produits de groupes	411
18.6	Groupes monogènes et groupes cycliques	415
18.7	Action d'un groupe sur un ensemble	416
	Exercices	421

Chapitre 19 • Ouvertures sur les anneaux commutatifs unitaires

19.1	Sous-anneau, extension de corps	445
19.2	Caractéristique	448
19.3	Quotient d'un anneau par un idéal	449
19.4	Exemples de quotients	450
19.5	Correspondance entre idéaux d'un anneau et idéaux d'un de ses quotients	455
19.6	Produits d'anneaux	456
19.7	Opérations sur les idéaux	458
19.8	Théorème chinois	459
19.9	Éléments inversibles	463
19.10	Divisibilité dans les anneaux intègres	465
19.11	Idéaux premiers et maximaux	468
19.12	Anneaux euclidiens	471

19.13	Anneaux factoriels	472
19.14	Théorème de Fermat pour $n = 3$	476
19.15	Corps des fractions d'un anneau intègre	480
	Exercices	484

Chapitre 20 • Ouvertures sur les polynômes

20.1	La A -algèbre $A[X]$	499
20.2	Corps de rupture et de décomposition	502
20.3	Si A factoriel, alors $A[X]$ factoriel	505
20.4	Recherche des facteurs irréductibles d'un polynôme	507
20.5	Décomposition en éléments simples dans $\mathbb{C}(X)$ et $\mathbb{R}(X)$	508
20.6	Méthodes pour prouver l'irréductibilité d'un polynôme de $\mathbb{Z}[X]$, de $\mathbb{Q}[X]$	512
20.7	Localisation des racines d'un polynôme de $\mathbb{R}[X]$	514
20.8	Polynômes à plusieurs indéterminées	518
20.9	Polynômes symétriques	520
20.10	Fractions continues	525
20.11	Géométrie algébrique	535
	Exercices	537

Chapitre 21 • Corps finis

21.1	Corps finis : généralités	555
21.2	Existence et unicité des corps finis	558
21.3	Loi de réciprocité quadratique	561
21.4	Factorisation dans $\mathbb{Z}[\iota]$, théorème des deux carrés	564
21.5	Algorithme de Berlekamp	566
	Exercices	569

Chapitre 22 • Codes correcteurs et cryptographie

22.1	Les débuts des codes correcteurs	581
22.2	Codes correcteurs	583
22.3	Exemples	584
22.4	Distance de Hamming	585
22.5	Codes linéaires	586
22.6	Codes cycliques	589
22.7	Codes BCH	592
22.8	Histoire de la cryptographie	596
22.9	Logarithme discret	600
22.10	La méthode RSA	601

Table des matières

22.11	Grands nombres premiers	604
22.12	Factorisation	606
22.13	Courbes elliptiques	608
	Exercices	612

Chapitre 23 • Formes bilinéaires symétriques et quadratiques

23.1	Compléments sur le groupe orthogonal d'un espace euclidien	621
23.2	Formes bilinéaires et bilinéaires symétriques	628
23.3	Formes quadratiques	632
23.4	Méthode de Gauss pour la décomposition en carrés	634
23.5	Décomposition d'une forme quadratique sur $\mathbb{C}$ ou $\mathbb{R}$	637
23.6	Diagonalisation simultanée de deux formes quadratiques	639
23.7	Orthogonalité	641
23.8	Espaces quadratiques réguliers	642
23.9	Quaternions	646
23.10	Recherches arithmétiques de Lagrange	652
	Exercices	658

Chapitre 24 • Annexes pour le plaisir, avec beaucoup de présent

24.1	Manjul Bhargava (né en 1974)	680
24.2	Groupe des points d'une courbe elliptique	681
24.3	La conjecture <i>abc</i>	686
24.4	Les fonctions <i>L</i> de Dirichlet	689
24.5	Autour de l'hypothèse de Riemann (HR)	695

Bibliographie	705

Index	707

AVANT-PROPOS

L'enseignement des mathématiques et, plus généralement, des matières scientifiques, pose problème à nos sociétés en mutation. Alors que la recherche se développe partout dans le monde, autant fondamentale que pour des applications extraordinairement nombreuses et diversifiées, l'enseignement des bases des mathématiques est déstructuré et appauvri.

Qu'on étudie pour devenir chercheur ou enseignant de mathématiques ou pour se diriger plus tard vers d'autres domaines, l'étude des mathématiques a un sens qui s'est obscurci et qu'il faut sans doute redéfinir.

Ce livre a différents aspects profondément liés qui, je l'espère, contribueront à lutter contre ces dérives. La plus grande partie du livre est consacrée à la présentation des notions d'algèbre linéaire et d'algèbre de base, comme beaucoup d'autres livres actuels, en cherchant à me mettre à la portée des étudiants de *tous* niveaux. Je cherche à en montrer la beauté et l'efficacité et à donner plein de plaisirs à mes lectrices et lecteurs. Je donne des éclairages, mathématiques ou anecdotiques, de divers moments de leur construction au cours du temps. Je donne enfin des applications récentes.

On devrait pouvoir penser aux mathématiques comme on pense, je donne quelques exemples parmi mille, à des tableaux de Rembrandt ou de Nicolas de Stael, des films de Lang ou Mizogushi, des textes de Rimbaud ou Perec, des musiques de Mozart ou Stockhausen, etc. (remplacez ces noms par ceux de vos artistes préférés), et je serais heureux si ce livre pouvait y contribuer.

La première édition de ce livre, en 2002, correspondait à deux années d'études après le baccalauréat. La mise en place d'une harmonisation des études au niveau européen, m'a conduit à ajouté cinq nouveaux chapitres pour couvrir la troisième année de licence, en apportant les modifications et corrections nécessaires aux 17 premiers chapitres. Le choix des thèmes de ces cinq nouveaux chapitres n'a pas été évident, chaque université ayant ses sujets favoris ; j'en ai développé quelques applications et j'ai dû renoncer à bien des idées, faute de place.

Pour avoir commencé à apprendre les mathématiques dans des livres, je peux dire que leur lecture est insuffisante. Je vous invite donc à parcourir autant qu'à lire ce livre, à vous raconter cent fois ce qu'il contient, à en discuter avec d'autres, à le confronter aux cours et exercices qui vous seront proposés (à l'Université pour beaucoup d'entre vous), afin que les mathématiques et les histoires qu'il présente deviennent vôtres, que vous ayez quelques idées générales permettant de voir les choses de plus haut, que vos efforts de mémoire ne portent pas sur des détails.

Ce livre comporte une sorte de petit roman, au chapitre 17, pour raconter la vie d'un des plus grands scientifiques de tous les temps, Karl Friedrich Gauss (1777-1855). Gauss est à l'origine de bien des idées étudiées ici.

Avant-propos

Dans les précédentes éditions, de nouveaux exercices, un chapitre sur les codes correcteurs d'erreurs et la cryptographie ont été introduits. Dans cette cinquième édition, le changement essentiel est l'ajout d'un dernier chapitre donnant quelques aperçus sur de beaux résultats récents et de beaux problèmes non résolus. C'est sans doute une démarche inhabituelle pour un livre d'enseignement, mais elle me semble nécessaire. La recherche mathématique est toujours en plein développement. J'ai choisi des problèmes qui ne me paraissaient pas trop difficiles à appréhender. Je l'ai rédigé pour le plaisir de celles et ceux qui sont intéressés par les recherches actuelles.

J'espère que tout cela vous donnera à toutes et tous envie de poursuivre l'étude des mathématiques.

Mes remerciements vont aux éditions Dunod, toujours prêtes à vous écouter, en particulier à Laetitia Hérin et Vanessa Beunèche pour cette dernière édition, à Ghislaine Gueudet-Chartier, Michel Viallard, Françoise Guimier qui ont relu et critiqué des parties de ce texte, à Michel Saez et Julien Samson, à Annette Houdebine-Paugam qui a tout relu... et tout critiqué, à ceux et celles qui m'ont indiqué par courriel des corrections à apporter au texte, et à tous les Rennais et Rennaises qui m'ont apporté des idées un jour ou l'autre.

À D. C. A. et à N., G., M. et M.,

Jean-Pierre Escofier

Novembre 2019

Les figures de ce livre ont été tracées à l'aide du logiciel fig4Tex développé par Yvon Lafranche et Daniel Martin de l'Université de Rennes 1.

Première année

L'algèbre linéaire est présente dans beaucoup de domaines des mathématiques comme la géométrie, l'analyse, l'analyse numérique, les statistiques. Ramener un problème de mathématiques à un problème d'algèbre linéaire (on dit qu'on linéarise le problème) permet souvent de pouvoir conduire des calculs, d'obtenir des solutions approchées, etc.

L'introduction à l'algèbre linéaire est le but du cours d'algèbre de première année. Les quatre premiers chapitres introduisent à l'algèbre linéaire en étudiant des situations où elle intervient. Avec les chapitres 5 à 10, on entre dans la théorie des espaces vectoriels (de dimension finie) et des applications linéaires, ce qui nous confronte à des problèmes nouveaux, qu'on ne peut pressentir en étudiant les exemples des quatre premiers chapitres et pour lesquels un effort d'adaptation à l'abstraction est nécessaire.

ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

1

Le but des quatre premiers chapitres est de présenter des situations où l'algèbre linéaire est utile. Dans les chapitres suivants, on verra comment les notions d'algèbre linéaire permettent de les envisager dans un même cadre.

1.1 SOMMES ET PRODUITS DE FONCTIONS

Définition 1 : notion de fonction, définition originelle. Wilhelm Gottfried von Leibniz (1646-1716) est sans doute le premier à utiliser le mot *fonction*. Jean Bernouilli (1667-1748), qui le suit dans ce choix, précise en 1718 ce qu'il entend par là : « On appelle fonction d'une grandeur variable, une quantité composée de quelque manière que ce soit de cette grandeur variable et de constantes ».

Définition 2 : notion de fonction, définition actuelle. Notre définition actuelle de fonction est plus précise : pour définir une fonction f , on se donne :

- 1) un ensemble A dit ensemble de départ ou source de la fonction ;
- 2) un ensemble B dit ensemble d'arrivée ou but de la fonction ;
- 3) pour chaque élément x de A , un élément de B qu'on note $f(x)$ et qu'on appelle image de x par f .

On peut préciser comment on réalise le 3) : on se donne un sous-ensemble G de l'ensemble produit $A \times B$ qu'on appelle graphe de la fonction, tel que, pour tout x de A , il existe un unique élément y de B tel que $(x, y) \in G$.

Notation. La notation $f: A \rightarrow B$, apparue dans les années 1930, est celle que nous utiliserons pour désigner la fonction f de source A et de but B . Pour ne pas alourdir les notations, on utilisera aussi la notation $x \mapsto f(x)$, par exemple $x \mapsto -2x$, lorsque le contexte indique clairement la source et le but de la fonction.

Différences entre les deux définitions. Les différences entre les conceptions sous-jacentes à ces deux définitions sont multiples. Jean Bernoulli, comme tous les mathématiciens du XVIII^e siècle, ne pense, en fait, pour définir ses « quantités composées », qu'à des formules algébriques comme des quotients de deux polynômes, ou analytiques, comme des sommes infinies (on parle de séries). Ce sont des mathématiciens comme Leonhard Euler (1707-1783) qui écrivent qu'il faut étendre la notion de fonction à des correspondances quelconques, sans préciser ce que cela veut dire exactement : cela « dépend de notre bon plaisir » (Euler a alors en tête de donner les réponses les plus générales possibles à des problèmes de physique). Les mathématiciens du début du XIX^e siècle, comme Joseph Fourier (1768-1830) dans sa théorie de la chaleur publiée en 1822, expliciteront cette idée : une fonction est « une suite de valeurs données, assujetties ou non à une loi commune, et qui répondent à toutes les valeurs de x comprises entre » les extrémités d'un intervalle. On notait alors une fonction par $f(x)$, $\phi(x)$, où x représentait la *variable*.

Autour de 1900. Les mathématiciens de cette époque, à la suite de travaux de Vito Volterra (1860-1940), de Ivar Fredholm (1866-1927), de Maurice Fréchet (1878-1973) commencent à considérer les fonctions comme des objets mathématiques sur lesquels on peut calculer, plus précisément comme des éléments d'un ensemble muni d'une structure. Cela leur permet, par exemple, de définir une fonction, notons-la F , sur l'ensemble E des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ en associant à

$f \in E$, sa primitive valant 0 pour $x = 0$: $x \mapsto \int_0^x f(t) dt$. Comment noter cette nouvelle

fonction ? Si on a noté $f(x)$ comme Fourier, on devrait écrire $F(f(x))$, mais c'est ambigu, puisque $f(x)$ désigne aussi l'élément image de x par f ; si on considère la fonction comme un objet à part entière, c'est la notation f qui est adaptée et son image par la fonction F se note naturellement $F(f)$. On écrira, par exemple :

$F(f)(x) = \int_0^x f(t) dt$. La nouvelle notation traduit donc un changement de point de vue

des mathématiciens vis à vis des fonctions que nous allons développer dans ce livre.

Remarque : source et but. Soulignons qu'une fonction n'est pas seulement une formule ou un procédé, mais aussi la donnée de l'ensemble de départ et de l'ensemble d'arrivée de la fonction. Ainsi, la fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^2$ est-elle une fonction différente de la fonction $g: [0, 1] \rightarrow \mathbb{R}$ définie par $g(x) = x^2$ ou de la fonction $h: \mathbb{R} \rightarrow \mathbb{R}_+$ définie par $h(x) = x^2$.

Définition 3 : fonction vide. Ce qui est lié à l'ensemble vide est parfois utile pour éviter d'avoir à distinguer, comme certains livres le font, des cas particuliers. Pour tout ensemble B , l'ensemble $\emptyset \times B$ est encore l'ensemble vide ; on pose $G = \emptyset$. Comme G vérifie la condition 3 de la définition de fonction, G définit une fonction, notée $\emptyset : \emptyset \rightarrow B$ et appelée fonction vide.

La notation $\emptyset$ vient d'une lettre norvégienne et est due à André Weil (1906-1998).

Définition 4 : sommes et produits. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ sont deux fonctions, on peut définir leur somme et leur produit. Ce sont de nouvelles fonctions de $\mathbb{R}$ dans $\mathbb{R}$, notées $f + g$ et fg , qui sont définies comme associant à tout x de $\mathbb{R}$ les éléments $f(x) + g(x)$ et $f(x)g(x)$, autrement dit, pour tout x de $\mathbb{R}$:

$$(f + g)(x) = f(x) + g(x)$$

$$(fg)(x) = f(x)g(x)$$

Ces définitions se généralisent à des sommes finies et à des produits finis de fonctions à valeurs dans $\mathbb{R}$.

Tout ce qui précède se généralise également à des fonctions à valeurs dans le corps $\mathbb{C}$ des nombres complexes ou aux fonctions à valeurs dans un corps K quelconque (pour la définition générale de corps, voir 12.5, définition 2).

Si a est un élément de $\mathbb{R}$, notons $c_a : \mathbb{R} \rightarrow \mathbb{R}$ la fonction constante définie par : $c_a(x) = a$ pour tout x de $\mathbb{R}$. Très naturellement, la somme $f + c_a$ est notée $f + a$ et le produit $c_a f$ est noté af ; ce sont les fonctions définies, pour tout x de $\mathbb{R}$, par $(f + a)(x) = f(x) + a$ et $(af)(x) = ax$. La fonction af est appelée produit de la fonction f par le scalaire a .

Pour ne pas alourdir les notations, la fonction c_a sera donc notée a ; c'est le contexte qui permettra de savoir si a représente le nombre réel a ou la fonction constante $x \mapsto a$.

Si $a = -1$, on pose $af = -f$; pour tout x de $\mathbb{R}$, on a donc $(-f)(x) = -f(x)$. On note $f + (-g) = f - g$, donc on a, pour tout x de $\mathbb{R}$, $(f - g)(x) = f(x) - g(x)$. On a : $f + (-f) = 0$, 0 représentant ici la fonction c_0 , puisque, pour tout x de $\mathbb{R}$:

$$(f + (-f))(x) = f(x) - f(x) = 0.$$

De même, pour noter un produit de deux fonctions comme $x \mapsto -2xf(x)$, on écrira simplement $-2x f$, sans chercher à donner un nom à la fonction $x \mapsto -2x$ dont la source et le but sont supposés être ceux de f .

Enfin, notons que les mots fonctions et applications sont aujourd'hui synonymes ; l'emploi de l'un ou l'autre est une question d'usage ou de circonstances.

1.2 ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES SANS SECOND MEMBRE

En classe de Premières ou Terminales, on rencontre des problèmes où on recherche des fonctions $f: \mathbb{R} \rightarrow \mathbb{R}$ satisfaisant des relations liant f , ses fonctions dérivée première f' et dérivée seconde f'' (les notations des physiciens ou des mécaniciens sont diverses). Par exemple :

$$f' - 2f = 0 \quad (E_1)$$

ou

$$f'' - 3f' + 2f = 0 \quad (E_2)$$

La première équation est appelée *équation différentielle linéaire à coefficients constants du premier ordre*, la seconde est appelée *équation différentielle linéaire à coefficients constants du second ordre*. Les coefficients sont 1 et -2 dans le premier cas, 1, -3 et 2 dans le second.

On peut aussi rencontrer des équations du type :

$$f' - 2xf = 0$$

où les coefficients sont 1 pour f' et la fonction $x \mapsto -2x$ pour f ; l'équation est encore appelée équation différentielle linéaire du premier ordre, mais elle n'est plus à coefficients constants.

Ces équations sont dites sans second membre ou homogènes pour indiquer que le second membre est nul.

1.3 RÉOLUTION DES ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES À COEFFICIENTS CONSTANTS

Un peu d'histoire. Les équations différentielles sont apparues en mathématiques dans la seconde moitié du XVII^e siècle, souvent en lien avec des problèmes de physique. En 1697, pour résoudre les équations différentielles linéaires du premier ordre, Jean Bernoulli propose la méthode de variation de la constante (voir un cours d'analyse). C'est alors un des mathématiciens les plus célèbres de l'Europe. C'est lui qui, quelques années plus tard, va conseiller Leonhard Euler, encore très jeune, dans ses premières lectures mathématiques. Le père d'Euler est un ancien condisciple de Jean Bernoulli. Il finit par accepter que son fils ne se tourne pas comme lui vers la théologie, mais vers les mathématiques. Leonhard ne tarde pas à publier ses premiers travaux. En 1726, quand le fils de Jean Bernoulli, Nicolas (1695-1726),

1.3 • Résolution des équations différentielles linéaires à coefficients constants

qui est membre de l'académie de Saint-Pétersbourg, meurt, c'est Euler qui est appelé à le remplacer. Euler devient rapidement le premier mathématicien de son temps : il s'intéresse à tous les sujets et innove dans tous les domaines ; son œuvre mathématique est la plus considérable jamais écrite : elle comporte une centaine de volumes.



Léonhard Euler (1707-1783)

Benjamin Holl, d'après A. Lorgna, BnF/Gallica

Les idées d'Euler. C'est en 1743 qu'Euler expose, dans un article écrit en latin, comment résoudre les équations différentielles linéaires à coefficients constants. Il prend tout de suite le cas général de l'équation d'ordre n sans second membre :

$$f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0f = 0$$

où les dérivées successives de f sont notées $f', f'', \dots, f^{(n)}$. Pour faciliter la compréhension, vous pouvez suivre sa méthode sur les équations E_1 et E_2 de 1.2.

Avant de montrer comment trouver des solutions, Euler relève quelques propriétés simples de ces équations qui sont des observations sur les aspects linéaires du problème. D'abord, il remarque que, si on connaît une solution f de l'équation différentielle, alors pour tout réel λ , la fonction λf est aussi solution de l'équation différentielle. En effet, la dérivée k -ième de λf est :

$$(\lambda f)^{(k)} = \lambda f^{(k)}.$$

Chapitre 1 • Équations différentielles linéaires

Puis Euler remarque que, si on connaît deux solutions f et g de l'équation différentielle, alors, pour tout λ et tout μ réels, λf et μg sont solutions de l'équation ainsi que $\lambda f + \mu g$. En effet, la dérivée k -ième de $\lambda f + \mu g$ est :

$$(\lambda f + \mu g)^{(k)} = \lambda f^{(k)} + \mu g^{(k)}.$$

Euler développe encore cette idée, en expliquant que, si f_1, f_2, f_3 , etc. sont des solutions de l'équation, alors $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 + \text{etc.}$ est encore une solution de l'équation pour tous réels $\lambda_1, \lambda_2, \lambda_3$, etc.

Il reste, bien sûr, à trouver des solutions de l'équation. Euler propose de chercher *a priori* des solutions de la forme $x \mapsto e^{rx}$, où r est un réel, dont les dérivées successives sont $x \mapsto re^{rx}$, $x \mapsto r^2 e^{rx}$, etc.

➤ Pour l'équation E_1 de 1.2, on doit avoir :

$$re^{rx} - 2e^{rx} = 0$$

pour tout x réel, donc $r - 2 = 0$ puisque e^{rx} est non nul. La fonction f définie par $f(x) = e^{2x}$ est solution de (E_1) et, par conséquent, toutes les fonctions de la forme λf avec λ réel sont solutions de (E_1) .

➤ Pour l'équation E_2 de 1.2, on obtient :

$$r^2 e^{rx} - 3re^{rx} + 2e^{rx} = 0$$

pour tout x réel, donc $r^2 - 3r + 2 = 0$. Comme cette équation du second degré a pour racines $r = 1$ et $r = 2$, les fonctions f_1 et f_2 définies par $f_1(x) = e^x$ et $f_2(x) = e^{2x}$ sont solutions de l'équation (E_2) et, par conséquent, toutes les fonctions de la forme $\lambda_1 f_1 + \lambda_2 f_2$ sont solutions de (E_2) . On retrouve les résultats vus en classes de Premières ou Terminales.

Équation caractéristique. Dans le cas général, Euler obtient une équation polynomiale de degré n en r que nous appellerons *équation caractéristique de l'équation différentielle* :

$$r^n + a_{n-1}r^{n-1} + \dots + a_0 = 0$$

Si cette équation caractéristique a n racines distinctes $r_1, \dots, r_n$ les fonctions $f_1, \dots, f_n$ définies par $f_1(x) = e^{r_1 x}, \dots, f_n(x) = e^{r_n x}$ sont solutions de l'équation et Euler en conclut que toute fonction de la forme $\lambda_1 f_1 + \dots + \lambda_n f_n$, avec $\lambda_1, \lambda_2, \lambda_3$, etc., réels, est solution.

Ensuite, Euler examine les cas où l'équation caractéristique a des racines doubles, triples, ..., ou des racines non réelles, ce que nous ne développerons pas. Cependant, Euler ne pose pas la question de savoir s'il a ainsi obtenu toutes les solutions de ses équations différentielles. C'est le cas, mais cette question ne sera résolue que plus tard.

1.4 COMBINAISONS LINÉAIRES ET ESPACE ENGENDRÉ

Définition : combinaison linéaire. Étant données deux fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$, nous appellerons toute fonction de la forme $\lambda f + \mu g$, avec λ et μ réels, combinaison linéaire à coefficients réels de f et g .

De même, étant données n fonctions $f_1 : \mathbb{R} \rightarrow \mathbb{R}, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$, nous appellerons toute fonction de la forme $\lambda_1 f_1 + \dots + \lambda_n f_n$ avec $\lambda_1, \dots, \lambda_n$ réels, combinaison linéaire à coefficients réels de $f_1, \dots, f_n$.

Notation. L'ensemble des combinaisons linéaires à coefficients réels de $f_1, \dots, f_n$ sera noté :

$$\text{Vect}(f_1, \dots, f_n).$$

et appelé espace engendré par les fonctions $f_1, \dots, f_n$.

Dans le cas où $n = 1$, $\text{Vect}(f)$ est simplement l'ensemble des fonctions de la forme λf où λ est un réel.

Dans son étude des équations différentielles linéaires à coefficients constants, Euler a su voir que, si $f_1, \dots, f_n$ sont des solutions de l'équation, alors tout élément de $\text{Vect}(f_1, \dots, f_n)$ est également solution de l'équation.

1.5 SOLUTIONS DES ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU SECOND ORDRE À COEFFICIENTS CONSTANTS SANS SECOND MEMBRE

La forme générale d'une équation différentielle linéaire du second ordre à coefficients constants sans second membre est

$$f'' + af' + bf = 0 \quad (E)$$

où a et b sont des réels. L'équation caractéristique de cette équation différentielle est :

$$r^2 + ar + b = 0$$

Notons $S(E)$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ solutions de E . On démontre les résultats suivants.

Proposition.

1) Si $a^2 - 4b > 0$, l'équation caractéristique admet deux racines réelles distinctes r_1 et r_2 et, si on définit f_1 et f_2 par $f_1(x) = e^{r_1 x}$, $f_2(x) = e^{r_2 x}$, on a : $S(E) = \text{Vect}(f_1, f_2)$.

2) Si $a^2 - 4b = 0$, l'équation caractéristique admet une racine double r et, si on définit f_1 et f_2 par $f_1(x) = e^{rx}$, $f_2(x) = xe^{rx}$, on a : $S(E) = \text{Vect}(f_1, f_2)$.

3) Si $a^2 - 4b < 0$, l'équation caractéristique admet deux racines imaginaires conjuguées $r_1 = s + it$ et $r_2 = s - it$, où s et t sont réels et, si on définit f_1 et f_2 par $f_1(x) = e^{sx} \cos tx$, $f_2(x) = e^{sx} \sin tx$, on a : $S(E) = \text{Vect}(f_1, f_2)$.

Démonstration. Une démonstration de cette proposition sera donnée au paragraphe 15.11.3. Dans chacun des trois cas, on montre par un calcul facile que tout élément de $\text{Vect}(f_1, f_2)$ vérifie bien l'équation différentielle E . La difficulté est de montrer qu'il n'y a pas d'autres fonctions solutions que les éléments de $\text{Vect}(f_1, f_2)$. $\square$

Commentaire. La réponse est finalement de la même forme dans les trois cas, ce sont les fonctions f_1 et f_2 qui changent.

On dit souvent que la solution générale de l'équation E est $\lambda_1 f_1 + \lambda_2 f_2$. Il faut comprendre, par cette expression, que toute fonction de cette forme est solution de E et que toute solution de E est de cette forme ; cela revient exactement à dire que l'ensemble des solutions de E est $\text{Vect}(f_1, f_2)$. Si on cherche une solution de E satisfaisant à des conditions particulières, comme cela arrive dans les problèmes de physique, par exemple, on écrit que $\lambda_1 f_1 + \lambda_2 f_2$ satisfait ces conditions, ce qui permet de trouver λ_1 et λ_2 .

Peut-on faire mieux ? La présentation de l'ensemble des solutions peut encore susciter une question. Ne pourrait-on pas avoir une présentation encore plus simple, de la forme $\text{Vect}(h)$? Autrement dit, toute solution de l'équation différentielle serait de la forme λh . Ce serait alors le cas de f_1 et de f_2 ; mais $f_1 = \lambda_1 h$ et $f_2 = \lambda_2 h$ impliqueraient, puisque λ_1 et λ_2 ne sont pas nuls, qu'il existe α non nul tel que $f_1 = \alpha f_2$. Dans chacun des trois cas, c'est impossible (voir exercice 1.2). Par conséquent, on ne peut avoir une description de $S(E)$ de la forme $\text{Vect}(h)$.

1.6 RÉSULTATS POUR LES ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES DU SECOND ORDRE À COEFFICIENTS CONSTANTS AVEC SECOND MEMBRE

Considérons l'équation différentielle :

$$f'' + af' + bf = g \quad (E')$$

où $g : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction donnée (par exemple : $g(x) = \sin x$, $g(x) = e^x$, etc.).

Cette équation est appelée équation différentielle linéaire du second ordre à coefficients constants avec second membre.

1.6 • Résultats pour les équations différentielles linéaires du second ordre

- Si $g = 0$ (rappelons que 0 fait référence ici à une fonction), nous pouvons appliquer les résultats du paragraphe 1.5.
- Si $g \neq 0$ (c'est-à-dire s'il existe x tel que $g(x) \neq 0$), la résolution de (E') comporte deux étapes tout à fait distinctes :

1) on résout l'équation différentielle sans second membre :

$$f'' + af' + bf = 0 \quad (E)$$

associée à l'équation E' et on obtient $S(E) = \text{Vect}(f_1, f_2)$ d'après la proposition du paragraphe 1.5 ;

2) on recherche une solution h (une seule suffit) de l'équation E' .

Proposition. *L'ensemble des solutions de E' est :*

$$S(E') = h + \text{Vect}(f_1, f_2) = h + S(E) = \{h + f \mid f \in S(E)\}$$

Commentaire. On peut aussi dire que la solution générale de l'équation avec second membre est la somme d'une solution particulière de l'équation avec second membre et de la solution générale de l'équation sans second membre.

Démonstration. 1) On a $h + S(E) \subset S(E')$ car si $f \in S(E)$, on a :

$$\begin{aligned} (h + f)'' + a(h + f)' + b(h + f) &= h'' + f'' + ah' + af' + bh + bf \\ &= (h'' + ah' + bh) + (f'' + af' + bf) = g, \end{aligned}$$

donc $h + f \in S(E')$.

2) Montrons maintenant que $S(E') \subset h + S(E)$. En effet, si $h_1 \in S(E')$, alors $h_1 - h$ est solution de E puisque :

$$\begin{aligned} (h_1 - h)'' + a(h_1 - h)' + b(h_1 - h) &= h_1'' - h'' + ah_1' - ah' + bh_1 - bh \\ &= (h_1'' + ah_1' + bh_1) - (h'' + ah' + bh) = g - g = 0. \end{aligned}$$

Par conséquent, $h_1 - h$ est un élément de $S(E)$, ce qui prouve que $h_1 \in h + S(E)$. □

Pour trouver une solution particulière de l'équation avec second membre, il faut un peu d'expérience. Donnons un exemple.

Si le second membre est de la forme $x \mapsto e^{\alpha x}$, on examine si α est racine de l'équation caractéristique.

- s'il ne l'est pas, on cherche une solution particulière de la forme $x \mapsto \lambda e^{\alpha x}$.
- si α est racine simple de l'équation caractéristique, on cherche une solution particulière de la forme $x \mapsto \lambda x e^{\alpha x}$,
- si α est racine double de l'équation caractéristique, on cherche une solution particulière de la forme $x \mapsto \lambda x^2 e^{\alpha x}$.

Principe de superposition

Quand la fonction g du second membre de (E') est une somme de deux fonctions : $g = g_1 + g_2$, on peut appliquer ce qu'on appelle le principe de superposition en cherchant des solutions particulières h_1, h_2 de (E') avec les seconds membres g_1, g_2 respectivement. Alors $h = h_1 + h_2$ est une solution particulière de (E') avec le second membre g . Le principe de superposition n'est qu'un vieux nom pour une simple technique de résolution des équations linéaires, nous en reparlerons en 7.8.

1.7. LE LOGICIEL **SAGE**

C'est en 2005 que le mathématicien William Stein (né en 1974) a la première idée du logiciel *Sage* (acronyme de *System for Arithmetic Geometry Experimentation*), le nom de la sauge en anglais. Il s'agit de créer un logiciel pour pouvoir effectuer toutes sortes de calculs et d'algorithmes des mathématiques. *Sage* est un logiciel libre, un logiciel que vous pouvez utiliser gratuitement, copier, modifier, améliorer, transmettre aux copines et copains, en signaler les bogues.... *Sage* s'appuie sur de nombreux logiciels existant. La communauté des utilisateurs de *Sage* l'enrichit sans cesse.

Sage utilise un langage de programmation sous licence libre : *Python*. Guido van Rossum (né en 1956) en a écrit les premières versions au début des années 1990 et il est toujours resté le principal responsable de ses développements. La NASA, Google et d'autres grands organismes disent utiliser ce langage.

Sage peut traiter aussi bien des problèmes d'analyse que d'algèbre ou de combinatoire. Il peut traiter des problèmes de calcul numérique ou des problèmes de calcul formel. Par exemple, si vous étudiez une équation du troisième degré, *Sage* peut vous donner l'expression des racines avec la formule de Cardan, la ou les racines réelles ou complexes avec l'approximation que vous souhaitez. Si vous faites de l'algèbre linéaire, *Sage* peut résoudre à peu près tous les problèmes que vous rencontrerez dans ce livre, par exemple résoudre les systèmes linéaires du chapitre 4, les problèmes de réduction des matrices du chapitre 15, les problèmes sur les corps finis du chapitre 21.

L'accès à *Sage* est très simple ; vous pouvez vous connecter sur le réseau à un serveur ou le télécharger sur votre ordinateur.

➤ Vers le chapitre 2

Le but de ce chapitre était de se familiariser avec la linéarité dans les résolutions d'équations différentielles linéaires. Le but du chapitre 2 est de développer la même idée pour les suites récurrentes linéaires.

Exercices

1.1 Fonctions paires et impaires

a) Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ des fonctions paires et a un réel. Préciser si les fonctions $-f, fg, f+1, f+g, af$ sont paires. Une fonction quelconque de $\text{Vect}(f, g)$ est-elle paire ?

b) Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ des fonctions impaires et a un réel. Préciser si les fonctions $-f, fg, f+1, f+g, af$ sont impaires. Une fonction quelconque de $\text{Vect}(f, g)$ est-elle impaire ?

1.2 Peut-on faire mieux ?

Montrer que les fonctions f_1 et f_2 ne sont pas proportionnelles dans les trois cas de la proposition du paragraphe 1.5.

1.3 Changement de fonction

On considère l'équation différentielle $(E) : f'' - 2f' + f = 0$ (E). On n'utilisera pas le résultat de la proposition de 1.5.

a) Déterminer l'ensemble des fonctions u telles que $x \mapsto u(x)e^x$ soit solution de (E).

b) En déduire l'ensemble des solutions de (E).

1.4 Résolutions d'équations

Déterminer l'ensemble des solutions des équations différentielles suivantes.

a) $f'' - f' - 2f = e^x$ (E_1)

b) $f'' + f' - 2f = e^x$ (E_2)

c) $f'' - 4f' + 4f = e^{2x}$ (E_3)

d) $f''' - 2f'' - f' + 2f = e^x$ (E_4)

e) $f'' - f' - 2f = e^x + e^{3x}$ (E_5)

f) $f' - af = b$ (E_6) où a et b sont des réels.

Solutions

1.1 a) La réponse est oui pour toutes les fonctions.

b) On voit que $-f$ est impaire, fg est paire, $f+g$, af et toutes les fonctions de $\text{Vect}(f, g)$ sont impaires.

Exemple de rédaction détaillée : si f et g sont des fonctions impaires, pour tout x réel, on a :

- $(fg)(-x) = f(-x)g(-x)$ par définition du produit fg ;
- donc $(fg)(-x) = (-f(x))(-g(x))$ car f et g sont impaires ;
- donc $(fg)(-x) = f(x)g(x)$;
- donc $(fg)(-x) = (fg)(x)$ par définition du produit fg ;
- donc fg est une fonction paire.

Étude de $f + 1$ si f est impaire : pour tout x réel, on a : $(f + 1)(-x) = f(-x) + 1 = -f(x) + 1$.

Par conséquent, $f + 1$ sera paire si, pour tout x réel, $f(x) + 1 = -f(x) + 1$, c'est-à-dire si $f = 0$, et $f + 1$ sera impaire si, pour tout x réel, $f(x) + 1 = -(-f(x) + 1) = f(x) - 1$, ce qui est impossible. Donc $f + 1$ n'est ni paire ni impaire pour $f \neq 0$.

1.2 Reprenons les 3 cas de la proposition du paragraphe 1.5.

a) On a $f_1(x) = e^{r_1 x}$ et $f_2(x) = e^{r_2 x}$. S'il existe α réel tel que $f_1 = \alpha f_2$ on a, pour tout x réel : $e^{(r_1 - r_2)x} = \alpha$, ce qui est impossible puisque $r_1 \neq r_2$.

b) On a $f_1(x) = e^{rx}$ et $f_2(x) = x e^{rx}$. S'il existe α réel tel que $f_1 = \alpha f_2$, on a, pour tout x réel : $1 = \alpha x$, ce qui est impossible, pour $x = 0$ par exemple.

c) On a $f_1(x) = e^{sx} \cos tx$ et $f_2(x) = e^{sx} \sin tx$. S'il existe α réel tel que $f_1 = \alpha f_2$, on a, pour tout x réel : $\cos tx = \alpha \sin tx$, ce qui est impossible (pour $x = 0$, par exemple).

1.3 a) En calculant les dérivées première et seconde de $x \mapsto u(x)e^x$, on voit que cette fonction est solution de (E) si et seulement si $u'' = 0$, donc si et seulement si u est de la forme $x \mapsto ax + b$.

b) Soit f une solution quelconque de (E) et définissons la fonction u par $x \mapsto e^{-x} f(x)$ (c'est possible car e^{-x} ne s'annule pas).

On a $f(x) = u(x)e^x$, donc f est de la forme $x \mapsto (ax + b)e^x$.

Par conséquent, l'ensemble des solutions de (E) est celui donné au cas 2 de la proposition du paragraphe 1.5.

1.4 La démarche est toujours la même : chercher la solution générale de l'équation sans second membre (E_i) et chercher une solution particulière de l'équation avec second membre, notée (E'_i) .

a) L'équation caractéristique de (E_1) admet deux racines distinctes : -1 et 2 , donc $S(E_1) = \text{Vect}(f_1, f_2)$ avec $f_1 : x \mapsto e^{-x}$ et $f_2 : x \mapsto e^{2x}$.

Le coefficient de l'exposant du second membre est 1 , il n'est pas racine de l'équation caractéristique. On cherche donc une solution particulière de la forme $x \mapsto \lambda e^x$.

On trouve $\lambda = -\frac{1}{2}$.

En posant $h_1 : x \mapsto -\frac{1}{2}e^x$, on trouve $S(E'_1) = h_1 + S(E_1)$.

Autrement dit, $S(E'_1)$ est l'ensemble des fonctions de la forme $x \mapsto -\frac{1}{2}e^x + \lambda_1 e^{-x} + \lambda_2 e^{2x}$ où λ_1 et λ_2 sont réels.

b) L'équation caractéristique de (E_2) admet deux racines distinctes : 1 et -2, donc $S(E_2) = \text{Vect}(f_3, f_4)$ avec $f_3 : x \mapsto e^x$ et $f_4 : x \mapsto e^{-2x}$.

Le coefficient de l'exposant du second membre est ici racine simple de l'équation caractéristique. On cherche donc une solution particulière de la forme $x \mapsto \lambda x e^x$.

On trouve $\lambda = -\frac{1}{3}$.

En posant $h_2 : x \mapsto \frac{1}{3}x e^x$, on trouve $S(E'_2) = h_2 + S(E_2)$. Autrement dit, $S(E'_2)$ est l'ensemble des fonctions de la forme $x \mapsto \frac{1}{3}x e^x + \lambda_1 e^x + \lambda_2 e^{-2x}$ où λ_1 et λ_2 sont réels.

c) L'équation caractéristique de (E_3) admet 2 comme racine double, donc $S(E_3) = \text{Vect}(f_2, f_5)$ avec $f_5 : x \mapsto x e^{2x}$.

Le coefficient de l'exposant du second membre est racine double de l'équation caractéristique. On cherche donc une solution particulière de la forme $x \mapsto \lambda x^2 e^{2x}$.

On trouve $\lambda = -\frac{1}{2}$.

En posant $h_3 : x \mapsto \frac{1}{2}x^2 e^{2x}$, on trouve $S(E'_3) = h_3 + S(E_3)$. Autrement dit, $S(E'_3)$ est l'ensemble des fonctions de la forme $x \mapsto \frac{1}{2}x^2 e^{2x} + \lambda_1 e^{2x} + \lambda_2 x e^{2x}$ où λ_1 et λ_2 sont réels.

d) L'équation caractéristique de (E_4) admet trois racines distinctes : -1, 1 et 2, donc $S(E_4) = \text{Vect}(f_1, f_2, f_3)$.

Le coefficient de l'exposant du second membre est 1, il est racine simple de l'équation caractéristique. En généralisant la méthode pour les équations du second ordre,

on cherche une solution particulière de la forme $x \mapsto \lambda x e^x$. On trouve $\lambda = -\frac{1}{2}$.

En posant $h_4 : x \mapsto -\frac{1}{2}x e^x$, on trouve $S(E'_4) = h_4 + S(E_4)$. Autrement dit, $S(E'_4)$ est l'ensemble des fonctions de la forme $x \mapsto -\frac{1}{2}x e^x + \lambda_1 e^{-x} + \lambda_2 e^x + \lambda_3 e^{2x}$ où λ_1 , λ_2 et λ_3 sont réels.

e) On applique le principe de superposition. Le second membre de E'_5 est une somme de deux fonctions et nous allons chercher une solution particulière de E'_5 comme somme de solutions particulières des équations $f'' - f' - 2f = e^x$ et $f'' - f' - 2f = e^{3x}$. Nous connaissons déjà une solution particulière de la première équation : h_1 . On peut chercher une solution particulière de la seconde de la forme $h_5 : x \mapsto \lambda e^{3x}$.

On trouve $\lambda = \frac{1}{4}$.

Donc $S(E'_5) = h_1 + h_5 + S(E_1)$. Autrement dit, $S(E'_5)$ est l'ensemble des fonctions de la forme $x \mapsto -\frac{1}{2}e^x + \frac{1}{4}e^{3x} + \lambda_1 e^{-x} + \lambda_2 e^{2x}$ où λ_1 et λ_2 sont réels.

f) L'ensemble des solutions de l'équation sans second membre est $S(E_6) = \text{Vect}(x \mapsto e^{ax})$. Si h est une solution de l'équation avec second membre, on a $S(E'_6) = h + S(E_6)$.

Si $a = 0$, l'équation avec second membre s'écrit $f' = b$ qui a pour solution particulière $x \mapsto bx$ et les solutions de E'_6 sont de la forme $x \mapsto bx + C$ où C est réel (on peut, bien sûr, raisonner directement dans ce cas).

Si $a \neq 0$, l'équation avec second membre s'écrit $f' - af = b$ qui a pour solution particulière la fonction constante $x \mapsto -\frac{b}{a}$ et les solutions de E'_6 sont de la forme $x \mapsto C e^{ax} - \frac{b}{a}$ où C est réel.

On peut appliquer dans ce dernier exercice la méthode des exercices précédents : l'équation caractéristique a pour racine a et le second membre est de la forme be^{0x} .

SUITES RÉCURRENTES LINÉAIRES

2

2.1 SOMMES ET PRODUITS DE SUITES

Notion de suite. La notion de suite est liée à l'idée d'énumération : une suite a un premier terme u_1 , un second u_2 , etc. La suite peut être finie ou infinie. Souvent, le premier terme de la suite est noté u_0 , le second u_1 , etc., son n -ième terme est alors le terme d'indice $n - 1$: u_{n-1} .

Il est important de savoir que les suites ne sont pas des objets mathématiques nouveaux ; ce sont simplement des fonctions. Par exemple, la suite infinie de nombres réels notée $u = (u_n)_{n \in \mathbb{N}}$ est une fonction $u : \mathbb{N} \rightarrow \mathbb{R}$. Avec la notation des fonctions, son premier terme devrait être noté $u(0)$, son second terme $u(1)$, etc., mais l'usage est de noter ces termes en utilisant des indices : $u_0, u_1, \dots, u_n$, etc., ce qui cache un peu la nature fonctionnelle des suites. Nous utiliserons cependant la notation $n \mapsto u_n$, par exemple : $n \mapsto 2^n + 3 \times (-1)^n - 3$. Le terme u_n est appelé terme général de la suite u .

Si la suite u est finie et possède p termes $u_1, \dots, u_p$, c'est en fait une fonction $u : \{1, \dots, p\} \rightarrow \mathbb{R}$. Dans la suite de ce chapitre, l'ensemble d'indices sera toujours $\mathbb{N}$.

Sommes et produits de suites. Puisque les suites sont des fonctions, on peut en faire la somme et le produit comme on l'a expliqué dans le chapitre 1.

La somme des suites $u = (u_n)_{n \in \mathbb{N}}$ et $v = (v_n)_{n \in \mathbb{N}}$ est donc la suite notée $u + v$ dont le terme d'indice n est $(u + v)(n) = u(n) + v(n)$, autrement dit :

$$(u + v)_n = u_n + v_n.$$

Chapitre 2 • Suites récurrentes linéaires

De même, le produit de ces deux suites est la suite notée uv dont le terme d'indice n est $(uv)(n) = u(n)v(n)$, autrement dit :

$$(uv)_n = u_nv_n.$$

Si a est un réel, la suite au est la suite définie par $(au)_n = au_n$, appelée produit de la suite u par le scalaire a .

Par exemple, si u et v sont définies par $u_n = 2n^2 - 3n + 4$ et $v_n = \frac{2n+1}{3}$, on a $(u+v)_4 = u_4 + v_4 = 27$, $(uv)_4 = u_4v_4 = 72$, $(5v)_4 = 5v_4 = 15$.

Une suite de la forme $\lambda_1 u + \lambda_2 v$ avec λ_1, λ_2 réels est appelée combinaison linéaire des deux suites u et v ; son terme général est donc $\lambda_1 u_n + \lambda_2 v_n$. L'ensemble des combinaisons linéaires de deux suites u et v se note $\text{Vect}(u, v)$. On peut généraliser cette notion à un nombre quelconque de suites.

Nous étudierons dans ce chapitre les suites définies par des relations de récurrence linéaires.

2.2 SUITES SATISFAISANT UNE RELATION DE RÉCURRENCE LINÉAIRE

On dit qu'une suite satisfait une relation de récurrence linéaire si le terme d'indice n est lié aux termes d'indice inférieur par une relation de la forme

$$u_n = a_1 u_{n-1} + \dots + a_k u_{n-k} \quad (R)$$

pour $n \geq k$, où k est un entier > 0 donné et $a_1, \dots, a_k$ des réels fixés.

La relation (R) peut aussi bien être écrite :

$$u_n - a_1 u_{n-1} - \dots - a_k u_{n-k} = 0 \quad (R)$$

ce qui correspond à notre écriture des équations différentielles linéaires du chapitre précédent. Une telle relation sera dite sans second membre. Une relation de récurrence linéaire avec second membre est une équation de la forme

$$u_n - a_1 u_{n-1} - \dots - a_k u_{n-k} = v_n \quad (R')$$

où v est une suite donnée.

L'étude des suites satisfaisant une relation de récurrence linéaire est tout à fait analogue à l'étude des équations différentielles vues au chapitre 1. Copions donc la démarche d'Euler. On peut d'abord remarquer que si s est une suite satisfaisant la relation R , alors, pour tout réel a , la suite as satisfait R . On remarque ensuite que si s et t sont deux suites satisfaisant R , alors, pour tous réels a et b , les suites as , bt et $as + bt$ satisfont R .

2.3 • Suites satisfaisant $u_n + au_{n-1} + bu_{n-2} = 0$

Après avoir fait ces remarques, on peut chercher des suites satisfaisant effectivement R ; on cherche d'abord des suites dont le terme général est de la forme r^n où r est un réel pour $n \geq 1$ et dont le premier terme est 1. Le nombre r doit satisfaire, pour $n \geq k$, l'équation :

$$r^n = a_1 r^{n-1} + \dots + a_k r^{n-k}$$

soit

$$r^k - a_1 r^{k-1} - \dots - a_k = 0$$

Cette équation polynomiale, appelée *équation caractéristique*, a des racines réelles ou complexes et on peut en déduire l'ensemble $S(R)$ des suites satisfaisant R .

Pour préciser une suite de cet ensemble, on se donne souvent les premiers termes de la suite ; ces conditions sont appelées conditions initiales.

2.3 SUITES SATISFAISANT $u_n + au_{n-1} + bu_{n-2} = 0$

Lorsque $k = 2$, la relation R est de la forme :

$$u_n + au_{n-1} + bu_{n-2} = 0 \quad (R)$$

avec a et b réels et l'équation caractéristique devient :

$$r^2 + ar + b = 0$$

(si la relation de récurrence se présente sous la forme $u_n = au_{n-1} + bu_{n-2}$, son équation caractéristique est $r^2 - ar - b = 0$).

Comme pour les équations différentielles linéaires du second ordre à coefficients constants, on distingue trois cas.

Proposition

- 1) Si $a^2 - 4b > 0$, l'équation caractéristique de R admet deux solutions réelles r_1 et r_2 distinctes et, si on définit les suites s et t par $s_n = r_1^n$, $t_n = r_2^n$, on a : $S(R) = \text{Vect}(s, t)$. Si $r_1 = 0$, on définit s par $s_n = 0$ pour tout n à l'exception de $s_0 = 1$.
- 2) Si $a^2 - 4b = 0$, l'équation caractéristique de R admet une racine double r . Si r est non nul, on définit s et t par $s_n = r^n$, $t_n = nr^n$ et on a $S(R) = \text{Vect}(s, t)$. Si r est nul, on définit les suites s et t par $s_n = t_n = 0$ pour tout n à l'exception de $s_0 = t_1 = 1$.
- 3) Si $a^2 - 4b < 0$, l'équation caractéristique de R admet deux solutions imaginaires conjuguées $\rho e^{\pm i\theta}$ où ρ est réel positif et θ est réel et, si on définit s et t par $s_n = \rho^n \cos n\theta$, $t_n = \rho^n \sin n\theta$, on a : $S(R) = \text{Vect}(s, t)$.

Chapitre 2 • Suites récurrentes linéaires

Commentaire. Comme au chapitre 1 pour les équations différentielles linéaires du second ordre à coefficients constants, la réponse a la même forme dans les trois cas, ce sont les suites s et t qui changent.

Démonstration. Faisons la démonstration dans le cas où l'équation caractéristique a deux racines réelles distinctes r_1 et r_2 .

Comme $r_1^2 + ar_1 + b = 0$, on a $r_1^n + ar_1^{n-1} + br_1^{n-2} = 0$, ce qui prouve que la suite s est dans $S(R)$. De même, la suite t est dans $S(R)$. On en déduit que toute combinaison linéaire des suites s et t est dans $S(R)$, autrement dit $\text{Vect}(s, t) \subset S(R)$.

Pour montrer $S(R) \subset \text{Vect}(s, t)$, il faut montrer que tout élément u de $S(R)$ est combinaison linéaire de s et t , autrement dit, montrer qu'il existe des réels λ_1 et λ_2 tels que :

$$\lambda_1 s + \lambda_2 t = u.$$

Cette égalité impose, pour les indices 0 et 1 :

$$\lambda_1 s_0 + \lambda_2 t_0 = u_0$$

$$\lambda_1 s_1 + \lambda_2 t_1 = u_1$$

c'est-à-dire

$$\lambda_1 + \lambda_2 = u_0$$

$$\lambda_1 r_1 + \lambda_2 r_2 = u_1,$$

ce qui donne, puisque r_1 et r_2 sont distincts, des valeurs de λ_1 et de λ_2 faciles à calculer.

Nous allons vérifier par récurrence que ces valeurs de λ_1 et de λ_2 conviennent pour tous les termes de la suite, c'est-à-dire que $u_n = \lambda_1 s_n + \lambda_2 t_n$ pour tout entier n . Nous venons de choisir λ_1 et λ_2 de façon que l'égalité soit vraie pour les indices 0 et 1. Supposons que :

$$\lambda_1 s_k + \lambda_2 t_k = u_k$$

soit vraie pour tout entier $k \leq n$ avec $n \geq 1$. Pour l'entier $n + 1$, on a :

$$\begin{aligned} u_{n+1} &= -au_n - bu_{n-1} \\ &= -a(\lambda_1 s_n + \lambda_2 t_n) - b(\lambda_1 s_{n-1} + \lambda_2 t_{n-1}) \\ &= -a(\lambda_1 r_1^n + \lambda_2 r_2^n) - b(\lambda_1 r_1^{n-1} + \lambda_2 r_2^{n-1}) \\ &= \lambda_1 r_1^{n-1}(-ar_1 - b) + \lambda_2 r_2^{n-1}(-ar_2 - b) \\ &= \lambda_1 r_1^{n+1} + \lambda_2 r_2^{n+1} \\ &= \lambda_1 s_{n+1} + \lambda_2 t_{n+1} \end{aligned}$$

On peut donc conclure que $u = \lambda_1 s + \lambda_2 t$, d'où $S(R) = \text{Vect}(s, t)$.

Nous laissons les deux autres cas aux lectrices et lecteurs. □

Relations avec second membre. Pour compléter l'analogie entre les équations différentielles linéaires à coefficients constants et les suites définies par une relation de récurrence linéaire, traitons le cas particulier des suites satisfaisant, pour $n \geq 2$, la relation de récurrence :

$$u_n + au_{n-1} + bu_{n-2} = v_n \quad (R')$$

où (v_n) est une suite donnée. Si on connaît une suite $w = (w_n)$ satisfaisant R' , on a $S(R') = w + S(R)$. Autrement dit, la solution générale de (R') est la somme d'une solution particulière de (R') et de la solution générale de (R) . La démonstration suit pas à pas celle donnée au paragraphe 1.6 pour les équations différentielles.

Quand la fonction v du second membre de (R') est une somme de deux suites : $v = v_1 + v_2$, on peut utiliser le principe de superposition présenté en 1.6 et chercher des solutions particulières w_1, w_2 de (R') avec les seconds membres v_1, v_2 respectivement. Alors $w = w_1 + w_2$ est une solution particulière de (R') avec le second membre v .

2.4 UN PEU D'HISTOIRE

Un cas particulier de suite satisfaisant une relation de la forme ci-dessus est très connu. C'est celui de la suite de Léonard de Pise (Leonardo Pisano, vers 1170-1250) que nous appelons *suite de Fibonacci*, définie par la relation de récurrence :

$$F_n = F_{n-1} + F_{n-2}$$

avec $F_0 = F_1 = 1$. Pour décrire cette suite, Léonard de Pise avait imaginé des couples de lapins : chaque couple produit régulièrement tous les mois un couple de petits lapins capables, deux mois après et les mois suivants, d'engendrer à leur tour des couples de petits lapins, etc. Le premier jour, Léonard imagine un couple de lapins. Un mois plus tard, un couple naît et il y aura en tout deux couples de lapins ; un mois plus tard, le couple d'origine produira un nouveau couple mais le jeune couple ne produira encore rien : il y aura donc trois couples de lapins, ... La relation de récurrence est bien celle que nous venons d'indiquer. Léonard de Pise arrête son calcul au bout d'un an. Si on veut conserver le modèle pour la suite infinie, il faut supposer les lapins immortels... et infatigables¹.

1. Notons qu'Acharya Hemachandra (1087-1172), qui vivait au Gudjerat, avait obtenu cette suite en comptant les cadences de longueur n dans la poésie sanscrite : elles s'obtiennent à partir des cadences de longueur $n - 1$ en ajoutant une syllabe courte ou des cadences de longueur $n - 2$ en ajoutant une syllabe longue.

À cette époque, il y a très peu de mathématiciens en Europe. La chance de Léonard est d'accompagner son père, représentant des marchands pisans, à Bougie (Bejaïa) en Algérie. Il peut ainsi entrer en contact avec les mathématiques du monde arabe qui sont alors extrêmement brillantes. De retour en Italie, il écrit plusieurs traités dans lesquels il expose les méthodes et les résultats qu'il vient d'apprendre ou de développer.

Contrairement à ce que j'ai affirmé, après d'autres, dans les précédentes éditions, le nom de Fibonacci n'a pas été inventé par le « mathématicien-escroc », Guillaume Libri (1803-1869) dans son grand traité d'histoire des mathématiques italiennes (1838-1841). On trouve mention de ce nom bien avant Libri. Je peux tout de même raconter que Guillaume Libri, fort de son prestige d'académicien français, avait été chargé de s'occuper des fonds des bibliothèques de France. Il en profita pour y voler durant plusieurs années un nombre considérable de livres anciens rares et de manuscrits qu'il vendit en Angleterre pour l'équivalent de 1 500 000 euros environ. Il dû s'enfuir en 1848. Une partie seulement des documents volés ont pu être restitués, certains très récemment.

2.5 ÉTUDE DE LA SUITE DE FIBONACCI

La suite de Fibonacci $F = (F_n)$ est définie, pour $n \geq 2$, par la relation de récurrence :

$$F_n = F_{n-1} + F_{n-2}$$

et par ses premiers termes $F_0 = 1$ et $F_1 = 1$.

On cherche d'abord l'ensemble E des suites u vérifiant la relation de récurrence :

$$u_n = u_{n-1} + u_{n-2}$$

L'équation caractéristique est :

$$r^2 - r - 1 = 0$$

ce qui donne $r_1 = \frac{1+\sqrt{5}}{2}$ et $r_2 = \frac{1-\sqrt{5}}{2}$ donc, si on définit les suites $s = (r_1^n)$ et $t = (r_2^n)$, on a :

$$E = \text{Vect}(s, t)$$

La suite $as + bt$, a et b réels, de E dont les deux premiers termes sont égaux à 1 vérifie :

$$a + b = 1$$

$$ar_1 + br_2 = 1$$

On trouve :

$$a = \frac{r_1}{\sqrt{5}}$$

$$b = -\frac{r_2}{\sqrt{5}}$$

d'où, pour tout $n \geq 0$:

$$F_n = \frac{1}{\sqrt{5}}(r_1^{n+1} - r_2^{n+1})$$

On peut s'étonner de cette expression pour une suite de nombres entiers, mais on peut vérifier avec la formule du binôme que les termes en $\sqrt{5}$ s'éliminent.

Le nombre $\frac{1+\sqrt{5}}{2}$ est appelé nombre d'or et souvent noté φ ; il est irrationnel et son développement décimal commence par 1,618. On trouve parfois des textes donnant à ce nombre une importance absurde dans l'architecture de monuments anciens ou dans la structure des tableaux. Certains croient le reconnaître en faisant suffisamment de tracés et en obtenant des nombres plus ou moins proches. Il s'agit là d'élucubrations relativement récentes, remontant aux années 1830-1930 (voir, par exemple, le livre de M. Neveux et H.E. Huntley : *Le nombre d'or*, Seuil). Parmi d'autres bêtises, l'égalité : $(1 + \varphi) \times 12 = 10\pi$; vérifiez-le !

➤ Vers le chapitre 3

Dans ces deux chapitres, nous avons vu que l'ensemble des solutions des équations différentielles linéaires ou des relations de récurrence linéaires sans second membre s'exprime comme l'ensemble des combinaisons linéaires de certaines solutions de ces problèmes. Nous allons retrouver le même type de résultats pour la résolution des systèmes linéaires, ce qui va nous conduire à présenter au chapitre 5 la notion d'espace vectoriel, notion qui permet d'unifier les points de vue. Mais d'abord, nous allons étudier les espaces vectoriels $\mathbb{R}^n$.

Exercices

2.1 Relation de récurrence linéaire avec second membre

Donner une démonstration du résultat sur les relations avec second membre $u_n = au_{n-1} + bu_{n-2} + v_n$ présenté à la fin du paragraphe 2.3.

2.2 Résolution de relations de récurrence linéaire

Déterminer l'ensemble des solutions des relations de récurrence suivantes (on pourra s'inspirer du chapitre 1 pour trouver des solutions particulières).

- a) $u_n = u_{n-1} + 2u_{n-2} + 3$ (R'_1)
- b) $u_n = -u_{n-1} + 2u_{n-2} + 3$ (R'_2)
- c) $u_n = 4u_{n-1} - 4u_{n-2} + 2^n$ (R'_3)
- d) $u_n = 2u_{n-1} + u_{n-2} - 2u_{n-3} + 1$ (R'_4)
- e) $u_n = u_{n-1} + 2u_{n-2} + 3 + 3^n$ (R'_5)
- f) $u_n = au_{n-1} + b$ (R'_6) avec a, b réels.

2.3 Déterminer l'ensemble des solutions de la relation de récurrence

$$u_n = u_{n-1} + u_{n-2} + v_n$$

dans les cas suivants :

- a) $v = a$ avec a réel ;
- b) $v_n = n$;
- c) $v_n = r^n$ (discuter suivant les valeurs de r).

2.4 Équation fonctionnelle

On cherche l'ensemble E des fonctions continues sur $\mathbb{R}$ vérifiant l'équation fonctionnelle suivante :

$$f(x+y) + f(x-y) = 2(f(x) + f(y)) \quad (R).$$

- a) Vérifier que, si f et g sont deux fonctions de E , leur somme est une fonction de E et que, si λ est un nombre réel, alors λf est une fonction de E .
- b) Soit $f \in E$. Déterminer $f(0)$, puis montrer que f est paire.
- c) Déterminer $f(nx)$ pour tout $n \in \mathbb{N}$ par récurrence sur n , puis $f((p/q)x)$ pour tout rationnel p/q .
- d) Conclure.

Solutions

2.1 Montrons d'abord que $w + S(R) \subset S(R')$. En effet, si $u \in S(R)$, on a :

$$\begin{aligned} (w+u)_n &= w_n + u_n = aw_{n-1} + bw_{n-2} + v_n + au_{n-1} + bu_{n-2} \\ &= a(w+u)_{n-1} + b(w+u)_{n-2} + v_n. \end{aligned}$$

Réciproquement, soit $w' \in S(R')$ et posons $u' = w' - w$. On a $u' \in S(R)$ car :

$$\begin{aligned} u'_n &= (w' - w)_n = w'_n - w_n = (aw'_{n-1} + bw'_{n-2} + v_n) - (aw_{n-1} + bw_{n-2} + v_n) \\ &= a(w' - w)_{n-1} + b(w' - w)_{n-2} = au'_{n-1} + bu'_{n-2}. \end{aligned}$$

Donc $S(R') \subset w + S(R)$ et la conclusion.

2.2 La démarche est toujours la même : chercher la solution générale de la relation sans second membre (R_i) et chercher une solution particulière de la relation avec second membre (R'_i). L'expérience montre que, si la suite v du second membre est de la forme $n \mapsto ar^n$ avec a réel, on cherchera une solution particulière de la forme $n \mapsto \lambda r^n$, $n \mapsto \lambda n \times r^n$, $n \mapsto \lambda n^2 \times r^n$, etc. suivant que r n'est pas racine de l'équation caractéristique, en est racine simple, en est racine double, etc. (remarquer que $r = 0$ correspond à une suite constante). Pour la suite du e), on peut appliquer le principe de superposition.

a) L'équation caractéristique de R_1 admet deux racines distinctes : -1 et 2 , donc $S(R_1) = \text{Vect}(s, t)$ avec $s_n = (-1)^n$ et $t_n = 2^n$.

On cherche une solution particulière sous la forme d'une suite constante $n \mapsto \lambda$ et on trouve $\lambda = -\frac{3}{2}$.

En posant $w_n = -\frac{3}{2}$, on a $S(R'_1) = w + S(R_1)$ et les suites solutions sont de la forme $n \mapsto \lambda_1(-1)^n + \lambda_2 2^n - \frac{3}{2}$ avec λ_1 et λ_2 réels.

Nous ne donnons que cette dernière forme pour les autres questions.

b) $n \mapsto \lambda_1 + \lambda_2(-2)^n + n$ avec λ_1 et λ_2 réels.

c) $n \mapsto \lambda_1 2^n + \lambda_2 n 2^n + n^2 2^{n-1}$ avec λ_1 et λ_2 réels.

d) $n \mapsto \lambda_1 + \lambda_2(-1)^n + \lambda_3 2^n - \frac{n}{2}$ avec λ_1, λ_2 et λ_3 réels.

e) $n \mapsto \lambda_1(-1)^n + \lambda_2 2^n - \frac{3}{2} + \frac{3^{n+2}}{4}$ avec λ_1 et λ_2 réels.

f) Si $a \neq 1$, on trouve les suites $n \mapsto Ca^n + \frac{b}{1-a}$ où C est réel.

Si $a = 1$, on trouve les suites $n \mapsto bn + C$.

2.3 La relation sans second membre a pour ensemble de solutions $\text{Vect}(s, t)$ avec

$$s_n = r_1^n, \quad t_n = r_2^n, \quad r_1 = \frac{1+\sqrt{5}}{2} \text{ et } r_2 = \frac{1-\sqrt{5}}{2} \text{ comme on l'a vu au paragraphe 2.5.}$$

Il reste à trouver une solution particulière de la relation avec second membre dans chacun des trois cas. On trouve :

a) $n \mapsto \lambda_1 r_1^n + \lambda_2 r_2^n - a.$

b) $n \mapsto \lambda_1 r_1^n + \lambda_2 r_2^n - n - 3.$

c) Si $r \neq r_1$ et $r \neq r_2$, on trouve :

$$n \mapsto \lambda_1 r_1^n + \lambda_2 r_2^n + \frac{1}{r^2 - r - 1} r^{n+2}.$$

Si $r = r_1$ ou si $r = r_2$, on trouve les suites de la forme :

$$n \mapsto \lambda_1 r_1^n + \lambda_2 r_2^n + \frac{r^2}{r+2} n r^n.$$

2.4 a) Si f et g vérifie la relation R , on a, pour tous x, y réels, d'après la définition de la somme de deux fonctions :

$$(f + g)(x + y) = f(x + y) + g(x + y),$$

$$(f + g)(x - y) = f(x - y) + g(x - y),$$

$$(f + g)(x) = f(x) + g(x),$$

$$(f + g)(y) = f(y) + g(y).$$

On vérifie facilement que $f + g$ satisfait la relation R , donc $f + g \in E$. On vérifie de même que $\lambda f \in E$. Ici encore, on peut donner des propriétés de l'espace des solutions sans en avoir déterminé aucune.

b) Pour x réel et $y = 0$, on a $2f(x) = 2(f(x) + f(0))$, d'où $f(0) = 0$.

Pour $x = 0$ et y réel, on a alors $f(y) + f(-y) = 2f(y)$, d'où la parité de f .

c) Des essais conduisent à montrer par récurrence que $f(nx) = n^2 f(x)$ pour tout $n \geq 1$, ce qui résulte de $f((n+1)x) + f((n-1)x) = 2(f(nx) + f(x))$ et de $f((n+1)x) = [2n^2 - (n-1)^2 + 2] f(x) = (n+1)^2 f(x)$.

Si $x' = (p/q)x$, avec p, q positifs, on a $qx' = px$, d'où $f(qx') = f(px)$; ce qui précède

d) donne $f((p/q)x) = (p/q)^2 f(x)$. Par parité, la même relation vaut si $p/q < 0$.

Posons $f(1) = K$. On a donc $f((p/q)x) = K (p/q)^2$ pour tout rationnel p/q . En approchant un réel x par une suite (p_n/q_n) de rationnels de limite x , la continuité de f montre que $f(x) = \lim_{n \rightarrow +\infty} f(p_n/q_n) = \lim_{n \rightarrow +\infty} K (p_n/q_n)^2 = Kx^2$. Par conséquent, f est nécessairement de la forme $x \mapsto Kx^2$. Il est enfin facile de vérifier que tout f de cette forme vérifie R . Par conséquent, E est l'ensemble des fonctions de la forme $x \mapsto Kx^2$.

L'ESPACE VECTORIEL $\mathbb{R}^n$

3

3.1 INTRODUCTION DE LA GÉOMÉTRIE À n DIMENSIONS

La géométrie euclidienne. La géométrie a été considérée comme la base des mathématiques depuis Euclide, 300 ans av. J.-C., jusqu'au XIX^e siècle. La géométrie euclidienne décrivait ce qui se passe dans un espace correspondant à celui qui nous entoure ; pour parler des courbes et de leurs propriétés, tout était fait géométriquement. Une conique était l'intersection d'un cône de révolution avec un plan, d'autres courbes étaient tracées à l'aide d'instruments plus ou moins compliqués ; les propriétés de ces courbes, la construction de leurs tangentes, etc., tout était traité avec des méthodes géométriques.

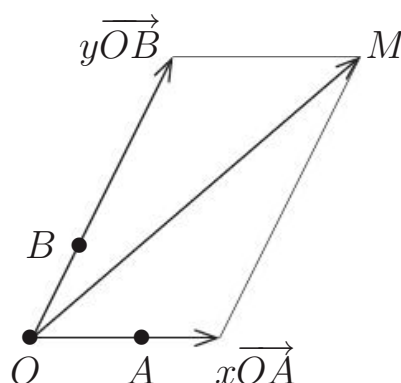
Descartes. René Descartes (1596-1650) a écrit le *Discours de la méthode* en 1637. On ne souligne peut-être pas assez le lien étroit entre la pensée philosophique de Descartes et ses travaux mathématiques. Il a annexé à son *Discours* un traité de *Géométrie* qui en constitue le principal exemple d'application. La *Géométrie* est le texte fondateur de nos notations actuelles et peut être lu aujourd'hui quasiment sans effort d'interprétation. C'est dans ce traité et dans des textes un peu antérieurs de Fermat qu'on trouve l'origine de la notion de repère et il est juste d'en avoir gardé la mémoire dans le mot *cartésien*. La notion de repère chez Descartes n'est pas entièrement dégagée et Descartes, qui refuse de parler de quantités négatives, s'arrange pour que ses x et ses y soient positifs.

Avec cette notion on peut :

- définir de nouvelles familles de courbes, en prenant des équations de toutes sortes, polynomiales ou non, liant l'abscisse x et l'ordonnée y d'un point dans un repère ;

- entreprendre l'étude des courbes planes avec des techniques nouvelles d'analyse ou d'algèbre.

Points du plan, de l'espace et leurs coordonnées. À partir de Descartes et Fermat, on peut donc parler des points, des droites, des plans de l'espace à l'aide de nombres. Par exemple, dans un plan, le choix de trois points O, A, B non sur une même droite permet de définir un repère affine. Si M est un point du plan, il existe des réels x et y uniques tels que $\overrightarrow{OM} = x\overrightarrow{OA} + y\overrightarrow{OB}$ (Descartes ne parle pas encore en terme de vecteurs) ; se donner le point M est exactement la même chose que se donner le couple (x, y) de nombres réels.



De même, dans l'espace, le choix de quatre points non coplanaires O, A, B, C permet d'identifier un point M de l'espace avec l'unique triplet (x, y, z) de nombres réels tels que $\overrightarrow{OM} = x\overrightarrow{OA} + y\overrightarrow{OB} + z\overrightarrow{OC}$.

Géométrie à n dimensions. Après l'introduction des repères cartésiens, l'identification des points d'une droite à des nombres réels, des points d'un plan à des couples de nombres réels, des points de l'espace à des triplets de nombres réels, n'a pas conduit immédiatement à généraliser les méthodes géométriques aux ensembles de n -uplets de nombres réels pour $n \geq 4$. Cette généralisation ne semble naturelle qu'*a posteriori*.

On remarquera également que le passage de la géométrie du plan et de l'espace à la géométrie dans l'ensemble des couples ou des triplets de nombres réels fait passer d'une géométrie où les données sont intuitives et non axiomatisées à une géométrie où les données sont basées sur la notion de nombre réel.

Dans un article publié dans le *Bulletin de la Société mathématique de France* (tome 3, 1875, pages 103-174), qui a été créée juste après la guerre de 1870 et joue toujours un rôle important dans la vie mathématique mondiale, Camille Jordan montre l'intérêt de cette généralisation. Avant de le lire, voici quelques mots sur ce grand mathématicien.



Camille Jordan (1838-1922)

Dieudonné J.- *Œuvres de Camille Jordan*, t1, Paris, Gauthier-Villars, 1961

Camille Jordan (1838-1922). Il est né dans une famille de scientifiques et d'hommes politiques. Le frère de sa mère est le peintre Puvis de Chavannes. Camille Jordan entre à l'École Polytechnique en 1855, avec 19,8 sur 20 de moyenne ; qui fera mieux ? Il en sort ingénieur et exerce cette profession pendant une quinzaine d'années durant lesquelles il trouve aussi le temps de poursuivre des recherches mathématiques très importantes. Nous parlerons au chapitre 11 de ses travaux fondateurs en théorie des groupes. En 1876, Camille Jordan devient professeur à l'École Polytechnique ; son cours d'analyse, plusieurs fois remanié, aura une grande influence.

Camille Jordan a eu deux filles et six fils. Trois d'entre eux sont tués durant la première guerre mondiale.

L'article de Jordan

Essai sur la géométrie à n dimensions

On sait que la fusion opérée par Descartes entre l'algèbre et la géométrie ne s'est pas montrée moins féconde pour l'une de ces sciences que pour l'autre. Car, si d'une part les géomètres ont appris, au contact de l'analyse, à donner à leurs recherches une généralité jusque-là inconnue, les analystes, de leur côté, ont trouvé un puissant secours dans les images de la géométrie, tant pour découvrir leurs théorèmes que pour les énoncer sous une forme simple et frappante.

Ce secours cesse lorsqu'on passe à la considération des fonctions de plus de trois variables ; aussi la théorie des fonctions est-elle relativement fort en retard. Le moment semble venu de combler cette lacune en généralisant les résultats déjà obtenus pour le cas de trois variables...

Bien que ces recherches soient purement algébriques, nous avons cru utile d'emprunter, ainsi que nos devanciers, quelques expressions à la géométrie. Ainsi, nous considérons un point comme défini dans l'espace à n dimensions, par les valeurs de n coordonnées $x_1, \dots, x_n$. Une équation linéaire entre ces coordonnées définira un plan...

Dans son article, Jordan va donc étudier comment généraliser à la géométrie à n dimensions les notions de distance, d'orthogonalité, d'angle et de rotation, en découvrant des théorèmes sur les nouveaux objets qu'il conçoit.

Affine et vectoriel. Pour bien comprendre ce qui suit, il est important de bien distinguer les notions de vectoriel et d'anneau.

Les droites, plans et espaces de la géométrie de l'enseignement secondaire sont des espaces de dimension 1, 2 ou 3 appelés affines et dont les éléments sont des points. Dans ces espaces, deux points a et b définissent un vecteur $\overline{ab}$. Les vecteurs $\overline{ab}$ et $\overline{cd}$ sont égaux si les points a, b, d et c forment un parallélogramme.

Le choix d'un repère d'origine O permet d'identifier ces points aux vecteurs d'origine O . Dans la théorie des espaces vectoriels, ce sont ces vecteurs que nous allons considérer et les points ne joueront aucun rôle.

Le rôle des dessins. Comme le dit Camille Jordan, nous allons conserver dans les espaces de dimension supérieure, « un puissant secours dans les images de la géométrie ». Chacun des lecteurs, chacune des lectrices adaptera cette maxime à sa propre personnalité ; un petit dessin aide souvent à mieux se représenter une situation, à voir les relations entre ses différents éléments, etc. Nous vous invitons fortement à faire cette démarche afin de faciliter votre compréhension de ce que nous allons exposer.

On constatera cependant que cette intuition géométrique n'est pas toujours possible ni utile et qu'il faut aussi se former aux caractères propres des objets et des démarches de l'algèbre linéaire.

3.2 FAMILLE D'ÉLÉMENTS, SUITES FINIES, n -UPLETS

Rappelons que, pour $n \geq 1$, une suite finie $x_1, \dots, x_n$ de n éléments d'un ensemble E définit un élément de l'ensemble produit E^n qu'on appelle n -uplet et qu'on note $(x_1, \dots, x_n)$. Si $n = 1$, un 1-uplet est simplement un élément de E , si $n = 2$, on parle de couple, si $n = 3$, on parle de triplet.

Définition : famille d'éléments. Étant donnés des ensembles E et I , on appelle famille d'éléments de E indexée par I une application $x : I \rightarrow E$; on la note $(x_i)_{i \in I}$. La famille est dite finie si I est fini. Dans ce cas, si I a n éléments, on pose très souvent $I = \{1, \dots, n\}$ et la famille s'identifie à la suite finie $x_1, \dots, x_n$ d'éléments de E ou au n -uplet $(x_1, \dots, x_n)$ qui est un élément de E^n .

Commentaire. Dans une famille, le même élément peut apparaître pour différentes valeurs des indices. Le fait de noter $1, \dots, n$ les éléments de I impose un ordre naturel sur les éléments de I qui est souvent utile, mais pas toujours nécessaire.

Le mot système est aussi utilisé, cela dépend un peu des habitudes de chacun. Il ne faut pas oublier la famille vide, ce qui permet d'éviter d'avoir à distinguer certains cas particuliers.

3.3 DÉFINITION DE $\mathbb{R}^n$

Définition : espace vectoriel $\mathbb{R}^n$. On appelle espace vectoriel $\mathbb{R}^n$ le triplet formé de :

- 1) l'ensemble des n -uplets $(x_1, \dots, x_n)$ de nombres réels ;
- 2) la loi d'addition de tels n -uplets :

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

- 3) la loi de multiplication de tels n -uplets par des nombres réels (souvent appelés scalaires) :

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

Les éléments de $\mathbb{R}^n$ considéré comme espace vectoriel sont appelés vecteurs. Le n -uplet $(0, \dots, 0)$ est noté 0 et appelé vecteur nul. Il ne faut pas le confondre avec le nombre réel 0 ; le contexte permet en général de lever cette ambiguïté. Certains livres conservent la notation $\vec{0}$ ou $0_{\mathbb{R}^n}$. La loi d'addition associe à deux vecteurs leur somme et la loi de multiplication est appelée produit par un scalaire.

Dans le cas $n = 1$, on parle de droite vectorielle ; dans le cas $n = 2$, on parle de plan vectoriel.

Notons que si $n = 0$, $\mathbb{R}^0$ ne possède qu'un élément : $\mathbb{R}^0 = \{0\}$.

Commentaire. On peut bien sûr définir d'autres opérations sur les n -uplets de nombres réels, par exemple le produit $(x_1, \dots, x_n)(y_1, \dots, y_n) = (x_1 y_1, \dots, x_n y_n)$, mais cela ne fait pas partie de la structure d'espace vectoriel de $\mathbb{R}^n$. Dans l'étude de l'espace vectoriel $\mathbb{R}^n$, nous ne retiendrons que la loi d'addition et la loi de multiplication par des scalaires.

Notations. Les flèches utilisées en géométrie pour noter les vecteurs ne seront pas utilisées pour noter les éléments des espaces vectoriels afin de ne pas alourdir les notations. Nous utiliserons souvent des pointillés, ce qui est parfois un peu abusif (par exemple, le n -uplet $(x_1, \dots, x_n)$ représente x_1 si $n = 1$, (x_1, x_2) si $n = 2$), mais suggère bien ce qui se passe.

3.4 COMBINAISONS LINÉAIRES ET ESPACE ENGENDRÉ

Nous allons reprendre les définitions des chapitres précédents dans le cadre de l'espace vectoriel $\mathbb{R}^n$.

Définition 1 : combinaison linéaire. Étant donnés des vecteurs $u_1, \dots, u_p$ de $\mathbb{R}^n$, nous appellerons combinaison linéaire à coefficients réels de $u_1, \dots, u_p$, tout vecteur u de $\mathbb{R}^n$ de la forme

$$u = \lambda_1 u_1 + \dots + \lambda_p u_p$$

avec $\lambda_1, \dots, \lambda_p$ réels. On dit que u dépend linéairement de $u_1, \dots, u_p$.

Si $p = 1$, une combinaison linéaire de u_1 est un vecteur de la forme λu_1 , si $p = 2$, une combinaison linéaire de u_1 et u_2 est un vecteur de la forme $\lambda u_1 + \mu u_2$ etc.

Définition 2 : espace engendré. L'ensemble des combinaisons linéaires à coefficients réels de $u_1, \dots, u_p$ est appelé espace engendré par $u_1, \dots, u_p$.

On le note $\text{Vect}(u_1, \dots, u_p)$. Cette notation pose un petit problème dans le cas d'un espace engendré par un seul vecteur $u = (x_1, \dots, x_n)$ de $\mathbb{R}^n$. On devrait noter $\text{Vect}(u) = \text{Vect}((x_1, \dots, x_n))$, avec des doubles parenthèses. On les omet souvent ; cet abus de notation n'est pas gênant si on sait bien de quoi on parle.

Premières propriétés. $\text{Vect}(u_1, \dots, u_p)$ ne dépend pas de l'ordre dans lequel sont donnés les vecteurs $u_1, \dots, u_p$.

Il est facile de vérifier que $\text{Vect}(u_1, \dots, u_p)$ est stable par somme, ce qui signifie que toute somme de deux vecteurs de $\text{Vect}(u_1, \dots, u_p)$ est dans $\text{Vect}(u_1, \dots, u_p)$.

De même, $\text{Vect}(u_1, \dots, u_p)$ est stable par produit par un réel, ce qui signifie que tout produit d'un vecteur de $\text{Vect}(u_1, \dots, u_p)$ par un réel est dans $\text{Vect}(u_1, \dots, u_p)$.

Par conséquent, $\text{Vect}(u_1, \dots, u_p)$ est stable par combinaison linéaire, ce qui signifie que toute combinaison linéaire de vecteurs de $\text{Vect}(u_1, \dots, u_p)$ est dans $\text{Vect}(u_1, \dots, u_p)$.

On en déduit que si $v_1, v_2 \in \text{Vect}(u_1, \dots, u_p)$, alors :

$$\text{Vect}(v_1, v_2) \subset \text{Vect}(u_1, \dots, u_p)$$

et, plus généralement, si $v_1, \dots, v_r \in \text{Vect}(u_1, \dots, u_p)$, alors :

$$\text{Vect}(v_1, \dots, v_r) \subset \text{Vect}(u_1, \dots, u_p).$$

Si un des u_i est nul, par exemple u_p , on a :

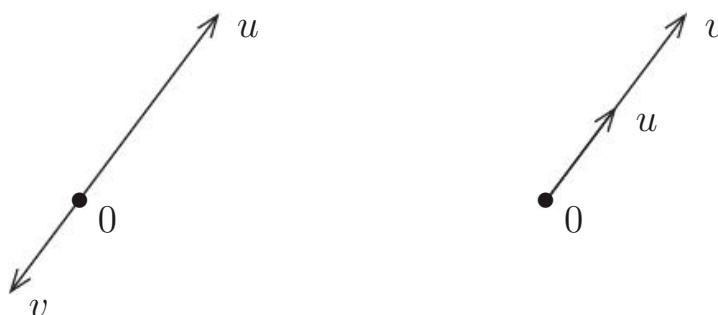
$$\text{Vect}(u_1, \dots, u_p) = \text{Vect}(u_1, \dots, u_{p-1}).$$

Si $u = 0$, $\text{Vect}(u) = \{0\}$, autrement dit $\text{Vect}(u)$ est réduit au vecteur nul. Notons, en passant, la distinction entre le vecteur nul, noté 0 , et l'ensemble de vecteurs réduit au seul vecteur nul, noté $\{0\}$.

L'espace engendré par un vecteur u non nul est une droite vectorielle

$$\text{Vect}(u) = \{\lambda u, \lambda \in \mathbb{R}\}.$$

Définition 3 : vecteurs colinéaires. On dit que des vecteurs u et v sont colinéaires si $u \in \text{Vect}(v)$ ou $v \in \text{Vect}(u)$, autrement dit s'il existe un réel λ tel que $u = \lambda v$ ou tel que $v = \lambda u$. Le vecteur 0 est colinéaire à tout vecteur. À part le cas particulier de $\mathbb{R}^0$, cela revient à dire que u et v appartiennent à une même droite vectorielle, ce qui correspond au sens du mot. On pourra penser aux dessins suivants qui représentent deux exemples de vecteurs u et v avec $v = \lambda u$:



Proposition 1 : critère de colinéarité. Deux vecteurs u et v sont colinéaires si et seulement s'il existe des réels λ et μ non tous deux nuls tels que :

$$\lambda u + \mu v = 0$$

Démonstration. Supposons que u et v sont colinéaires. Si $u \in \text{Vect}(v)$, il existe un réel α tel que $u = \alpha v$. En prenant $\lambda = 1$ et $\mu = -\alpha$, on a bien $\lambda u + \mu v = 0$. On raisonne de même si $v \in \text{Vect}(u)$.

Réciproquement, supposons la condition ci-dessus satisfaite. Si $\lambda \neq 0$, on a :

$$u = -\lambda^{-1}\mu v$$

ce qui prouve que $u \in \text{Vect}(v)$. On raisonne de même si $\mu \neq 0$. □

Proposition 2. Soient $p \geq 2$ un entier et $u_1, \dots, u_p$ des vecteurs de $\mathbb{R}^n$.

1) Soit $\lambda \neq 0$ un réel. On a alors :

$$\text{Vect}(u_1, u_2, \dots, u_p) = \text{Vect}(\lambda u_1, u_2, \dots, u_p).$$

2) Soit $u'_1 = u_1 + \sum_{2 \leq k \leq p} \lambda_k u_k$. On a alors :

$$\text{Vect}(u_1, u_2, \dots, u_p) = \text{Vect}(u'_1, u_2, \dots, u_p).$$

Commentaire. Cette proposition sera d'une grande utilité par la suite. On pourra se souvenir du 2) de la proposition sous une forme un peu plus générale (justifiée par le fait que l'espace engendré ne dépend pas de l'ordre des vecteurs) : on ne change pas l'espace engendré par un ensemble de vecteurs quand on change un des vecteurs en lui ajoutant une combinaison linéaire des autres vecteurs.

Démonstration. Pour montrer l'égalité de deux ensembles, il suffit de montrer que chacun d'eux est inclus dans l'autre.

1) Montrer la double inclusion est facile.

2) Un élément v de $\text{Vect}(u'_1, u_2, \dots, u_p)$ est une combinaison linéaire de $u'_1, u_2, \dots, u_p$. Comme u'_1 est une combinaison linéaire de $u_1, u_2, \dots, u_p$, il en est de même de v . Donc $\text{Vect}(u'_1, u_2, \dots, u_p) \subset \text{Vect}(u_1, u_2, \dots, u_p)$.

Un élément w de $\text{Vect}(u_1, u_2, \dots, u_p)$ est une combinaison linéaire de $u_1, u_2, \dots, u_p$. La définition de u'_1 permet d'exprimer u_1 comme combinaison linéaire de $u'_1, u_2, \dots, u_p$: $u_1 = u'_1 - \sum_{2 \leq k \leq p} \lambda_k u_k$.

Par conséquent, w est dans $\text{Vect}(u'_1, u_2, \dots, u_p)$.

Donc $\text{Vect}(u_1, u_2, \dots, u_p) \subset \text{Vect}(u'_1, u_2, \dots, u_p)$, d'où l'égalité des deux ensembles. $\square$

3.5 BASE CANONIQUE DE $\mathbb{R}^n$

Dans $\mathbb{R}^n$, pour $k = 1, \dots, n$, on note e_k le vecteur défini par le n -uplet dont toutes les composantes sont nulles sauf la k -ième qui est égale à 1.

Par exemple, dans $\mathbb{R}^4$, $e_1 = (1, 0, 0, 0)$, $e_2 = (0, 1, 0, 0)$, $e_3 = (0, 0, 1, 0)$ et $e_4 = (0, 0, 0, 1)$.

Définition. La famille de vecteurs $(e_1, \dots, e_n)$ est appelée base canonique de $\mathbb{R}^n$.

Notation. Pour tout n , on notera « can » la base canonique de $\mathbb{R}^n$, le contexte permettant de savoir quelle est la valeur de n sous-entendue.

Proposition 1 : première propriété de la base canonique. *Tout vecteur de $\mathbb{R}^n$ est une combinaison linéaire des e_k , $1 \leq k \leq n$, autrement dit : $\mathbb{R}^n = \text{Vect}(e_1, \dots, e_n)$.*

Démonstration. Soit $u = (x_1, \dots, x_n)$ un vecteur de $\mathbb{R}^n$; on peut écrire :

$$u = (x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n.$$

□

Les nombres $x_1, \dots, x_n$ sont appelés coordonnées de u dans la base $(e_1, \dots, e_n)$.

Proposition 2 : seconde propriété de la base canonique. *L'écriture d'un vecteur u comme combinaison linéaire de vecteurs de la base canonique de $\mathbb{R}^n$ est unique.*

Démonstration. Supposons que u s'écrive de deux façons comme combinaison linéaire de vecteurs de la base canonique, c'est-à-dire qu'il existe deux n -uplets $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ de nombres réels tels que

$$u = x_1 e_1 + \dots + x_n e_n = y_1 e_1 + \dots + y_n e_n$$

On a donc $u = x_1(1, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) + \dots = (x_1, \dots, x_n)$ et, de même, $u = (y_1, \dots, y_n)$. Les n -uplets $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ sont donc égaux, ce qui signifie que $y_k = x_k$ pour tout k , $1 \leq k \leq n$. □

Commentaire. Si on permute les vecteurs de la base canonique de $\mathbb{R}^n$, on obtient une famille de vecteurs possédant les deux mêmes propriétés que la base canonique. On appellera cette famille « base de $\mathbb{R}^n$ » en attendant une définition générale de la notion de base au chapitre 6. Ces bases vérifient les propositions 1 et 2 énoncées cidessus pour la base canonique.

3.6 FAMILLES TRIANGULAIRES ET ÉCHELONNÉES

Fixons un entier n et soit $(u_1, \dots, u_p)$ une famille de vecteurs de $\mathbb{R}^n$.

Pour noter les coordonnées de ces vecteurs par rapport à la base canonique $(e_1, \dots, e_n)$ de $\mathbb{R}^n$, nous allons faire usage de doubles indices. On écrira $u_1 = (x_{1,1}, \dots, x_{n,1})$ pour noter les n coordonnées du vecteur u_1 dans la base canonique, autrement dit : $u_1 = x_{1,1}e_1 + \dots + x_{n,1}e_n$. En général, $u_j = (x_{1,j}, \dots, x_{n,j})$. Nous supprimerons les virgules entre les deux indices quand cela ne créera pas de confusion.

Il est souvent commode de présenter ces coordonnées en les plaçant verticalement dans l'ordre des vecteurs et horizontalement dans l'ordre des vecteurs de base :

$$\begin{array}{ccccccc}
 & u_1 & u_2 & \dots & u_p & & \\
 e_1 & x_{1,1} & x_{1,2} & \dots & x_{1,p} & & \\
 e_2 & x_{2,1} & x_{2,2} & \dots & x_{2,p} & & \\
 \vdots & \vdots & \vdots & \dots & \vdots & & \\
 e_n & x_{n,1} & x_{n,2} & \dots & x_{n,p} & &
 \end{array}$$

Famille triangulaire. Nous dirons que la famille $(u_1, \dots, u_p)$ est triangulaire par rapport à la base canonique s'il existe un entier $r \leq p$ tel que :

- si $r < p$ on ait $u_j = 0$ pour $r < j \leq p$;
- $x_{ij} = 0$ pour tout couple (i, j) tel que $1 \leq i < j \leq r$;
- $x_{ii} \neq 0$ pour $1 \leq i \leq r$.

La famille est donc triangulaire si les coordonnées des vecteurs se présentent ainsi :

$$\begin{array}{ccccccc}
 u_1 & u_2 & \dots & u_r & u_{r+1} & \dots & u_p \\
 x_{1,1} & 0 & \dots & 0 & 0 & \dots & 0 \\
 x_{2,1} & x_{2,2} & \dots & 0 & \vdots & \dots & \vdots \\
 \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 x_{r,1} & x_{r,2} & \dots & x_{rr} & \vdots & \dots & \vdots \\
 \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 x_{n1} & x_{n2} & \dots & x_{nr} & 0 & \dots & 0
 \end{array}$$

où $x_{11}, \dots, x_{rr}$ sont non nuls. Les colonnes de 0 à droite peuvent ne pas exister.

Famille échelonnée. Cette forme est plus générale que la forme triangulaire. Pour tout vecteur u_k non nul de la famille $(u_1, \dots, u_p)$, notons i_k l'indice de sa première coordonnée non nulle, c'est-à-dire que $u_k = (0, \dots, 0, x_{i_k,k}, \dots, x_{n,k})$ avec $x_{i_k,k} \neq 0$.

Nous dirons que la famille est échelonnée par rapport à la base canonique s'il existe un entier $r \leq p$ tel que :

- si $r < p$, on ait $u_j = 0$ pour $r < j \leq p$;
- la suite $(i_1, \dots, i_r)$ soit strictement croissante.

La famille est donc échelonnée si les coordonnées des vecteurs se présentent avec une sorte d'escalier à marches de hauteurs variables marquant la séparation entre une zone composée uniquement de 0 et une zone où les parties des colonnes situées sous l'escalier commencent par des termes non nuls, comme dans l'exemple suivant d'une famille de 7 vecteurs de $\mathbb{R}^6$.

u_1	u_2	u_3	u_4	u_5	u_6	u_7
0	0	0	0	0	0	0
1	0	0	0	0	0	0
-1	3	0	0	0	0	0
2	-1	0	0	0	0	0
0	2	-1	0	0	0	0
1	0	1	5	0	0	0

Toute famille triangulaire par rapport à la base canonique est donc échelonnée par rapport à cette base.

Toute famille échelonnée peut être vue comme une famille triangulaire par rapport à une autre base. En effet, une famille échelonnée par rapport à la base canonique est triangulaire par rapport à la base définie par la famille $(e_{i_1}, \dots, e_{i_r})$, complétée par les vecteurs de la base canonique ne figurant pas dans cette liste. Par exemple, la famille de l'exemple précédent a une forme triangulaire par rapport à la base $(e_2, e_3, e_5, e_6, e_1, e_4)$, ce qu'on peut noter en rappelant dans une première colonne les vecteurs de la base choisie :

	u_1	u_2	u_3	u_4	u_5	u_6	u_7
e_2	1	0	0	0	0	0	0
e_3	-1	3	0	0	0	0	0
e_5	0	2	-1	0	0	0	0
e_6	1	0	1	5	0	0	0
e_1	0	0	0	0	0	0	0
e_4	2	-1	0	0	0	0	0

Notons enfin qu'une famille échelonnée de n vecteurs non nuls de $\mathbb{R}^n$ est triangulaire.

3.7 LA DROITE VECTORIELLE $\mathbb{R}$

L'espace vectoriel $\mathbb{R}^1 = \mathbb{R}$ est simple. Un élément de $\mathbb{R}$, autrement dit un vecteur de $\mathbb{R}$, est simplement un nombre réel ; la loi d'addition est l'addition de deux nombres réels et la loi de multiplication est la multiplication de deux nombres réels.

Dans cet espace vectoriel, la base canonique est réduite à $e_1 = 1$ et $\text{Vect}(e_1) = \mathbb{R}$.

On a $\text{Vect}(0) = \{0\}$ et, pour tout élément u non nul de $\mathbb{R}$, $\text{Vect}(u) = \mathbb{R}$. En effet, pour tout v de $\mathbb{R}$, il existe un réel λ tel que $v = \lambda u$ (prendre $\lambda = vu^{-1}$), donc $\text{Vect}(u) = \mathbb{R}$.

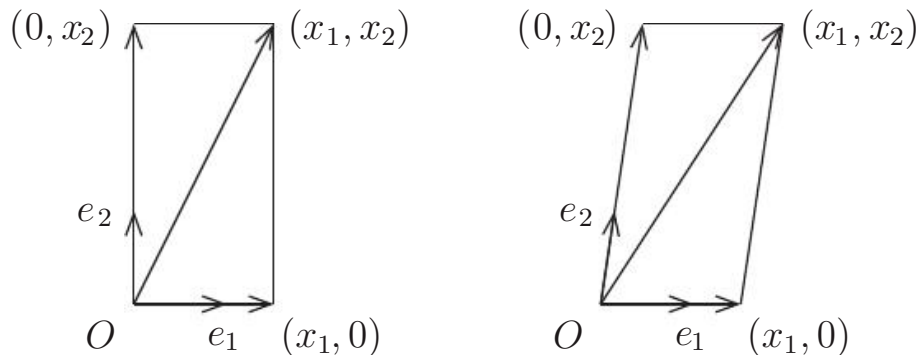
Par conséquent, deux vecteurs quelconques de $\mathbb{R}$ sont colinéaires. On pourra penser au dessin suivant.



3.8 ESPACES ENGENDRÉS DANS $\mathbb{R}^2$

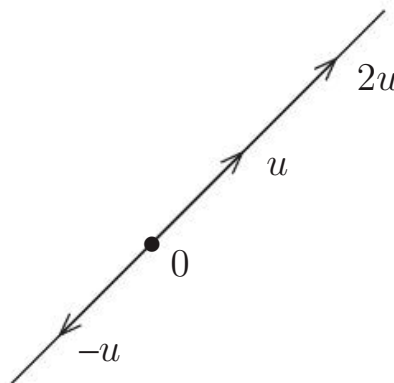
Le plan vectoriel $\mathbb{R}^2$ est l'ensemble des couples (x_1, x_2) de nombres réels avec la loi d'addition $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ et la loi de multiplication $\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$.

La représentation qu'on peut se faire de $\mathbb{R}^2$ (figure de droite) est celle d'une origine O correspondant au vecteur 0 et de vecteurs d'origine O du plan, en particulier les vecteurs $e_1 = (1, 0)$ et $e_2 = (0, 1)$ de la base canonique de $\mathbb{R}^2$, qu'on peut représenter éventuellement perpendiculaires (figure de gauche), même si cela ne correspond pas à quelque chose défini par la structure d'espace vectoriel.



Sous-espace engendré par un vecteur. Soit $u = (x_1, x_2)$ un vecteur de $\mathbb{R}^2$.

Si $u \neq 0$, $\text{Vect}(u) = \{\lambda(x_1, x_2), \lambda \in \mathbb{R}\}$ est une droite vectorielle. On peut penser à un dessin du genre suivant pour suggérer cet espace.



Sous-espace engendré par deux vecteurs. Soient $u = (x_1, x_2)$ et $v = (y_1, y_2)$ deux vecteurs de $\mathbb{R}^2$. Les cas suivants peuvent se produire :

1) u et v sont colinéaires.

a) $u = v = 0$, alors $\text{Vect}(u, v) = \{0\}$.

b) u et v ne sont pas tous deux nuls, alors $\text{Vect}(u, v)$ est une droite vectorielle de $\mathbb{R}^2$.

2) u et v ne sont pas colinéaires.

Proposition

1) Critère de colinéarité dans $\mathbb{R}^2$: les vecteurs u et v sont colinéaires dans $\mathbb{R}^2$ si et seulement si $x_1 y_2 - x_2 y_1 = 0$.

2) Si u et v ne sont pas colinéaires dans $\mathbb{R}^2$, alors $\text{Vect}(u, v) = \mathbb{R}^2$.

Démonstration. Nous détaillons cette démonstration qui présente sur un cas particulier un algorithme très important : l'algorithme du pivot de Gauss pour les vecteurs. Représentons les coordonnées des vecteurs u et v par rapport à la base canonique :

$$\begin{array}{cc} u & v \\ x_1 & y_1 \\ x_2 & y_2 \end{array}$$

Si $u = 0$, u et v sont colinéaires et $x_1 y_2 - x_2 y_1 = 0$.

Si u n'est pas nul, on a $x_1 \neq 0$ ou $x_2 \neq 0$.

a) Supposons d'abord $x_1 \neq 0$ et considérons le vecteur $v' = v - \frac{y_1}{x_1}u$. La proposition

3.4 permet d'affirmer que $\text{Vect}(u, v) = \text{Vect}(u, v')$. En posant $y'_2 = \frac{x_1 y_2 - x_2 y_1}{x_1}$, on

a $v' = (0, y'_2)$. Le système (u, v') est sous forme triangulaire par rapport à la base canonique :

$$\begin{array}{cc} u & v' \\ x_1 & 0 \\ x_2 & y'_2 \end{array}$$

Deux cas se présentent :

➤ si $y'_2 = 0$, c'est-à-dire si $x_1 y_2 - x_2 y_1 = 0$, on a $v' = 0$, donc $\text{Vect}(u, v) = \text{Vect}(u)$; l'espace engendré par u et v est la droite vectorielle engendrée par u et on a

$$u = \frac{y_1}{x_1} u ; u \text{ et } v \text{ sont colinéaires.}$$

➤ si $y'_2 \neq 0$, c'est-à-dire si $x_1 y_2 - x_2 y_1 \neq 0$, on a $v' \neq 0$ et u et v ne sont pas colinéaires. Pour montrer que $\text{Vect}(u, v) = \mathbb{R}^2$, il suffit de montrer que

$\text{Vect}(u, v') = \mathbb{R}^2$. Il suffit donc de montrer qu'un vecteur $w = (z_1, z_2)$ quelconque de $\mathbb{R}^2$ est une combinaison linéaire de u et de v' , c'est-à-dire de montrer l'existence de réels λ et μ tels que $\lambda u + \mu v' = w$. Cette équation est équivalente au système :

$$\begin{cases} \lambda x_1 & = z_1 \\ \lambda x_2 + \mu y'_2 & = z_2 \end{cases}.$$

Comme x_1 n'est pas nul, la première équation permet de déterminer λ ; comme y'_2 n'est pas nul, la seconde équation permet alors de déterminer μ .

Nous venons de revenir aux systèmes linéaires pour montrer que $w = (z_1, z_2)$ est une combinaison linéaire de u et de v' , mais il est possible de continuer dans l'esprit de l'algorithme. Pour cela, on écrit en colonnes les coordonnées de u , v' et w dans la base canonique :

$$\begin{array}{ccc} u & v' & w \\ x_1 & 0 & z_1 \\ x_2 & y'_2 & z_2 \end{array}$$

On fait alors apparaître des 0 à la place de z_1 puis z_2 , en calculant $w' = w - \frac{z_1}{x_1}u = (0, z'_2)$,

puis $w'' = w' - \frac{z'_2}{y'_2}v' = (0, 0)$:

$$\begin{array}{ccccc} u & v' & w & w' & w'' \\ x_1 & 0 & z_1 & 0 & 0 \\ x_2 & y'_2 & z_2 & z'_2 & 0 \end{array}$$

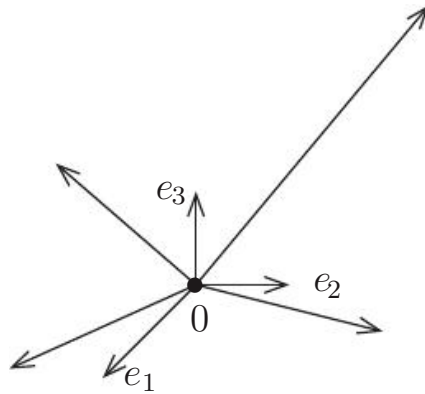
Comme $w'' = 0$, on a $w'' = \frac{z'_2}{y'_2}v'$, ce qui permet d'exprimer w comme combinaison linéaire de u et v .

b) Si $x_1 = 0$, on a $x_2 \neq 0$ et on a les mêmes résultats en permutant les rôles des deux vecteurs de la base canonique. $\square$

3.9 ESPACES ENGENDRÉS DANS $\mathbb{R}^3$

L'espace vectoriel $\mathbb{R}^3$ est l'ensemble des triplets (x_1, x_2, x_3) de nombres réels avec la loi d'addition : $(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$ et la multiplication par un scalaire : $\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$.

On peut se représenter cet espace en imaginant une origine O , représentant le vecteur nul, dans l'espace ordinaire et les vecteurs d'origine O , en particulier les vecteurs $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$ de la base canonique, qu'on peut éventuellement représenter comme perpendiculaires.



Nous allons donner un second exemple d'application de l'algorithme du pivot pour déterminer l'espace engendré par trois vecteurs de $\mathbb{R}^3$. La méthode pour déterminer l'espace engendré par un ensemble fini quelconque de vecteurs est analogue, nous la décrirons dans le cadre de $\mathbb{R}^n$ au paragraphe suivant.

Soient donc trois vecteurs non nuls de $\mathbb{R}^3$: $u = (x_1, x_2, x_3)$, $v = (y_1, y_2, y_3)$ et $w = (z_1, z_2, z_3)$.

L'idée est de construire un système sous forme échelonnée par rapport à la base canonique qui engendre le même espace que u, v, w .

1) Supposons $x_1 \neq 0$. Posons $v' = v - \frac{y_1}{x_1}u$, $w' = w - \frac{z_1}{x_1}u$. D'après la proposition 3.4

appliquée deux fois, on a $\text{Vect}(u, v, w) = \text{Vect}(u, v', w) = \text{Vect}(u, v', w')$. Les coordonnées de u, v', w' peuvent être représentées par :

$$\begin{array}{ccc} u & v' & w' \\ x_1 & 0 & 0 \\ x_2 & y'_2 & z'_2 \\ x_3 & y'_3 & z'_3 \end{array}$$

a) Supposons $y'_2 \neq 0$ et posons $w'' = w' - \frac{z'_2}{y'_2}v'$. D'après la proposition 3.4, on a

$\text{Vect}(u, v', w') = \text{Vect}(u, v', w'')$. Les coordonnées de u, v', w'' peuvent être représentées par :

$$\begin{array}{ccc} u & v' & w'' \\ x_1 & 0 & 0 \\ x_2 & y'_2 & 0 \\ x_3 & y'_3 & z''_3 \end{array}$$

On a alors deux cas :

- si $w'' = 0$, on a $\text{Vect}(u, v, w) = \text{Vect}(u, v')$; comme u et v' ne sont pas colinéaires, l'espace $\text{Vect}(u, v, w)$ est un plan vectoriel dans $\mathbb{R}^3$;
- si $w'' \neq 0$, la famille (u, v', w'') est triangulaire et on a $\text{Vect}(u, v, w) = \mathbb{R}^3$ par un raisonnement analogue à celui fait dans $\mathbb{R}^2$.

b) Supposons $y'_2 = 0$. Deux cas peuvent se produire.

- $z'_2 \neq 0$. On permute v et w et on se retrouve dans le cas précédent.
- $z'_2 = 0$. Les coordonnées de u, v', w' peuvent se représenter par :

$$\begin{array}{ccc} u & v' & w' \\ x_1 & 0 & 0 \\ x_2 & 0 & 0 \\ x_3 & y'_3 & z'_3 \end{array}$$

Si $y'_3 \neq 0$, on pose $w'' = w' - \frac{z'_3}{y'_3} v' = 0$, donc $\text{Vect}(u, v, w) = \text{Vect}(u, v')$ est un plan vectoriel de $\mathbb{R}^3$. Il en va de même si $z'_3 \neq 0$ en échangeant les rôles de v' et de w' .

Si $y'_3 = z'_3 = 0$, alors $\text{Vect}(u, v, w) = \text{Vect}(u)$ est une droite vectorielle de $\mathbb{R}^3$.

2) Supposons $x_1 = 0$. Plusieurs cas sont possibles.

- La première coordonnée de v est non nulle. On permute u et v et on se retrouve dans le cas précédent. Il en va de même si la première coordonnée de w est non nulle.
- Les premières coordonnées des trois vecteurs sont nulles. On reprend la même méthode que ci-dessus en distinguant les cas $x_2 \neq 0$, puis $x_2 = 0$, etc.

3.10 ALGORITHME DU PIVOT DE GAUSS DANS $\mathbb{R}^n$

On note $e_1, \dots, e_n$ les vecteurs de la base canonique de $\mathbb{R}^n$.

Soient $p \geq 1$ un entier et $u_1, \dots, u_p$ des vecteurs de $\mathbb{R}^n$. Notre but est de donner dans cette section un algorithme, autrement dit une méthode, pour déterminer un ensemble de vecteurs engendrant $\text{Vect}(u_1, \dots, u_p)$ et formant une famille triangulaire ou échelonnée. On vient de le faire dans les cas où $n = 2$, $p = 2$ ou $n = 3$, $p = 3$. Cela devrait permettre de comprendre la méthode générale sans en donner une description détaillée.

Notion de pivot. Dans cette méthode, on recherche à chaque étape une coordonnée non nulle. Le choix de la coordonnée la plus grande en valeur absolue est meilleur pour éviter des problèmes délicats d'erreurs de l'ordinateur qui calcule avec des valeurs approchées. Ces coordonnées s'appellent les pivots et la méthode générale est appelée algorithme, ou méthode, du pivot de Gauss. Pour la distinguer de la méthode du pivot utilisée pour résoudre des systèmes linéaires et présentée au chapitre 4, on précise parfois : *méthode du pivot de Gauss pour les vecteurs*.

L'algorithme du pivot pour les vecteurs. Si un des vecteurs $u_1, \dots, u_p$ est nul, on peut le placer à la fin de la famille $(u_1, \dots, u_p)$ sans changer l'espace engendré. Cette remarque vaudra à toutes les étapes de l'algorithme que nous allons décrire.

Disposons verticalement les coordonnées des vecteurs dans la base canonique en utilisant des doubles indices :

$$\begin{array}{cccc} & u_1 & u_2 & \dots & u_p \\ e_1 & x_{1,1} & x_{1,2} & \dots & x_{1,p} \\ e_2 & x_{2,1} & x_{2,2} & \dots & x_{2,p} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ e_n & x_{n,1} & x_{n,2} & \dots & x_{n,p} \end{array}$$

Soit toutes les premières coordonnées sont nulles et on est ramené à un problème sur les $n - 1$ dernières lignes, soit l'une des premières coordonnées est non nulle et, en permutant éventuellement deux vecteurs, on peut se ramener au cas où c'est la première coordonnée de u_1 qui est non nulle. Elle nous sert de premier pivot : on

pose $u'_j = u_j - \frac{x_{1,j}}{x_{1,1}} u_1$ pour $2 \leq j \leq p$, ce qui permet d'obtenir des 0 sur la première ligne des colonnes 2 à p . On peut alors représenter les vecteurs par :

$$\begin{array}{cccc} u_1 & u'_2 & \dots & u'_p \\ x_{1,1} & 0 & \dots & 0 \\ x_{2,1} & * & \dots & * \\ \vdots & \vdots & \dots & \vdots \\ x_{n,1} & * & \dots & * \end{array}$$

On reprend la même démarche sur les $p - 1$ dernières colonnes. Soit toutes les secondes coordonnées de $u'_2, \dots, u'_p$ sont nulles et on est ramené à un problème sur les $n - 2$ dernières lignes, soit l'une d'entre elles est non nulle et, en permutant éventuellement des vecteurs, on peut se ramener à ce que ce soit celle de u'_2 ; elle nous sert de second pivot. On continue ainsi jusqu'à l'obtention d'une famille échelonnée par rapport à la base canonique, donc une famille triangulaire par rapport à une base de $\mathbb{R}^n$, en permutant les vecteurs de la base canonique, comme on l'a vu en 3.6 (on n'écrit pas les vecteurs nuls) :

$$\begin{array}{cccccc}
 & u_1 & u'_2 & u''_3 & \dots & u_r^{(r-1)} \\
 x_{1,1} & 0 & 0 & \dots & 0 \\
 x_{2,1} & * & 0 & \dots & 0 \\
 \dots & \dots & \dots & \dots & \dots \\
 x_{r,1} & * & \dots & \dots & * \\
 \dots & \dots & \dots & \dots & \dots \\
 x_{n,1} & * & \dots & \dots & *
 \end{array}$$

Bilan des opérations. Pour transformer la famille de vecteurs initiale en une famille échelonnée, nous avons utilisé un certain nombre de fois les deux seuls types d'opérations suivantes :

- la permutation de deux colonnes entre elles ;
- l'addition d'un multiple d'une colonne à une autre.

Pour transformer la famille échelonnée en une famille triangulaire, nous avons permuté un certain nombre de fois deux lignes entre elles.

Conseil. Quand on utilise la méthode de Gauss, il faut soigneusement séparer les différents tableaux les uns des autres et noter tout aussi soigneusement les égalités définissant $u'_2, \dots, u'_p$ puis $u''_2, \dots, u''_p$, etc. Ces égalités peuvent être utiles, non seulement pour vérifier ses calculs, mais également pour certaines applications.

Concept de rang. Le nombre r est appelé rang de la famille de vecteurs $(u_1, \dots, u_p)$.

Il n'est pas évident, au stade où nous en sommes, que le nombre r de vecteurs non nuls obtenus soit le même si on ne prend pas les vecteurs dans le même ordre au départ. C'est cependant vrai. Nous justifierons cela dans le paragraphe 6.7. Le concept de rang est un des concepts importants de la théorie des espaces vectoriels de dimension finie ; il a mis beaucoup de temps à être dégagé. Nous en reparlerons à plusieurs reprises.

Cas où l'espace engendré est $\mathbb{R}^n$. Si $r = n$, un raisonnement, que nous avons déjà fait dans le cas de trois vecteurs, montre que $\text{Vect}(u_1, \dots, u_p) = \mathbb{R}^n$. Nous montrerons au chapitre 6 qu'il est ici nécessaire, pour que $\text{Vect}(u_1, \dots, u_p) = \mathbb{R}^n$, d'avoir $p \geq n$.

Appartenance à un « Vect ». Voici un premier exemple d'application de la méthode du pivot.

Étant donnés une famille de vecteurs $(u_1, \dots, u_p)$ de $\mathbb{R}^n$ et un vecteur v de $\mathbb{R}^n$, une question naturelle est de savoir si $v \in \text{Vect}(u_1, \dots, u_p)$.

On peut commencer par appliquer la méthode du pivot à la famille $(u_1, \dots, u_p)$ pour obtenir une famille triangulaire $(u_1, u'_2, u''_3, \dots, u_r^{(r-1)})$ qui engendre le même espace. Le problème est de savoir si $v \in \text{Vect}(u_1, u'_2, u''_3, \dots, u_r^{(r-1)})$, ce qui se fait

en plaçant v en colonne après le vecteur $u_r^{(r-1)}$ et en faisant apparaître des 0 dans les lignes successives de v . Comme la famille $(u_1, u'_2, u''_3, \dots, u_r^{(r-1)})$ est triangulaire, les calculs sont faciles. Si le résultat final est 0, c'est que v est dans $\text{Vect}(u_1, \dots, u_p)$. Dans le cas contraire, il ne s'y trouve pas. Ce calcul donne une expression de v en fonction des vecteurs $(u_1, u'_2, u''_3, \dots, u_r^{(r-1)})$ et on peut en déduire une expression de v en fonction de $(u_1, \dots, u_p)$ à l'aide des égalités ayant servi à définir les vecteurs aux différentes étapes des calculs.

Relation non triviale entre des vecteurs. Voici un second exemple d'application de la méthode du pivot.

L'algorithme du pivot appliqué à une famille $(u_1, \dots, u_p)$ de vecteurs de $\mathbb{R}^n$ fait parfois apparaître une colonne de 0. Si c'est la s -ième colonne qui devient nulle, cela signifie que u_s est une combinaison linéaire de $u_1, \dots, u_{s-1}$. On obtient cette relation en revenant aux définitions des différents vecteurs (voir l'exemple ci-dessous).

Donnons maintenant (enfin !) un exemple numérique pour bien voir fonctionner la méthode.

Exemple. Considérons la famille (u_1, u_2, u_3, u_4) des quatre vecteurs de $\mathbb{R}^5$ définie par :

u_1	u_2	u_3	u_4
0	0	0	0
1	-1	2	-1
2	-2	4	-2
-1	3	-2	5
3	-2	6	1

La méthode du pivot de Gauss conduit à calculer $u'_2 = u_2 + u_1$, $u'_3 = u_3 - 2u_1$, $u'_4 = u_4 + u_1$ pour obtenir :

u_1	u'_2	u'_3	u'_4
0	0	0	0
1	0	0	0
2	0	0	0
-1	2	0	4
3	1	0	4

L'égalité $u'_3 = 0$ donne, en revenant à la définition de u'_3 : $u_3 - 2u_1 = 0$ ou $u_3 = 2u_1$, relation non triviale entre les vecteurs de la famille (u_1, u_2, u_3, u_4) .

On permute alors u'_3 et u'_4 en posant $u''_4 = u'_3$ puis on calcule $u''_3 = u'_4 - 2u'_2$. Pour aboutir à :

u_1	u'_2	u''_3	u''_4
0	0	0	0
1	0	0	0
2	0	0	0
-1	2	0	0
3	1	2	0

La famille $(u_1, u'_2, u''_3, u''_4)$ a une forme échelonnée par rapport à la base canonique. Son rang est 3. Elle a une forme triangulaire par rapport à la base $(e_2, e_4, e_5, e_1, e_3)$:

	u_1	u'_2	u''_3	u''_4
e_2	1	0	0	0
e_4	-1	2	0	0
e_5	3	1	2	0
e_1	0	0	0	0
e_3	2	0	0	0

➤ Vers le chapitre 4

La résolution de systèmes linéaires à 2 ou 3 équations à 2 ou 3 inconnues vous est déjà connue. Nous allons les revoir, les préciser et les généraliser avec une méthode très proche de celle développée dans ce chapitre : l'algorithme du pivot de Gauss pour les systèmes linéaires.

Exercices

3.1 Soient $u = (2, 1)$ et $v = (-1, 1)$ deux vecteurs de $\mathbb{R}^2$.

a) Calculer $w = 2u + v$.

b) Les vecteurs u et v sont-ils colinéaires ?

c) Représenter u, v, w , Vect (u), Vect (v) et Vect (w) sur un dessin.

d) Comparer Vect (u) et Vect (v, w).

3.2 Stabilité de sous-ensembles de $\mathbb{R}^2$

Représenter graphiquement les ensembles suivants de $\mathbb{R}^2$. Étudier s'ils sont stables par somme, par produit par un scalaire et par combinaison linéaire. S'ils sont stables par combinaison linéaire, les écrire avec la notation Vect.

- $A_1 = \{(x, y) \mid x, y \in \mathbb{R}, 2x + 3y = 0\}$
- $A_2 = \{(x, y) \mid x, y \in \mathbb{R}, 2x + 3y = 1\}$
- $A_3 = \{(x, y) \mid x, y \in \mathbb{R}, x^2 + y^2 \leq 1\}$
- $A_4 = \{(a + b, a - b) \mid a, b \in \mathbb{R}\}$
- $A_5 = \{(x, y) \mid x, y \in \mathbb{R}, x \neq 0 \text{ et } y \neq 0\} \cup \{0\}$
- $A_6 = \{(x, y) \mid x, y \in \mathbb{R}, x \geq 0 \text{ et } y \geq 0\}$
- $A_7 = \{(x, y) \mid x, y \in \mathbb{Z}\}$

3.3 Dans $\mathbb{R}^3$, on considère les vecteurs $u = (-2, 3, 7)$, $v = (1, -2, -3)$, $w = (-1, -1, 6)$. Montrer que $\text{Vect}(u, v, w) = \text{Vect}(u, v) = \text{Vect}(v, w) = \text{Vect}(w, u)$.

3.4 Soient $u_1, \dots, u_p$ des vecteurs de $\mathbb{R}^n$. On suppose que u_1 est une combinaison linéaire de $u_2, \dots, u_p$. Que peut-on en déduire pour $\text{Vect}(u_1, \dots, u_p)$?

3.5 Détermination de rang et appartenance à un « Vect »

Dans les exercices suivants, si le rang de la famille est strictement inférieur au nombre de vecteurs de la famille, on déterminera une ou des relations non triviales entre les vecteurs de la famille.

a) Déterminer le rang dans $\mathbb{R}^2$ de la famille : $u_1 = (3, 1)$, $u_2 = (-1, 5)$.

Le vecteur $e_1 = (1, 0)$ appartient-il à $\text{Vect}(u_1, u_2)$? Si oui, l'exprimer comme combinaison linéaire de u_1 et u_2 .

b) Déterminer le rang dans $\mathbb{R}^3$ de la famille : $u_1 = (-1, 1, -3)$, $u_2 = (1, 2, 5)$, $u_3 = (1, 7, 1)$.

Le vecteur $e_1 = (1, 0, 0)$ appartient-il à $\text{Vect}(u_1, u_2, u_3)$? Si oui, l'exprimer comme combinaison linéaire de u_1, u_2, u_3 .

c) Déterminer le rang dans $\mathbb{R}^3$ de la famille : $u_1 = (1, 4, -3)$, $u_2 = (2, 5, 3)$, $u_3 = (3, 0, -3)$.

Le vecteur $e_1 = (1, 0, 0)$ appartient-il à $\text{Vect}(u_1, u_2, u_3)$? Si oui, l'exprimer comme combinaison linéaire de u_1, u_2, u_3 .

3.6 Comparaison de deux « Vect »

a) Dans $\mathbb{R}^3$, on pose $u_1 = (2, 3, -1)$, $u_2 = (1, -1, -2)$, $v_1 = (3, 7, 0)$, $v_2 = (5, 0, -7)$. Montrer que $\text{Vect}(u_1, u_2) = \text{Vect}(v_1, v_2)$.

b) Dans $\mathbb{R}^4$, on pose $u_1 = (1, 0, 1, 1)$, $u_2 = (-1, -2, 3, -1)$, $u_3 = (-5, -3, 1, -5)$, $v_1 = (-1, -1, 1, -1)$, $v_2 = (4, 1, 2, 4)$. Montrer que $\text{Vect}(u_1, u_2, u_3) = \text{Vect}(v_1, v_2)$.

Solutions

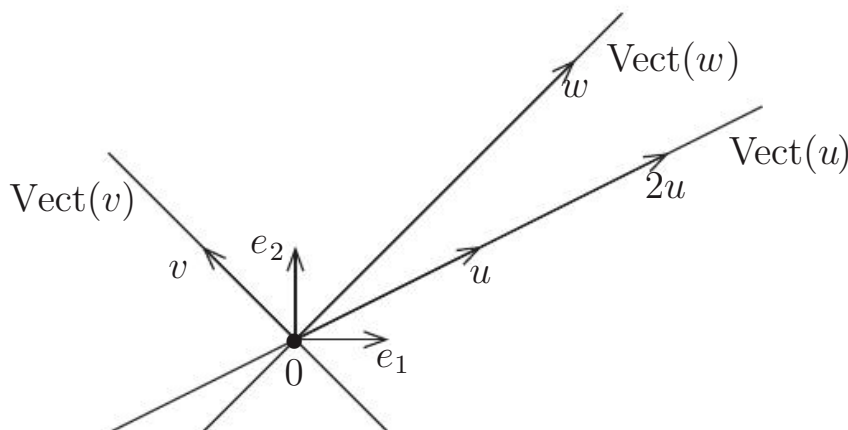
3.1 a) $w = (3, 3)$.

b) Si u et v étaient colinéaires, il existerait des réels λ et μ non nuls tels que

$$\lambda u + \mu v = 0, \text{ soit } (2\lambda - \mu, \lambda + \mu) = 0. \text{ La résolution du système } \begin{cases} 2\lambda - \mu = 0 \\ \lambda + \mu = 0 \end{cases}$$

donne $\lambda = \mu = 0$. Les deux vecteurs u et v ne sont pas colinéaires. On peut aussi appliquer le critère de 3.8.

c)



d) En posant $w' = w + 3v = (0, 6)$, on a $\text{Vect}(v, w) = \text{Vect}(v, w') = \mathbb{R}^2$ d'après la proposition 3.8, donc $\text{Vect}(u)$ est strictement contenu dans $\text{Vect}(v, w)$.

3.2 Les dessins sont laissés aux lectrices et lecteurs. Pour les cas où il n'y a pas stabilité, un seul contre-exemple suffit.

• $A_1 = \{(x, y) \mid x, y \in \mathbb{R}, 2x + 3y = 0\}$

Stabilité par rapport à la somme. Pour tous couples (x, y) et (x', y') de A_1 , on a $(x, y) + (x', y') = (x + x', y + y')$. Ce vecteur appartient à A_1 car :

$$2(x + x') + 3(y + y') = (2x + 3y) + (2x' + 3y'), \text{ et cette dernière somme est nulle d'après l'hypothèse.}$$

Stabilité par produit par un scalaire. Pour tout réel λ et tout couple (x, y) de A_1 , on a $\lambda(x, y) = (\lambda x, \lambda y)$. Ce vecteur appartient à A_1 car :

$$2(\lambda x) + 3(\lambda y) = \lambda(2x + 3y) = 0.$$

Stabilité par combinaison linéaire. Il suffit de dire que la stabilité par somme et la stabilité par produit par un scalaire entraînent la stabilité par combinaison linéaire.

On peut aussi faire une démonstration directe : pour tous couples (x, y) et (x', y') de A_1 et tous réels λ et μ , on a $\lambda(x, y) + \mu(x', y') = (\lambda x + \mu x', \lambda y + \mu y')$ et ce vecteur appartient à A_1 car :

$$2(\lambda x + \mu x') + 3(\lambda y + \mu y') = \lambda(2x + 3y) + \mu(2x' + 3y') = 0.$$

Pour écrire A_1 avec la notation Vect, on remarque que $A_1 = \left\{ \left(x, -\frac{2}{3}x \right) \mid x \in \mathbb{R} \right\}$,
donc $A_1 = \text{Vect} \left(\left(1, -\frac{2}{3} \right) \right)$.

On peut également faire cet exercice d'entraînement à l'envers : voir directement que A_1 est Vect $\left(\left(1, -\frac{2}{3} \right) \right)$, d'où découlent les trois propriétés de stabilité sans aucun calcul, d'après les propriétés de stabilité des espaces engendrés.

- A_2 n'est pas stable par somme comme le montre l'exemple du vecteur $(-1, 1)$ de A_2 , dont la somme avec lui-même $(-2, 2)$ n'y est pas.

A_2 n'est pas stable par produit par un scalaire, comme le montre l'exemple du produit de $(-1, 1)$ par 2.

- A_3 n'est pas stable par somme, comme le montre l'exemple des vecteurs $(1, 0)$ et $(0, 1)$, tous deux dans A_3 , mais dont la somme $(1, 1)$ n'y est pas.

A_3 n'est pas stable par produit par un scalaire, comme le montre l'exemple du produit de $(1, 0)$ par 3.

- Comme $A_4 = \{a(1, 1) + b(1, -1) \mid a, b \in \mathbb{R}\}$, on a $A_4 = \text{Vect}((1, 1), (1, -1))$ qui est stable par combinaison linéaire.

- A_5 est stable par produit par un scalaire : si x et y sont non nuls ainsi que λ , alors les deux composantes de $\lambda(x, y)$ sont non nulles, et si λ est nul, on obtient le vecteur 0 qui est aussi dans A_5 . Par contre, A_5 n'est pas stable par somme, comme le montre l'exemple de $(-1, 1) + (1, 1)$.

Pour se représenter A_5 , on pensera que les seuls vecteurs qui ne sont pas dans A_5 sont les vecteurs non nuls colinéaires aux deux vecteurs de la base canonique de $\mathbb{R}^2$.

- A_6 est stable par somme mais non par produit par des nombres strictement négatifs, par exemple $(-1)(1, 1) \notin A_6$
- A_7 est stable par somme mais non par produit par des nombres non entiers, par exemple $\sqrt{2}(1, 1) \notin A_7$.

3.3 L'algorithme du pivot montre que $w - 5v - 3u = 0$. On peut tirer de cette relation l'un des vecteurs en fonction des deux autres, ce qui permet de prouver que les espaces engendrés par deux des trois vecteurs sont égaux à Vect (u, v, w) .

3.4 Vect $(u_1, \dots, u_p) = \text{Vect}(u_2, \dots, u_p)$, d'après la proposition 2 de 3.4.

3.5 Cet exercice est un exercice d'entraînement à la méthode du pivot et on adoptera la présentation de l'exemple donné dans 3.10. Les égalités entre « Vect » résultent de la proposition 2 de 3.4, ce qui ne sera pas mentionné.

a) On permute l'ordre de u_1 et u_2 pour faciliter le calcul : on pose donc $u'_1 = u_2$ et $u'_2 = u_1$:

$$\begin{array}{cc} u'_1 & u'_2 \\ -1 & 3 \\ 5 & 1 \end{array}$$

On pose alors $u''_2 = u'_2 + 3u'_1$:

$$\begin{array}{cc} u'_1 & u''_2 \\ -1 & 0 \\ 5 & 16 \end{array}$$

On a $\text{Vect}(u_1, u_2) = \text{Vect}(u'_1, u''_2)$. Comme la famille (u'_1, u''_2) est triangulaire, son rang est 2 ; on a vu que cela implique que (u_1, u_2) engendre $\mathbb{R}^2$. Par conséquent, $e_1 \in \text{Vect}(u_1, u_2)$.

Pour exprimer e_1 en fonction de u_1 et de u_2 , on peut appliquer la méthode du pivot à la famille (u'_1, u''_2, e_1) ; en posant $e'_1 = e_1 + u'_1$, on obtient :

$$\begin{array}{ccc} u'_1 & u''_2 & e'_1 \\ -1 & 0 & 0 \\ 5 & 16 & 5 \end{array}$$

d'où $e'_1 = \frac{5}{16}u''_2$; on en déduit $e_1 + u'_1 = \frac{5}{16}(u'_2 + 3u'_1)$, puis $e_1 = \frac{1}{16}(5u'_2 - u'_1)$
 $= \frac{1}{6}(5u_1 - u_2)$.

Une autre méthode pour obtenir l'expression de e_1 en fonction de u_1 et u_2 est de chercher directement les réels α et β tels que $e_1 = \alpha u_1 + \beta u_2$, ce qui conduit au système linéaire :

$$\begin{cases} 3\alpha - \beta = 1 \\ \alpha + 5\beta = 0 \end{cases}$$

qui donne la même relation.

b) On part de :

$$\begin{array}{ccc} u_1 & u_2 & u_3 \\ -1 & 1 & 1 \\ 1 & 2 & 7 \\ -3 & 5 & 1 \end{array}$$

On pose $u'_2 = u_2 + u_1$ et $u'_3 = u_3 + u_1$, ce qui donne :

$$\begin{array}{ccc} u_1 & u'_2 & u'_3 \\ -1 & 0 & 0 \\ 1 & 3 & 8 \\ -3 & 2 & -2 \end{array}$$

En posant $u_3'' = u_3' - \frac{8}{3}u_2'$, on obtient :

$$\begin{array}{ccc} u_1 & u_2' & u_3'' \\ -1 & 0 & 0 \\ 1 & 3 & 0 \\ -3 & 2 & -\frac{22}{3} \end{array}$$

On a $\text{Vect}(u_1, u_2, u_3) = \text{Vect}(u_1, u_2', u_3'')$. Comme la famille (u_1, u_2', u_3'') est triangulaire, son rang est 3 et elle engendre $\mathbb{R}^3$.

On peut aussi, après la première étape, permuter le rôle des vecteurs de base e_2 et e_3 : les coordonnées des vecteurs s'écrivent alors :

$$\begin{array}{ccc} u_1 & u_2' & u_3' \\ e_1 & -1 & 0 & 0 \\ e_3 & -3 & 2 & -2 \\ e_2 & 1 & 3 & 8 \end{array}$$

et on voit immédiatement que la famille (u_1, u_2', u_3''') , avec $u_3''' = u_2' + u_3'$ est triangulaire. On conclut de la même façon.

Pour trouver l'expression de e_1 en fonction de u_1, u_2, u_3 , on peut utiliser la famille triangulaire u_1, u_2', u_3''' , écrire :

$$\begin{array}{ccc} u_1 & u_2' & u_3'' & e_1 \\ -1 & 0 & 0 & 1 \\ 1 & 3 & 0 & 0 \\ -3 & 2 & -\frac{22}{3} & 0 \end{array}$$

et calculer successivement $e_1' = e_1 + u_1, e_1'' = e_1' - \frac{1}{3}u_2'$:

$$\begin{array}{cccccc} u_1 & u_2' & u_3'' & e_1 & e_1' & e_1'' \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 & 0 \\ -3 & 2 & -\frac{22}{3} & 0 & -3 & -\frac{11}{3} \end{array}$$

On en déduit $e_1'' = \frac{1}{2}u_3''$, puis $e_1' - \frac{1}{3}u_2' = \frac{1}{2}\left(u_3' - \frac{8}{3}u_2'\right)$ et enfin $e_1 = -\frac{3}{2}u_1 - u_2 + \frac{1}{2}u_3$, résultat qu'il est facile de vérifier sur les trois coordonnées.

Une autre méthode pour obtenir cette relation est de chercher des réels α, β et γ tels que $e_1 = \alpha u_1 + \beta u_2 + \gamma u_3$, ce qui conduit à la résolution du système :

$$\begin{cases} -\alpha + \beta + \gamma = 1 \\ \alpha + 2\beta + 7\gamma = 0 \\ -3\alpha + 5\beta + \gamma = 0 \end{cases}$$

La résolution de ce système (un algorithme est donné au chapitre 4) donne bien sûr les valeurs de α, β et γ trouvées ci-dessus.

On trouve que la famille est de rang 3 et engendre $\mathbb{R}^3$. On a

$$e_1 = -\frac{1}{6}u_1 + \frac{2}{15}u_2 + \frac{3}{10}u_3.$$

3.6 a) Pour montrer l'égalité des deux ensembles, on va prouver les deux inclusions réciproques : $\text{Vect}(v_1, v_2) \subset \text{Vect}(u_1, u_2)$ et $\text{Vect}(u_1, u_2) \subset \text{Vect}(v_1, v_2)$.

D'après les propriétés des « Vect », pour montrer la première inclusion, il suffit de montrer que $v_1, v_2 \in \text{Vect}(u_1, u_2)$. On utilise la méthode du pivot, en partant de la famille u_2, u_1, v_1, v_2 :

$$\begin{array}{cccc} u_2 & u_1 & v_1 & v_2 \\ 1 & 2 & 3 & 5 \\ -1 & 3 & 7 & 0 \\ -2 & -1 & 0 & -7 \end{array}$$

On pose $u'_1 = u_1 - 2u_2, v'_1 = v_1 - 3u_2, v'_2 = v_2 - 5u_2$, ce qui donne :

$$\begin{array}{cccc} u_2 & u'_1 & v'_1 & v'_2 \\ 1 & 0 & 0 & 0 \\ -1 & 5 & 10 & 5 \\ -2 & 3 & 6 & 3 \end{array}$$

Sans continuer la méthode, on voit que $v'_1 = 2u'_1$ et que $v'_2 = u'_1$. Comme $u'_1 \in \text{Vect}(u_1, u_2)$, on en déduit que $v'_1, v'_2 \in \text{Vect}(u_1, u_2)$ donc $v_1, v_2 \in \text{Vect}(u_1, u_2)$ d'où la première inclusion.

L'inclusion réciproque se montre de la même manière.

b) Faire comme pour le **a**).

4.1 HISTOIRE ANCIENNE

Babylone, – 1800 av. J.-C. Quelques tablettes babyloniennes donnent à résoudre, sous la forme de problèmes concrets, ce que nous appelons maintenant des systèmes linéaires de deux équations à deux inconnues. Par exemple :

$$\left\{ \begin{array}{l} x - y = 10 \\ \frac{2}{3}x - \frac{y}{2} = 8 + \frac{1}{3} \end{array} \right. \quad \left\{ \begin{array}{l} \frac{x}{7} + \frac{y}{11} = 1 \\ \frac{6x}{7} = \frac{10y}{11} \end{array} \right.$$

Liu Hui. Le mathématicien chinois Liu Hui vivait vers 250-300. Il est célèbre pour avoir écrit le *Jiuzhang suanshu*. Le titre de ce livre peut se traduire par : *Neuf chapitres sur la science du calcul*. C'est le grand classique des mathématiques anciennes de l'Extrême Orient. Le texte tel qu'il nous est parvenu présente 246 problèmes répartis par thèmes. Il a été édité dans un livre magnifique de plus de 1100 pages par Karine Chemla et Guo Shuchun (Dunod, 2004). Dans le chapitre 8, qui contient 18 problèmes, Liu Hui présente la résolution de systèmes concrets d'équations linéaires ayant jusqu'à 6 inconnues, en exposant en quelques phrases incisives (son style n'a rien à voir avec celui d'Euclide) la méthode *fangcheng* que nous appelons maintenant l'algorithme ou la méthode du pivot de Gauss.

Suivons sa démarche pour résoudre le problème suivant, où il considère des bottes de céréales de différentes qualités :

- 3 bottes de qualité supérieure, 2 de qualité moyenne et 1 de qualité médiocre donnent 39 *dou* de grains ;
- 2 bottes de qualité supérieure, 3 de qualité moyenne et 1 de qualité médiocre donnent 34 *dou* de grains ;
- 1 botte de qualité supérieure, 2 de qualité moyenne et 3 de qualité médiocre donnent 26 *dou* de grains.

Combien de grains donne chaque type de bottes ?

En notant x , y et z les quantités de grains données par les bottes de qualité supérieure, moyenne et médiocre, ce problème équivaut à la résolution du système d'équations :

$$\begin{cases} 3x + 2y + z = 39 & l_1 \\ 2x + 3y + z = 34 & l_2 \\ x + 2y + 3z = 26 & l_3 \end{cases}$$

Démarche de Liu Hui. Les calculs que fait Liu Hui, qui dispose les coefficients en colonnes, peuvent être interprétés de la façon suivante. D'abord, calculer $3l_2$ et soustraire du résultat $2l_1$, ce qui donne :

$$\begin{cases} 3x + 2y + z = 39 & l_1 \\ 5y + z = 24 & l'_2 = 3l_2 - 2l_1 \\ x + 2y + 3z = 26 & l_3 \end{cases}$$

Puis calculer $l'_3 = 3l_3 - l_1$:

$$\begin{cases} 3x + 2y + z = 39 & l_1 \\ 5y + z = 24 & l'_2 = 3l_2 - 2l_1 \\ 4y + 8z = 39 & l'_3 = 3l_3 - l_1 \end{cases}$$

Enfin calculer $l''_3 = 5l'_3 - 4l'_2$:

$$\begin{cases} 3x + 2y + z = 39 & l_1 \\ 5y + z = 24 & l'_2 \\ 36z = 99 & l''_3 = 5l'_3 - 4l'_2 \end{cases}$$

Ces calculs permettent à Liu Hui d'obtenir z dans l''_3 , puis y , en reportant la valeur trouvée pour z dans l'_2 , et enfin x à l'aide de l_1 .

Décrite sur un exemple simple, minimal en quelque sorte, mais qui a valeur d'exposé d'une méthode générale, c'est, 1 500 ans avant Gauss, la méthode de résolution des systèmes linéaires que nous allons exposer dans ce chapitre.

4.2 LEIBNIZ, CRAMER, GAUSS

Leibniz. L'introduction des lettres par François Viète (1540-1603) à la fin du ^{xvi}^e siècle permettait d'écrire des équations générales, alors que, jusque-là, on ne présentait que des exemples numériques.

Les systèmes linéaires posaient cependant des problèmes de notation particuliers.

Dans un manuscrit de 1678 retrouvé 300 ans plus tard, Leibniz trouve une idée. Le 28 avril 1693, il la décrit au marquis Guillaume de l'Hospital (1661-1704) : l'usage des lettres $a, b, c, \dots$ ne permet pas d'exprimer les « harmonies » d'un système linéaire, écrit-il, entendant par là qu'elles ne reflètent pas le parallélisme des calculs dans les différentes lignes quand, par exemple, on utilise la première ligne pour faire disparaître les termes en x des lignes suivantes. Il explique à son ami qu'il utilise pour sa part des chiffres : 10, 11, 12, 20, ..., qu'il ne considère pas comme des nombres véritables. Par exemple, il écrit le système :

$$\begin{cases} 10 + 11x + 12y = 0 \\ 20 + 21x + 22y = 0 \\ 30 + 31x + 32y = 0 \end{cases}$$

C'est un essai de notation indicielle et, pour notre part, nous remplacerions, par exemple, 10 par $a_{1,0}$, 11 par $a_{1,1}$, 12 par $a_{1,2}$, 20 par $a_{2,0}$, etc., en écrivant :

$$\begin{cases} a_{1,0} + a_{1,1}x + a_{1,2}y = 0 \\ a_{2,0} + a_{2,1}x + a_{2,2}y = 0 \\ a_{3,0} + a_{3,1}x + a_{3,2}y = 0 \end{cases}$$

Leibniz remarque que les calculs pour éliminer x entre la première et la troisième équation se déduisent des calculs pour éliminer x entre la première et la seconde équation en remplaçant les 2 par des 3.

Cramer. En 1750, Gabriel Cramer (1704-1752) donne lui aussi des formules pour les solutions de systèmes d'équations linéaires. Sa notation ne comporte qu'un seul indice, placé en fait en exposant, mais son texte va ouvrir un nouveau domaine : la théorie des déterminants ; nous en reparlerons au chapitre 14.

Les formules de Cramer nécessitent de nombreuses multiplications pour résoudre un système linéaire ; ce nombre croît rapidement avec le nombre d'inconnues. C'était, et cela reste, un grave inconvénient.

Gauss. La contribution de Gauss à la résolution de systèmes linéaires est beaucoup plus complexe à comprendre, et on ne retrouve dans ses textes que des parties de l'algorithme que nous allons exposer dans ce chapitre et qui porte son nom. Les problèmes posés par la résolution des systèmes linéaires dans les années 1800 sont liés aux observations astronomiques et géodésiques (voir 16.8).

4.3 SYSTÈMES LINÉAIRES

Définition. Soient $n, p \geq 1$ des entiers. Un système linéaire L à coefficients réels de n équations linéaires $l_1, \dots, l_n$ à p inconnues $x_1, \dots, x_p$ est de la forme :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = b_1 & l_1 \\ \vdots & \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = b_n & l_n \end{cases}$$

où les coefficients a_{ij} et b_i sont des réels (dans les exemples, les coefficients seront presque toujours entiers, pour conduire à des calculs simples).

Le système :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = 0 \\ \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = 0 \end{cases}$$

sera dit système linéaire sans second membre associé à L .

Nous noterons $S(L)$ l'ensemble des p -uplets $(x_1, \dots, x_p)$ vérifiant toutes les équations de L , autrement dit l'ensemble des solutions de L .

Définition : systèmes linéaires équivalents. Deux systèmes linéaires L_1 et L_2 à p inconnues $x_1, \dots, x_p$ sont dits équivalents si $S(L_1) = S(L_2)$.

Commentaire. Les méthodes de résolution des systèmes linéaires présentées dans ce chapitre vont consister à se ramener à un système équivalent au système initial et plus facile à résoudre.

On peut se demander ce que voudrait dire résoudre un système linéaire dans les cas $n = 0$ ou $p = 0$. Ces cas peuvent effectivement se produire ; on les précisera en 7.9 ; le lecteur ou la lectrice évitera de croire que la compréhension de ces cas est essentielle.

4.4 EXEMPLES DE RÉOLUTIONS

Avant de définir précisément ce qu'est résoudre un système linéaire, nous allons donner quelques exemples où nous noterons les inconnues $x, y, \dots$ et les coefficients $a, b, \dots$

Exemple 1 : $p = n = 1$. Le système L se réduit à une équation :

$$ax = b$$

- Si $a \neq 0$, alors le système a une solution unique : $S(L) = \left\{ \frac{b}{a} \right\}$.
- Si $a = 0$ et si $b \neq 0$, le système n'a aucune solution : $S(L) = \emptyset$.
- Si $a = 0$ et si $b = 0$, tout réel satisfait le système : $S(L) = \mathbb{R}$.

On voit déjà sur cet exemple simple la diversité des possibilités.

Exemple 2 : $p = 2, n = 1$. Le système L se réduit à une équation :

$$ax + by = c$$

- Si $a = b = c = 0$, alors $S(L) = \mathbb{R}^2$.
- Si $a = b = 0$ et $c \neq 0$, alors $S(L) = \emptyset$.
- Si $a \neq 0$, les couples (x, y) solutions sont de la forme $\left(-\frac{b}{a}y + \frac{c}{a}, y \right)$ où y est un réel quelconque, donc $S(L) = \left\{ \left(-\frac{b}{a}y + \frac{c}{a}, y \right) \mid y \in \mathbb{R} \right\} = u_0 + \text{Vect}(u_1)$ avec $u_0 = \left(\frac{c}{a}, 0 \right)$, $u_1 = \left(-\frac{b}{a}, 1 \right)$. On voit que u_0 est une solution particulière de l'équation avec second membre et que $u = yu_1$ décrit l'ensemble des solutions de l'équation sans second membre quand y varie dans $\mathbb{R}$.
- Si $b \neq 0$, les couples (x, y) solutions sont de la forme $\left(x, -\frac{a}{b}x + \frac{c}{b} \right)$ où x est un réel quelconque, donc $S(L) = v_0 + \text{Vect}(v_1)$ avec $v_0 = \left(0, \frac{c}{b} \right)$, $v_1 = \left(1, -\frac{a}{b} \right)$. Cet ensemble coïncide avec le précédent si $ab \neq 0$.

Exemple 3 : $p = n = 2$. Le système L s'écrit :

$$\begin{cases} ax + by = c & l \\ a'x + b'y = c' & l' \end{cases}$$

- 1) On peut commencer par le cas où $a = b = a' = b' = 0$:
 - a) si $c \neq 0$ ou si $c' \neq 0$, le système contient l'équation $0 = c$ ou l'équation $0 = c'$ et $S(L) = \emptyset$.
 - b) si $c = c' = 0$, le système se réduit à l'équation $0 = 0$ et $S(L) = \mathbb{R}^2$.
- 2) Supposons que l'un au moins des quatre coefficients a, b, a', b' soit non nul. On peut supposer que c'est a , en permutant si besoin les inconnues x et y ou les équations l et l' . Le système L est équivalent au système L' obtenu en soustrayant $\frac{a'}{a}$ fois la première ligne de la seconde :

$$\begin{cases} ax + by = c & l \\ \left(b' - \frac{a'}{a}b\right)y = c' - \frac{a'}{a}c & l'' = l' - \frac{a'}{a}l \end{cases}$$

En effet, si (x, y) vérifie l et l' , il vérifie l et l'' et, réciproquement, s'il vérifie l et l'' , il vérifie l et l' car $l' = l'' + \frac{a'}{a}l$. On a donc $S(L') = S(L)$.

Le système L' est équivalent au système L'' obtenu en multipliant la seconde ligne par a :

$$\begin{cases} ax + by = c & l \\ (ab' - a'b)y = ac' - a'c & al'' = al' - a'l \end{cases}$$

On est amené à distinguer différents cas.

- a) Si $ab' - a'b = 0$ et $ac' - a'c \neq 0$, le système n'a aucune solution et $S(L) = \emptyset$.
- b) Si $ab' - a'b = ac' - a'c = 0$, alors le système est équivalent au système formé de l'unique équation l et on est ramené à l'exemple 2 : $S(L) = u_0 + \text{Vect}(u_1)$ avec

$$u_0 = \left(\frac{c}{a}, 0\right), u_1 = \left(-\frac{b}{a}, 1\right).$$

- c) Si $ab' - a'b \neq 0$, alors le système a une unique solution :

$$S(L) = \left\{ \left(\frac{bc' - b'c}{ab' - a'b}, \frac{ac' - a'c}{ab' - a'b} \right) \right\}.$$

4.5 SYSTÈMES ÉQUIVALENTS

Proposition. Soit L un système linéaire à coefficients réels de $n \geq 2$ équations linéaires $l_1, \dots, l_n$ à p inconnues $x_1, \dots, x_p$.

- 1) Soit λ un réel non nul. Alors L est équivalent au système L' formé par les équations $\lambda l_1, l_2, \dots, l_n$.
- 2) On note L' le système obtenu en remplaçant l_1 par l'équation $l'_1 = l_1 + \sum_{2 \leq i \leq n} \lambda_i l_i$. Alors L et L' sont des systèmes équivalents.

Démonstration. Dans les deux cas, pour montrer l'égalité des ensembles $S(L)$ et $S(L')$, il suffit de montrer que chacun d'eux est inclus dans l'autre, ce qui est facile pour le premier cas.

Dans le second cas, si $(x_1, \dots, x_p)$ est une solution de L , c'est visiblement une solution de L' et si $(x_1, \dots, x_p)$ est une solution de L' , c'est une solution de L car

$$l_1 = l'_1 - \sum_{2 \leq i \leq n} \lambda_i l_i. \quad \square$$

Commentaire. La proposition ci-dessus, analogue à la proposition 2 de 3.4, sera d'une grande utilité par la suite. On pourra s'en souvenir sous une forme un peu plus générale :

- 1) on obtient un système équivalent à un système donné en multipliant une ligne par un scalaire non nul ;
- 2) on obtient un système équivalent à un système donné en ajoutant à une des équations du système une combinaison linéaire des autres équations.

Les débutant(e)s prendront cependant garde à ne pas succomber à l'ivresse des combinaisons simultanées dans tous les sens : il faut sans cesse avoir en tête l'idée de passer d'un système à un système équivalent, et les propriétés précédentes sont celles qui permettent de s'en assurer.

Remarque. Quand on permute les lignes d'un système, on obtient un système équivalent. Par exemple, le système :

$$\begin{cases} ax + by = c & l \\ a'x + b'y = c' & l' \end{cases}$$

est équivalent au système :

$$\begin{cases} a'x + b'y = c' & l' \\ ax + by = c & l \end{cases}$$

4.6 SYSTÈMES TRIANGULAIRES ET ÉCHELONNÉS

Systèmes triangulaires. Un système linéaire :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = b_1 & l_1 \\ \vdots & \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = b_n & l_n \end{cases}$$

est triangulaire s'il existe un entier $r \leq n$ tel que :

- si $r < n$ on ait $a_{ij} = 0$ pour $r < i \leq n$ et $1 \leq j \leq p$;
- $a_{ij} = 0$ pour tout couple (i, j) tel que $1 \leq j < i \leq r$;
- $a_{ii} \neq 0$ pour $1 \leq i \leq r$.

Un système est donc triangulaire s'il se présente ainsi :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r + \dots + a_{1p}x_p = b_1 & l_1 \\ \quad a_{22}x_2 + \dots + a_{2r}x_r + \dots + a_{2p}x_p = b_2 & l_2 \\ \quad \quad \quad \dots \quad \quad \quad \dots \quad \quad \quad \dots & \dots \\ \quad \quad \quad \quad a_{rr}x_r + \dots + a_{rp}x_p = b_r & l_r \\ \quad \quad \quad \quad \quad \quad \quad 0 = b_{r+1} & l_{r+1} \\ \quad \quad \quad \quad \quad \quad \quad \dots & \dots \\ \quad \quad \quad \quad \quad \quad \quad 0 = b_n & l_n \end{cases}$$

où les a_{ii} , $1 \leq i \leq r$, sont non nuls. Les lignes $l_{r+1}, \dots, l_n$ peuvent ne pas exister.

Systèmes échelonnés. Cette forme est un peu plus générale que la forme triangulaire et se rencontre naturellement dans les applications. Pour toute équation l_i , notons j_i l'indice du premier coefficient non nul, s'il existe, c'est-à-dire que l_i est de la forme : $a_{ij_i}x_{j_i} + \dots + a_{ip}x_p = b_i$ avec $a_{ij_i} \neq 0$.

Nous dirons que le système est échelonné s'il existe un entier $r \leq n$ tel que :

- si $r < n$, on a $a_{ij} = 0$ pour $r < i \leq n$ et $1 \leq j \leq p$;
- pour $i \leq r$, j_i existe et la suite $j_1, \dots, j_r$ est strictement croissante.

Autrement dit, un système est échelonné si les coefficients non nuls des équations se présentent avec une sorte d'escalier à marches de longueurs variables marquant la séparation entre une zone composée uniquement de 0 et une zone où les lignes situées à droite de l'escalier commencent par des termes non nuls, comme dans l'exemple suivant d'un système de 5 équations à 6 inconnues :

$$\left\{ \begin{array}{cccccccl} 5x_1 - x_2 & -x_3 & +2x_4 & & +x_6 & = b_1 & l_1 \\ & 3x_3 & -x_4 & +2x_5 & & = b_2 & l_2 \\ & & & -x_5 & +x_6 & = b_3 & l_3 \\ & & & & 5x_6 & = b_4 & l_4 \\ & & & & 0 & = b_5 & l_5 \end{array} \right.$$

Tout système échelonné peut être vu comme un système triangulaire en les variables $x_{j_1}, \dots, x_{j_r}$, en faisant passer les termes correspondant aux autres variables à la fin des équations. Pour le système échelonné précédent, on obtient un système triangulaire en faisant passer les termes en x_2 et x_4 des deux premières équations après le terme en x_6 .

4.7 MÉTHODE DU PIVOT DE GAUSS

Dans cette section, nous décrivons la méthode appelée méthode ou algorithme du pivot de Gauss pour la résolution des systèmes linéaires.

Soit L un système linéaire :

$$\left\{ \begin{array}{cccccl} a_{11}x_1 + & \dots & + a_{1p}x_p & = b_1 & l_1 \\ \vdots & & \vdots & \vdots & \vdots \\ a_{n1}x_1 + & \dots & + a_{np}x_p & = b_n & l_n \end{array} \right.$$

où les coefficients a_{ij} et b_i sont des réels.

- 1) Si tous les coefficients a_{ij} et b_i sont nuls, alors $S(L) = \mathbb{R}^p$.
- 2) S'il existe un i tel que $a_{ij} = 0$ pour tout j , $1 \leq j \leq p$, et que $b_i \neq 0$, autrement dit l_i est de la forme $0x_1 + \dots + 0x_p = b_i$, on a $S(L) = \emptyset$; le système n'a aucune solution ; on dit qu'il est incompatible.

- 3) Sinon, on retire les équations dont tous les coefficients sont nuls, c'est-à-dire les équations de la forme $0x_1 + \dots + 0x_p = 0$. Notons toujours L le système obtenu.
- 4) On considère la première inconnue ayant un coefficient non nul dans l'une au moins des équations. On peut supposer que c'est dans la première équation, en changeant si besoin l'ordre des équations (ce qui ne change pas $S(L)$). On appelle premier pivot ce coefficient. On suppose que c'est a_{11} pour simplifier, quitte à renuméroter les inconnues. Le système L est équivalent au système L' formé par les équations l_1 et, pour $2 \leq i \leq n$, $l'_i = l_i - \frac{a_{i1}}{a_{11}}l_1$.

Ce système a la forme suivante :

$$\left\{ \begin{array}{ccccccc} a_{11}x_1 + & & \dots & + a_{1p}x_p & = b_1 & l_1 \\ & a'_{22}x_2 + & \dots & + a'_{2p}x_p & = b'_2 & l'_2 \\ & \vdots & & \vdots & \vdots & \vdots \\ & a'_{n2}x_2 + & \dots & + a'_{np}x_p & = b'_n & l'_n \end{array} \right.$$

On s'occupe alors du système formé par $l'_2, \dots, l'_n$ et on recommence exactement la même procédure.

- 1') On retire les équations dont tous les coefficients a'_{ij} et b'_i sont nuls. S'il ne reste aucune équation, on a terminé la première partie de l'algorithme.
- 2') S'il existe un $i \geq 2$ tel que $a'_{ij} = 0$ pour $1 \leq j \leq p$ et que $b'_i \neq 0$, le système est incompatible et $S(L) = \emptyset$.
- 3') Sinon, l'un des coefficients $a'_{i2}, i \geq 2$ de la seconde colonne est non nul (si ce n'est pas le cas, on regarde les coefficients $a'_{i3}, i \geq 2$ de la troisième colonne, etc.) ; on peut supposer que c'est a'_{22} en faisant les changements d'ordre d'équations nécessaires (sans toucher à l_1). On l'appelle second pivot.

Le système L est équivalent au système L'' formé par les équations l_1, l'_2 et des équations où les coefficients de x_1 et de x_2 sont nuls, obtenues par les formules

$$l''_i = l'_i - \frac{a'_{i2}}{a'_{22}}l'_2 \text{ pour } i \geq 3.$$

On continue cet algorithme jusqu'à rencontrer une incompatibilité ou obtenir un système échelonné de $r \leq n$ équations équivalent à L sans ligne de la forme $0 = 0$ et tel que, si j_i note l'indice du premier coefficient non nul de l'équation l_i , la suite $j_1, \dots, j_r$ soit strictement croissante.

Une nouvelle partie de l'algorithme commence alors. On passe tout d'abord dans le second membre tout ce qui concerne les x_j d'indice différent de $j_1, \dots, j_r$. Pour ne pas multiplier les problèmes de notation, supposons que $j_i = i$ pour $1 \leq i \leq r$, ce qui revient à renuméroter les inconnues, et utilisons les notations a_{ij} pour les coefficients

de ce nouveau système. La forme du système est, en reprenant la notation l_i pour ces nouvelles lignes :

$$\left\{ \begin{array}{ccccccc} a_{11}x_1 + & \cdots & \cdots & = b_1 & + \cdots & l_1 \\ & a_{22}x_2 + & \cdots & = b_2 & + \cdots & l_2 \\ & & \vdots & & \vdots & \\ & & a_{rr}x_r & = b_r & + \cdots & l_r \end{array} \right.$$

où $a_{11}, \dots, a_{rr}$ sont non nuls.

On pourra alors calculer x_r dans la dernière ligne, reporter cette valeur dans l_{r-1} pour obtenir x_{r-1} , etc. Mais on peut aussi procéder différemment : on utilise la dernière équation pour faire disparaître les termes en x_r des équations précédentes (il suffit de calculer, pour $1 \leq i < r$, $l'_i = l_i - \alpha_i l_r$ avec α_i réel convenable), puis on utilise la $(r-1)$ -ième équation qui vient d'être obtenue et qui ne contient plus de terme en x_r pour faire disparaître les termes en x_{r-1} des équations précédentes, etc. On obtient un système équivalent au système L de la forme :

$$\left\{ \begin{array}{ccc} a_{11}x_1 & = b_1 & + \cdots \\ a_{22}x_2 & = b_2 & + \cdots \\ \vdots & \vdots & \\ a_{rr}x_r & = b_r & + \cdots \end{array} \right.$$

En calculant $x_1, \dots, x_r$ dans ce système en fonction des b_i et des $x_{r+1}, \dots, x_p$, on obtient un système de la forme :

$$\left\{ \begin{array}{cccc} x_1 & = c_1 + d_{1,r+1}x_{r+1} & + \cdots + d_{1,p}x_p \\ x_2 & = c_2 + d_{2,r+1}x_{r+1} & + \cdots + d_{2,p}x_p \\ \vdots & \vdots & \vdots & \vdots \\ x_r & = c_r + d_{r,r+1}x_{r+1} & + \cdots + d_{r,p}x_p \end{array} \right.$$

Les solutions de ce système sont les vecteurs $(x_1, \dots, x_p)$ définis par un choix arbitraire de $x_{r+1}, \dots, x_p$ et par les formules précédentes. On peut le décrire de la façon suivante. Posons :

- $s = (c_1, \dots, c_r, 0, \dots, 0)$
- $u_{r+k} = (d_{1,r+k}, \dots, d_{r,r+k}, 0, \dots, 0) + e_{r+k}$ pour $1 \leq k \leq p-r$. Ici, e_{r+k} désigne le $(r+k)$ -ième vecteur de la base canonique de $\mathbb{R}^p$: il indique un 1 dans la suite des coordonnées de u_{r+k} au rang $r+k$: $u_{r+k} = (d_{1,r+k}, \dots, d_{r,r+k}, 0, \dots, 1, \dots, 0)$.

On voit alors que l'ensemble des solutions de L a une description de la même forme que pour les exemples de 4.4 :

$$S(L) = s + \text{Vect}(u_{r+1}, \dots, u_p).$$

Bilan des opérations. Pour transformer le système linéaire initial en un système échelonné, nous avons utilisé un certain nombre de fois les deux seuls types d'opérations suivants :

- la permutation de deux lignes entre elles ;
- l'addition d'un multiple d'une ligne à une autre.

Pour la résolution de petits systèmes où les coefficients sont entiers, d'autres opérations sont très utiles : diviser tous les coefficients par un facteur commun (en particulier, changer le signe de tous les coefficients), ou multiplier par un dénominateur commun un système où apparaissent des fractions.

Rang d'un système. On verra en 7.9 que le nombre r , que nous appellerons rang du système linéaire, ne dépend pas des choix faits dans l'application de la méthode du pivot. Cette invariance est un résultat théorique important.

On peut résumer les résultats obtenus par le théorème suivant.

Théorème. Soit L un système linéaire et L' le système sans second membre associé ; nous utilisons les notations de 4.3.

1) L'ensemble des solutions $S(L)$ peut être vide.

Si $S(L)$ n'est pas vide et si on note $s = (s_1, \dots, s_p)$ un élément de $S(L)$ (qu'on peut appeler solution particulière de L), on obtient les résultats suivants.

2) Le p -uplet $(x_1, \dots, x_p)$ de $\mathbb{R}^p$ appartient à $S(L)$ si et seulement si le p -uplet $(y_1, \dots, y_p)$ défini par $y_i = x_i - s_i$ pour $1 \leq i \leq p$ est dans $S(L')$.

3) Il existe un entier r , appelé rang du système linéaire, et des vecteurs s et $u_{r+1}, \dots, u_p$ de $\mathbb{R}^n$ tels que $S(L') = \text{Vect}(u_{r+1}, \dots, u_p)$ et que :

$$S(L) = s + \text{Vect}(u_{r+1}, \dots, u_p).$$

4) Le système a une solution unique si et seulement si $r = p$.

Commentaire. Le 2) énonce un résultat du même type que ceux vus dans les chapitres 1 et 2 : la solution du système L , le système « avec second membre », est la somme d'une solution particulière de ce système et de la solution générale du système sans second membre. Un énoncé unifiant ces résultats sera présenté en 7.9.

La méthode de résolution des systèmes linéaires est très semblable à celle du chapitre précédent pour déterminer le rang d'un système de vecteurs. La seule différence notable est que les vecteurs étaient placés verticalement et qu'on faisait apparaître des 0 dans des fins de ligne, alors qu'ici les coefficients sont disposés horizontalement et qu'on fait apparaître des 0 sur des fins de colonnes. L'unification de ces deux méthodes attendra le paragraphe 10.3.

Dernier commentaire. L'idée de la méthode du pivot est finalement simple : utiliser un coefficient d'une équation pour éliminer une inconnue du reste du système et recommencer. Quelques exemples permettront, on l'espère, de comprendre comment s'applique ce discours général.

4.8 EXEMPLES

➤ Résolvons le système L suivant :

$$\begin{cases} x + 3y - z = 9 & l_1 \\ 3x - y + z = -1 & l_2 \\ -2x + y - 3z = 6 & l_3 \end{cases}$$

En choisissant le coefficient de x dans l_1 comme premier pivot, le système est équivalent au système L' :

$$\begin{cases} x + 3y - z = 9 & l_1 \\ -10y + 4z = -28 & l'_2 = l_2 - 3l_1 \\ 7y - 5z = 24 & l'_3 = l_3 + 2l_1 \end{cases}$$

En choisissant le coefficient de y dans l'_2 comme second pivot, le système est équivalent au système L'' :

$$\begin{cases} x + 3y - z = 9 & l_1 \\ -10y + 4z = -28 & l'_2 \\ -\frac{11}{5}z = \frac{22}{5} & l''_3 = l'_3 + \frac{7}{10}l'_2 \end{cases}$$

On trouve $z = -2$ avec l''_3 , puis en reportant cette valeur dans l'_2 , on obtient $y = 2$, enfin, en reportant ces deux valeurs dans l_1 , on obtient $x = 1$, autrement dit $S(L) = \{(1, 2, -2)\}$.

➤ Transformons le système précédent en modifiant la seconde équation :

$$\begin{cases} x + 3y - z = 9 & l_1 \\ 3x + 9y - z = -1 & l_2 \\ -2x + y - 3z = 10 & l_3 \end{cases}$$

Avec le même premier pivot que précédemment, on trouve que ce système est équivalent au système :

$$\begin{cases} x + 3y - z = 9 & l_1 \\ 2z = -28 & l'_2 = l_2 - 3l_1 \\ 7y - 5z = 28 & l'_3 = l_3 + 2l_1 \end{cases}$$

Il suffit alors de permuter les deux dernières lignes pour retrouver un système triangulaire :

$$\left\{ \begin{array}{rcl} x + 3y - z & = & 9 \quad l_1 \\ 7y - 5z & = & 28 \quad l_2'' = l_3' \\ 2z & = & -28 \quad l_3'' = l_2' \end{array} \right.$$

On trouve $S(L) = \{(13, -6, -14)\}$.

➤ Considérons maintenant le système L :

$$\left\{ \begin{array}{rcl} x + 3y - z & = & 9 \quad l_1 \\ 3x + 9y - 3z & = & -1 \quad l_2 \\ -2x + y - 3z & = & 6 \quad l_3 \end{array} \right.$$

Avec le même premier pivot que précédemment, on trouve que ce système est équivalent au système :

$$\left\{ \begin{array}{rcl} x + 3y - z & = & 9 \quad l_1 \\ 0 & = & -28 \quad l_2' = l_2 - 3l_1 \\ 7y - 5z & = & 24 \quad l_3' = l_3 + 2l_1 \end{array} \right.$$

qui est clairement incompatible, donc $S(L) = \emptyset$.

➤ Considérons enfin le système L :

$$\left\{ \begin{array}{rcl} x + 3y - z & = & 9 \quad l_1 \\ 3x + 9y - 3z & = & 27 \quad l_2 \\ -2x + y - 5z & = & 10 \quad l_3 \end{array} \right.$$

Comme $l_2 - 3l_1$ est une équation de la forme $0 = 0$, le système est équivalent au système :

$$\left\{ \begin{array}{rcl} x + 3y - z & = & 9 \quad l_1 \\ 7y - 7z & = & 28 \quad l_3' = l_3 + 2l_1 \end{array} \right.$$

Ce système est équivalent à :

$$\left\{ \begin{array}{rcl} x + 3y & = & 9 + z \\ y & = & 4 + z \end{array} \right.$$

ou encore à :

$$\left\{ \begin{array}{rcl} x & = & -3 - 2z \\ y & = & 4 + z \end{array} \right.$$

Posons $s = (-3, 4, 0)$ et $u_3 = (-2, 1, 1)$. On a $S(L) = s + \text{Vect}(u_3)$.

4.9 SYSTÈMES AVEC PARAMÈTRES

On rencontre des systèmes d'équations linéaires où les coefficients ne sont pas des réels fixés, mais des paramètres réels. Suivant les valeurs de ces paramètres, le rang du système peut varier et l'ensemble des solutions avoir des descriptions différentes. Le but de ces exercices est surtout d'entraîner les étudiant(e)s à organiser leurs calculs et à avoir une démarche rigoureuse qui n'oublie aucun cas.

Du point de vue de leur résolution, les systèmes avec paramètres ne posent pas de problèmes nouveaux. Il convient simplement, dans l'application de la méthode du pivot, de distinguer les cas particuliers qui apparaissent et de les traiter un par un. On essaie en général d'organiser les calculs pour que ces cas apparaissent le plus tard possible, c'est-à-dire en évitant le plus longtemps possible d'utiliser des pivots dépendant des paramètres.

Exemple. Examinons suivant les valeurs de a et de b , le système L :

$$\begin{cases} ax + by + z = 1 & l_1 \\ x + aby + z = b & l_2 \\ x + by + az = 1 & l_3 \end{cases}$$

Pour éviter d'avoir à distinguer les cas $a = 0$ et $a \neq 0$ (on verra plus tard qu'il n'y a pas lieu de les distinguer), permutons les rôles de l_1 et de l_3 :

$$\begin{cases} x + by + az = 1 & l'_1 = l_3 \\ x + aby + z = b & l'_2 = l_2 \\ ax + by + z = 1 & l'_3 = l_1 \end{cases}$$

puis prenons comme premier pivot le coefficient de x dans l'_1 . Le système L est équivalent au système :

$$\begin{cases} x + by + az = 1 & l'_1 \\ b(a-1)y + (1-a)z = b-1 & l''_2 = l'_2 - l'_1 \\ b(1-a)y + (1-a^2)z = 1-a & l''_3 = l'_3 - al'_1 \end{cases}$$

qui est équivalent au système :

$$\begin{cases} x + by + az = 1 & l'_1 \\ b(a-1)y + (1-a)z = b-1 & l''_2 \\ (1-a)(2+a)z = b-a & l'''_3 = l''_3 + l''_2 \end{cases}$$

On examine les cas où le pivot $(1-a)(2+a)$ de l'''_3 est nul. Si $a = 1$ et $b \neq 1$, le système est incompatible et $S(L) = \emptyset$. De même si $a = -2$ et $b \neq -2$.

- Si $a = b = 1$, le système L est équivalent au système formé par la seule équation :

$$x + y + z = 1$$

soit :

$$x = 1 - y - z$$

L'ensemble des solutions est $S(L) = \{(1 - y - z, y, z)\}$, soit $S(L) = (1, 0, 0) + \text{Vect}((-1, 1, 0), (-1, 0, 1))$.

- Si $a = b = -2$, le système L est équivalent au système de deux équations :

$$\begin{cases} x - 2y - 2z = 1 \\ 2y + z = -1 \end{cases}$$

L'ensemble des solutions est $S(L) = \{(-1 - 2y, y, -1 - 2y)\}$, soit $S(L) = (-1, 0, -1) + \text{Vect}((2, -1, 2))$.

Le pivot de l_2'' est nul pour $a = 1$, cas que nous avons déjà étudié, et pour $b = 0$. Il reste donc à étudier le cas $b = 0$, qui conduit à une incompatibilité, et le cas général, qui est un calcul laissé en exercice.

4.10 PROBLÈMES ACTUELS

Le développement de l'informatique durant les cinquantes dernières années a permis de résoudre des systèmes linéaires de plus en plus grands. Ces systèmes se rencontrent dans des problèmes de modélisations très variés : conception d'une aile d'avion, prévisions météorologiques, systèmes mécaniques, etc.

Les mathématiciens doivent donc inventer des méthodes de résolution les plus efficaces possibles. Actuellement, on résout, avec la méthode du pivot de Gauss, des systèmes linéaires de 1 000 équations à 1 000 inconnues, systèmes qui ont de l'ordre de 1 000 000 de coefficients non nuls.

Il y a 60 ans, tout cela était parfaitement impossible et il fallait le génie visionnaire de John (Janos) von Neumann (1903-1957), d'origine hongroise, un des plus grands mathématiciens du xx^e siècle, pour comprendre alors comment les ordinateurs allaient bouleverser les possibilités et les méthodes de calcul, pour étudier l'effet de l'accumulation massive d'erreurs d'arrondis qui perturbent les résultats dans l'algorithme de Gauss, etc.

Systèmes linéaires particuliers. Bien souvent, les systèmes linéaires auxquels conduisent les problèmes pratiques de la physique, de l'industrie, etc. ont des caractéristiques particulières. Ce sont, par exemple, des systèmes où les coefficients de la diagonale (les a_{ii}) et une bande de coefficients autour de cette diagonale sont non

nuls, alors que les coefficients des parties triangulaires inférieures et supérieures restantes du système sont nuls. Pour ces systèmes, il existe des méthodes de résolution spécifiques. On peut atteindre des systèmes de 1 000 000 d'équations à 1 000 000 d'inconnues, voire plus. L'architecture parallèle de certains ordinateurs est aussi à prendre en compte pour le choix d'une méthode.

Nombre d'opérations. Il est nécessaire de compter soigneusement le nombre d'opérations à effectuer pour résoudre un système. Commençons le calcul pour la méthode de Gauss en comptant le nombre d'opérations pour trianguler le système. Plaçons-nous dans le cas de la résolution d'un système de n équations à n inconnues conduisant à une solution unique et reprenons la démarche du paragraphe 4.7. On calcule d'abord l'inverse de a_{11} , ce qui fait un calcul d'inverse. Pour calculer $l'_2 = l_2 - \frac{a_{21}}{a_{11}} l_1$, on effectue la multiplication de cet inverse par a_{21} , ce qui fait une multiplication, puis on calcule les coefficients de $\frac{a_{21}}{a_{11}} l_1$, ce qui fait $n - 1$ multiplications (le coefficient de x_1 ne nécessite aucun calcul). On doit enfin effectuer $n - 1$ additions ou soustractions. Pour calculer $l'_2, \dots, l'_n$, on doit donc faire $n(n - 1)$ multiplications et $(n - 1)^2$ additions ou soustractions. Puis on recommence avec un système de $n - 1$ équations avec $n - 1$ inconnues, etc. Récapitulons : pour trianguler, on devra faire :

- $n - 1$ calculs d'inverses ;
- $n^2 - n + (n - 1)^2 - (n - 1) + \dots = \frac{n(n^2 - 1)}{3}$ multiplications ;
- $(n - 1)^2 + (n - 1) + (n - 2)^2 + (n - 2) + \dots = \frac{n(n^2 - 1)}{3}$ additions ou soustractions.

Ces calculs utilisent les formules $1 + 2 + \dots + n = \frac{n(n + 1)}{2}$ et $1^2 + 2^2 + \dots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$ qu'on démontre par récurrence.

Comme les additions et soustractions coûtent un temps bien moindre que les multiplications, on voit que la triangulation d'un système de n équations à n inconnues par la méthode du pivot de Gauss nécessite de l'ordre de $\frac{n^3}{3}$ multiplications. Il faut encore compter le nombre d'opérations de la seconde phase de l'algorithme (il y en a moins), mais aussi les calculs sur les seconds membres. On arrive finalement à une estimation de la complexité de l'algorithme de l'ordre de $\frac{n^3}{3}$ multiplications (en ne comptant pas les termes en $n^k, k \leq 2$). Pour $n = 1\,000$, cela donne 333 333 000 opé-

rations. Pour $n = 10$, la méthode du pivot de Gauss nécessite donc de l'ordre de 330 multiplications. On verra au chapitre 14 que les formules de Cramer nécessitent de l'ordre de $(n + 2) !$ multiplications ; pour $n = 10$, l'application de ces formules demande donc plus de 479 000 000 multiplications. La méthode du pivot de Gauss leur est donc nettement supérieure.

Choix des pivots. Bien que cela ne soit pas le but de ce livre, ajoutons deux mots sur la difficulté de traiter de grands systèmes linéaires à coefficients réels. Même si la taille des nombres est faible au départ, il y a un grand risque que les calculs explosent : les nombres qui apparaissent deviennent soit extrêmement grands soit extrêmement petits, les erreurs d'arrondi se propagent, et il est parfois difficile de savoir ce que signifie le résultat donné par un ordinateur. Ces problèmes font l'objet de recherches très importantes ; différentes stratégies sont possibles pour choisir les pivots.

Notons que les calculatrices formelles actuelles donnent des résultats exacts mais sans discussion dans le cas de systèmes avec paramètres.

➤ Vers le chapitre 5

Nous avons vu l'intérêt des combinaisons linéaires de fonctions dans le chapitre 1, des combinaisons linéaires de suites dans le chapitre 2, des combinaisons linéaires de vecteurs de $\mathbb{R}^n$ dans le chapitre 3, des combinaisons linéaires d'équations dans ce chapitre. L'analogie est profonde. Pour pouvoir préciser cette analogie, il faut faire un pas vers l'abstraction. Il faut considérer les fonctions du chapitre 1, les suites du chapitre 2 ainsi que les vecteurs du chapitre 3 comme des exemples de la même notion. Nous allons conserver le nom de vecteur à cette notion abstraite, et pour dire comment nous travaillons avec ces vecteurs, nous allons définir la notion d'espace vectoriel.

Exercices

4.1 Résolution de systèmes linéaires

Résoudre les systèmes suivants en utilisant la méthode du pivot de Gauss et en donnant les solutions sous la forme $S(L) = \{s + u \mid u \in \text{Vect}(u_{r+1}, \dots, u_p)\}$. Discuter, s'il y a lieu, suivant les valeurs du paramètre.

$$L_1 \quad \begin{cases} 3x + y = 5 \\ 4x + y = 7 \end{cases} \quad L_2 \quad \begin{cases} 2x + 5y = -4 \end{cases}$$

$$L_3 \begin{cases} 3x + y = 5 \\ 4x + y = 7 \\ 2x + 5y = a \end{cases} \quad L_4 \begin{cases} x + 2y + 3z = 1 \\ 2x + 5y + 8z = 0 \\ 3x + 4y + 10z = 1 \end{cases}$$

$$L_5 \begin{cases} x + 2y + 3z = 6 \\ 2x - 3y - 8z = 19 \\ 3x + y + 4z = -2 \end{cases} \quad L_6 \begin{cases} x + 2y + 3z = 6 \\ 2x - 3y - 8z = 19 \\ 3x + y - z = -2 \end{cases}$$

$$L_7 \begin{cases} x + 2y + 3z + 4t + 5u = 6 \\ x + 3y + 5z + 7t + 4u = 13 \\ x + 4y + 7z + 10t + 3u = 20 \\ x + 2y + 4z + 6t + 8u = 13 \end{cases}$$

$$L_8 \begin{cases} x + z + t = 0 \\ ax + y + (a-1)z + t = 0 \\ x - y + 2z + at = 0 \\ 2x + ay + z + 2t = 0 \end{cases}$$

$$L_9 \begin{cases} x + y + (1-a)z = a+2 \\ (1+a)x - y + 2z = 0 \\ 2x - ay + 3z = a+2 \end{cases}$$

4.2 Vrai, faux

- a)** Un système linéaire de 3 équations à 4 inconnues dont les seconds membres sont nuls a des solutions non nulles.
- b)** Un système linéaire de 4 équations à 3 inconnues dont les seconds membres sont nuls n'a que la solution nulle.

4.3 Appartenance à un « Vect »

En se ramenant à la résolution d'un système linéaire, résoudre les questions suivantes. On comparera avec les méthodes données dans l'exercice 3.6.

- a)** On pose $e_1 = (1, 0)$, $u_1 = (3, 1)$ et $u_2 = (-1, 5)$.

Montrer que $e_1 \in \text{Vect}(u_1, u_2)$ en écrivant explicitement e_1 comme combinaison linéaire de u_1, u_2 .

b) On pose $e_1 = (1, 0, 0)$, $u_1 = (-1, 1, -3)$, $u_2 = (1, 2, 5)$ et $u_3 = (1, 7, 1)$.

Montrer, avec la même méthode, que $e_1 \in \text{Vect}(u_1, u_2, u_3)$.

4.4 Intersection de deux Vect dans $\mathbb{R}^3$

On pose $u_1 = (3, 4, 2)$, $u_2 = (0, 1, -1)$, $v_1 = (-1, 2, 1)$ et $v_2 = (-4, 2, 0)$.

Déterminer les vecteurs appartenant à l'intersection $\text{Vect}(u_1, u_2) \cap \text{Vect}(v_1, v_2)$.

On se ramènera à la résolution d'un système linéaire.

Solutions

4.1 Quand on résout un système linéaire, il est très important de noter que les différents systèmes qu'on est amené à écrire en appliquant la méthode du pivot sont équivalents, c'est-à-dire ont tous le même ensemble de solutions, ce qu'une rédaction détaillée doit toujours mentionner.

Pour être sûr(e), quand on résout un système, de l'équivalence des systèmes qu'on est amené à écrire, on pourra n'utiliser une équation que pour modifier celles qui sont placées en-dessous, comme dans la méthode du pivot. On voit souvent des débutant(e)s faire des opérations dans tous les sens sans aucun contrôle, ce qui peut parfois réussir, mais pas toujours (et ce n'est, de toutes façons, pas rigoureux).

$$S(L_1) = \{(2, -1)\}.$$

$$S(L_2) = s + \text{Vect}(v) \text{ avec } s = (-2, 0) \text{ et } v = \left(-\frac{5}{2}, 1\right).$$

Détaillons l'exemple de la résolution du système L_3 . Ce système s'écrit :

$$\begin{cases} y + 3x = 5 & l_1 \\ y + 4x = 7 & l_2 \\ 5y + 2x = a & l_3 \end{cases}$$

Il est équivalent au système

$$\begin{cases} y + 3x = 5 & l'_1 \\ x = 2 & l'_2 = l_2 - l_1 \\ -13x = a - 25 & l'_3 = l_3 - 5l_1 \end{cases}$$

qui est lui-même équivalent au système formé de l_1, l'_2 et :

$$l''_3 = l'_3 + 13l'_2 : 0 = a + 1.$$

Si $a \neq -1$, on a $S(L_3) = \emptyset$; si $a = -1$, on retrouve $S(L_1)$.

$$S(L_4) = \left\{ \left(\frac{19}{5}, \frac{2}{5}, -\frac{6}{5} \right) \right\}.$$

$$S(L_5) = \{(3, 9, -5)\}.$$

$$S(L_6) = \emptyset.$$

$S(L_7) = s + \text{Vect}(v, w)$ avec $s = (-1, -7, 7, 0, 0)$, $v = (0, 1, -2, 1, 0)$ et $w = (-10, 7, -3, 0, 1)$.

Si $a = 1$, $S(L_8) = \text{Vect}(v, w)$ avec $v = (-1, 1, 1, 0)$ et $w = (-1, 0, 0, 1)$.

Si $a \neq 1$, $S(L_8) = \text{Vect}(v)$ avec $v = (a - 1, -1, -a, 1)$.

$$S(L_9) = \left\{ \left(\frac{1}{a-2}, \frac{-a-3}{a-2}, \frac{-a-2}{a-2} \right) \right\} \text{ sauf :}$$

si $a = 2$, $S(L_9) = \emptyset$;

si $a = 0$, $S(L_9) = (1, 1, 0) + \text{Vect}((-3, 1, 2))$;

si $a = -2$, $S(L_9) = \{(x, -x, 0) \mid x \in \mathbb{R}\} = \text{Vect}((1, -1, 0))$.

4.2 a) Vrai : le système est de rang ≤ 3 .

b) Faux : penser, par exemple, au cas où toutes les équations sont identiques.

4.3 On se ramène à la résolution des systèmes :

$$\text{a) } \begin{cases} 3\lambda - \mu = 1 \\ \lambda + 5\mu = 0 \end{cases}$$

$$\text{b) } \begin{cases} -\lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \lambda_1 + 2\lambda_2 + 7\lambda_3 = 0 \\ -3\lambda_1 + 5\lambda_2 + \lambda_3 = 0 \end{cases}$$

4.4 Un vecteur de l'intersection s'écrit à la fois $xu_1 + yu_2$ et $zv_1 + tv_2$, ce qui donne le système :

$$\begin{cases} 3x = -z - 4t \\ 4x + y = 2z + 2t \\ 2x - y = z \end{cases}$$

La méthode du pivot donne : $x = -\frac{2}{3}t$, $y = \frac{2}{3}t$, $z = -2t$. L'intersection recherchée est $\text{Vect}((1, -1, 3))$.

GÉNÉRALITÉS SUR LES ESPACES VECTORIELS

5

5.1 INTRODUCTION

Nous avons déjà dit que l'algèbre linéaire, c'est-à-dire l'étude des espaces vectoriels et des applications linéaires, présentait de grandes difficultés pour un certain nombre d'étudiant(e)s. Cela tient peut-être à l'effort d'abstraction nécessaire à cette étude. Nous avons essayé de le préparer en présentant dans les chapitres précédents des exemples d'espaces vectoriels dans lesquels toutes celles et tous ceux qui utilisent des techniques d'algèbre linéaire ont l'habitude de travailler.

Nous avons souligné que les chapitres 1, 2 et 3 se ressemblent. Pourquoi donc refaire dans chacun d'eux les mêmes démarches quand on sent que c'est toujours la même chose ? Ne vaut-il pas mieux essayer de dégager ce qui fait cette parenté et créer un cadre dans lequel on pourra énoncer et démontrer des résultats une fois pour toutes ?

Ce qu'il y a de commun entre les différents exemples est ce qu'on appelle une structure d'espace vectoriel. Pour pouvoir la dégager, il faut examiner précisément ce qui est commun aux différents exemples, et choisir des axiomes, aussi simples que possible, pour la décrire.

Cependant, on s'aperçoit que ces nouveaux objets posent à leur tour des problèmes. Ces problèmes ne se posaient pas vraiment au niveau des différents exemples d'espaces vectoriels, mais ils se posent au niveau théorique ; il va falloir se forger une intuition à un autre niveau, celui des espaces vectoriels, pour pouvoir les résoudre.

5.2 UN PEU D'HISTOIRE

À la fin du XIX^e siècle, l'idée de structure en mathématiques est une idée récente qui ne s'est pas encore imposée. Évariste Galois (1811-1832) a sans doute été le premier à en dégager les contours pour les corps et les groupes (voir chapitres 11 et 12) dans son travail sur la résolution des équations algébriques.

En ce qui concerne les espaces vectoriels, c'est Hermann Grassmann (1809-1877), un autodidacte à l'écart de la communauté mathématique de son temps, qui va le premier avoir une approche formaliste des problèmes de la géométrie. Son traité de 1844, très obscur, mêlant différents types de problèmes géométriques, ne fut ni lu ni compris. Grassmann rédigea ses travaux dans une forme plus achevée en 1862, mais sans plus de succès.

C'est Giuseppe Peano (1858-1932) qui, en relisant Grassmann, dégage de son texte de 1862 la notion d'espace vectoriel, qu'il appelle « *sistemi lineari* ». Sa définition est moins précise que celle que nous allons donner en 5.3, mais elle contient les mêmes idées. Peano donne quelques résultats sur les espaces vectoriels et les applications linéaires que nous allons retrouver dans les chapitres qui suivent : définition de la dimension, composée d'applications linéaires, etc. Après lui, quelques mathématiciens de l'école italienne continuent à explorer cette direction, comme Salvatore Pincherle (1853-1936), sans que leurs idées aient beaucoup d'influence.



Giuseppe Peano (1858-1932)

Peano G., *Opere scelte*, vol 1, Rome, Edizioni Cremonese, 1957.

C'est un autre domaine qui va montrer l'importance des notions d'espace vectoriel et d'application linéaire. Dans les années 1900, de nombreuses recherches de l'école allemande étudient les espaces de fonctions, en liaison en particulier avec la théorie des séries de Fourier et d'autres que nous ne pouvons évoquer à ce niveau (le premier à avoir considéré les fonctions comme des objets (voir 1.1) est le mathématicien Vito Volterra (1860-1940), en 1886).

En 1920, Stefan Banach (1892-1945) fait, dans une thèse qui marque le début des travaux de la grande école polonaise de mathématiques de l'entre-deux guerres, la synthèse magistrale des travaux de l'école allemande. Banach écrit :

L'ouvrage présent a pour but d'établir quelques théorèmes valables pour différents champs fonctionnels, que je spécifie dans la suite. Toutefois, afin de ne pas être obligé à les démontrer isolément pour chaque champ particulier, ce qui serait bien pénible, j'ai choisi une voie différente que voici : je considère d'une façon générale les ensembles d'éléments dont je postule certaines propriétés, j'en déduis des théorèmes et je démontre ensuite de chaque champ fonctionnel particulier que les postulats adoptés sont vrais pour lui.

(Stefan Banach, « Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales », *Fundamenta Mathematicae* 3, 1922).

Les motivations pour développer les notions d'algèbre linéaire sont donc données par des espaces de dimension infinie. La synthèse entre les idées de Grassmann et Peano et ces nouvelles idées se fait dans le livre fondateur, en allemand, de Bartel Leendert Van der Waerden (1903-1996) : *Moderne Algebra* (premières éditions en 1930 et 1937). Elle inspirera les célèbres traités de Nicolas Bourbaki. L'importance unificatrice de l'algèbre linéaire, qui permet de traiter de la même manière des problèmes de domaines éloignés, est alors reconnue par tous.

Nous avons essayé de faire percevoir ces aspects de l'algèbre linéaire à travers les premiers chapitres, mais il est impossible de donner encore une idée des problèmes d'analyse fonctionnelle qui ont tant contribué à l'élaboration de tous les concepts de l'algèbre linéaire. Le lecteur et la lectrice doivent être patients. Ce livre ne pourra pas aller assez loin pour les leur montrer.

5.3 STRUCTURE DE $\mathbb{R}$ -ESPACE VECTORIEL

En reprenant le langage utilisé dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, nous appellerons vecteurs les éléments d'un espace vectoriel. Les vecteurs pourront être des fonctions, comme dans le chapitre 1, des suites, comme dans le chapitre 2 ou encore des n -uplets de nombres réels, comme dans le chapitre 3. Nous avons additionné ces éléments et nous les avons multipliés par des nombres, rien d'autre. Ces additions et multiplications ont quelques propriétés qu'il faut mettre en évidence ; c'est ce que nous allons faire.

Définition de $\mathbb{R}$ -espace vectoriel :

1) **Les données.** Un $\mathbb{R}$ -espace vectoriel (on oubliera souvent le $\mathbb{R}$ dans les chapitres suivants) est un triplet $(E, +, \cdot)$ formé d'un ensemble E dont les éléments seront appelés vecteurs, d'une loi d'addition qui est une application $+: E \times E \rightarrow E$ qui associe à deux vecteurs u et v le vecteur $u + v$ qu'on appellera somme des deux vecteurs, et d'une loi de multiplication par un scalaire qui est une application notée par un point : $\mathbb{R} \times E \rightarrow E$ qui associe à un nombre réel λ et à un vecteur u le vecteur λu qu'on appellera produit du vecteur u par le scalaire λ (ou $\lambda \cdot u$).

Ces deux lois vérifient les axiomes ci-dessous.

2) **Les axiomes pour la somme.** Le couple $(E, +)$ est un groupe abélien, ce qui signifie que la somme possède les propriétés suivantes :

➤ **Associativité :**

pour tous u, v, w de E , on a $u + (v + w) = (u + v) + w$

➤ **Commutativité :**

pour tous u, v de E , on a $u + v = v + u$

➤ **Existence d'un élément neutre pour l'addition :**

il existe un élément noté 0 dans E tel que :

pour tout u de E , on a $u + 0 = u$

➤ **Existence d'un élément opposé :**

pour tout u de E , il existe un élément v de E tel que :

$$u + v = 0$$

Conséquences immédiates des axiomes de la somme.

➤ L'associativité permet de ne pas mettre de parenthèses dans des sommes de plusieurs termes ;

➤ la commutativité permet de changer l'ordre des termes dans une somme ;

➤ compte tenu de la commutativité de la somme, on a :

pour tout u de E : $u + 0 = u$ et $0 + u = u$;

➤ l'élément opposé de u est noté $-u$ et la somme $u + (-v)$ est notée $u - v$; on peut donc écrire $u - u = 0$;

➤ enfin, on peut faire des simplifications comme :

si $u + v = u + w$ alors $v = w$,

il suffit d'ajouter $-u$ aux deux membres ; par conséquent, $u + v = u$ implique $v = 0$.

3) Les axiomes de compatibilité pour la multiplication.

➤ Avec le produit des réels :

pour tout u de E et tous λ et μ de $\mathbb{R}$ on a $(\lambda\mu)u = \lambda(\mu u)$

➤ Avec la somme des réels :

pour tout u de E et tous λ et μ de $\mathbb{R}$ on a $(\lambda + \mu)u = \lambda u + \mu u$

➤ Avec la somme des vecteurs :

pour tous u et v de E et tout λ de $\mathbb{R}$ on a $\lambda(u + v) = \lambda u + \lambda v$

➤ Avec l'unité :

pour tout u de E , on a $1u = u$

Conséquences immédiates des axiomes de la multiplication. Ces axiomes impliquent les règles de calcul dont on a l'habitude.

Par exemple :

➤ $0u = 0$ (où le premier 0 désigne le réel 0 et où le second 0 désigne l'élément neutre de E) : en effet, les axiomes montrent que

$$u + 0u = 1u + 0u = (1 + 0)u = 1u = u ;$$

➤ $\lambda 0 = 0$ car $\lambda 0 + \lambda 0 = \lambda(0 + 0) = \lambda 0$;➤ $(-1)u = -u$ car $u + (-1)u = (1 - 1)u = 0u = 0 = u + (-u)$;➤ $2u = u + u$, etc.

Commentaire. Ce début ne doit pas effrayer : dire qu'on se donne un triplet est la meilleure façon de ne pas oublier qu'on se donne trois choses, l'ensemble des éléments et les deux lois d'addition des éléments entre eux et de multiplication des éléments par des réels. En fait, les lois d'addition et de multiplication une fois précisées, tout le monde parle de l'espace vectoriel E sans citer ni $\mathbb{R}$ ni les deux lois ; on dira aussi que l'ensemble E est muni par ces deux lois d'une structure d'espace vectoriel. On abrège aussi parfois l'appellation *espace vectoriel* en *espace*.

5.4 EXEMPLES FONDAMENTAUX

Nous avons donné plusieurs exemples importants de $\mathbb{R}$ -espaces vectoriels au cours des chapitres précédents (nous ne reprécisons pas les lois d'addition des vecteurs et de multiplication par un scalaire) :

➤ Pour tout entier $n \geq 0$, l'espace vectoriel $\mathbb{R}^n$ présenté au chapitre 3.➤ L'espace $\text{Fonct}(I, \mathbb{R})$ des fonctions quelconques d'un intervalle I de $\mathbb{R}$ dans $\mathbb{R}$ (éventuellement $I = \mathbb{R}$) présenté au chapitre 1.

- L'espace des suites de nombres réels présenté au chapitre 2.
- L'espace $\mathbb{R}[X]$ des polynômes à coefficients réels. Le chapitre 13 sera consacré à une construction algébrique des polynômes. Nous n'en n'avons pas besoin pour le moment car, dans le cas de coefficients réels, on peut identifier les polynômes et les fonctions polynômes. Pour nous, un polynôme sera donc une fonction $\mathbb{R} \rightarrow \mathbb{R}$ de la forme $x \mapsto \sum_{0 \leq k \leq n} a_k x^k$. En notant $X : x \mapsto x$ la fonction $\text{id}_{\mathbb{R}}$, la fonction X^k est la fonction $x \mapsto x^k$ et la fonction $x \mapsto \sum_{0 \leq k \leq n} a_k x^k$ est la fonction $\sum_{0 \leq k \leq n} a_k X^k$.
- L'espace $\text{Fonc}(X, E)$ des applications d'un ensemble X dans un espace vectoriel E (voir exercice 5.6).

Pour vérifier que ces exemples conviennent, il faut vérifier à chaque fois les 8 axiomes, ce qui est facile mais un peu long. On dispose souvent d'un moyen très efficace d'abrégé ces vérifications pour d'autres espaces, car ce sont des sous-ensembles de l'un des exemples précédents, et il suffit de vérifier qu'ils en forment des sous-espaces vectoriels, ce qui est en général très simple à vérifier, voir paragraphe 5.6.

5.5 PRÉCISIONS SUR LES CORPS

Dans les premiers chapitres de ce livre, nous nous limitons, à quelques exceptions près, aux espaces vectoriels sur le corps des réels. La raison en est que nous ne souhaitons pas créer un obstacle supplémentaire à la compréhension en parlant d'espaces vectoriels sur d'autres corps. Une fois comprises les notions dans ce cadre, on peut facilement les transposer aux espaces vectoriels sur d'autres corps : les théorèmes et les méthodes sont les mêmes. Pour ceux qui veulent cependant déjà entendre parler de ce point de vue plus général, voici quelques précisions.

Les nombres réels s'additionnent et se multiplient ; les deux lois sont associatives et commutatives ; il existe des éléments 0 et 1 qui jouent le rôle d'éléments neutres pour l'addition et la multiplication ; tout réel a un opposé pour l'addition ; tout réel non nul a un inverse pour la multiplication. Tout cela paraît peut-être clair, mais les nombres réels n'ont été définis que relativement récemment. Jusqu'au milieu du XIX^e siècle, la notion de nombre s'appuie sur la notion de mesure de segments qui vient de la géométrie grecque et ne concerne que les nombres positifs. C'est Richard Dedekind (1831-1916) qui a l'idée, en 1858, de construire les nombres réels à partir des nombres rationnels : un nombre réel est défini en se donnant tous les nombres rationnels qui lui sont inférieurs et tous ceux qui lui sont supérieurs : c'est la construction dite par « coupures ».

Quant aux réels négatifs et à la règle des signes, c'est aussi à cette époque qu'ils ont cessé de tourmenter les mathématiciens. Jusque-là, beaucoup d'entre eux refusaient les nombres négatifs, car on ne peut « ôter quelque chose de rien » (Lazare

Carnot (1753-1823), 1803). Il y avait là une difficulté à distinguer le zéro absolu et le zéro comme origine sur un axe. En fait, les problèmes d'ordre ne relèvent pas de la structure de corps.

On a alors quelques exemples d'objets qui se comportent comme les nombres réels (les nombres complexes, les nombres rationnels, etc.) ou presque comme les nombres réels (les *quaternions* inventés par William Hamilton (1805-1865) en 1843, voir 22.10). Dedekind les appelle des *corps* en 1871 (*Körper* en allemand, d'où la notation K)¹. Les propriétés qui définissent cette structure sont présentées au chapitre 12, où des exemples d'autres corps seront également donnés.

Les résultats que nous allons énoncer pour les espaces vectoriels sur le corps $\mathbb{R}$ restent valables, sans modification aucune, pour les espaces vectoriels sur le corps $\mathbb{C}$ formé par les nombres complexes ou les espaces vectoriels sur le corps $\mathbb{Q}$ formé par les nombres rationnels.

5.6 SOUS-ESPACES VECTORIELS

Essayons de donner une idée de cette notion en regardant ce qui se passe dans l'espace vectoriel $\mathbb{R}^3$. Cet espace contient des droites et des plans vectoriels : ce sont des espaces vectoriels contenus dans $\mathbb{R}^3$ pour lesquels l'addition des vecteurs et la multiplication par des scalaires se font comme dans $\mathbb{R}^3$.

Définition. Soit $(E, +, \cdot)$ un espace vectoriel. Un espace vectoriel $(F, +', \cdot')$ est appelé sous-espace vectoriel de $(E, +, \cdot)$ si $F \subset E$, si la loi d'addition $+' : F \times F \rightarrow F$ est la restriction de la loi $+$ (c'est-à-dire si, pour tout u et tout v de F , on a $u +' v = u + v$) et si la loi $\cdot' : \mathbb{R} \times F \rightarrow F$ est la restriction de la loi $\cdot$ de multiplication par un scalaire de E (c'est-à-dire si, pour tout u de F et tout λ de $\mathbb{R}$, on a $\lambda \cdot' u = \lambda \cdot u$).

Proposition 1 : critère pour les sous-espaces vectoriels. Soit E un espace vectoriel. Un sous-ensemble F de E définit un sous-espace vectoriel de E s'il est non vide et s'il est stable par les restrictions à F de l'addition et de la multiplication par un scalaire de E , autrement dit si :

- pour tous u et v de F : $u + v \in F$,
- pour tout u de F et tout λ de $\mathbb{R}$: $\lambda u \in F$.

Démonstration. Il faut se convaincre que les 8 axiomes de la définition sont vérifiés, ce qui est facile. La condition « F non vide », toujours vérifiée dans la pratique, permet de montrer que $0 \in F$ puisqu'il existe un élément u de F et que $0u = 0 \in F$. □

1. On dit *cuerpo* en espagnol, *field* en anglais...

Commentaire. C'est souvent la vérification de $0 \in F$ qui permet de voir que F est non vide.

Au lieu des deux conditions de la définition précédente, on peut donner une condition unique, toujours en supposant F non vide :

pour tous u et v de F , tous λ et μ de $\mathbb{R}$: $\lambda u + \mu v \in F$.

Proposition 2. Intersection de sous-espaces vectoriels. *Soit E un espace vectoriel. L'intersection d'une famille de sous-espaces vectoriels de E est un sous-espace vectoriel de E .*

Démonstration. Soit $(F_i)_{i \in I}$ une famille de sous-espaces vectoriels de E , u et v des éléments de $G = \bigcap_{i \in I} F_i$, λ et μ des réels. Pour tout i de I , u et v sont des éléments de F_i , donc $\lambda u + \mu v$ est dans F_i . Par conséquent, $\lambda u + \mu v \in G$. $\square$

5.7 EXEMPLES DE SOUS-ESPACES VECTORIELS

La liste des sous-espaces vectoriels des exemples fondamentaux est très longue et beaucoup d'exemples, qui demandent des connaissances plus avancées, ne peuvent être encore présentés. Les lecteurs et lectrices sont invités à vérifier, à l'aide des résultats de 5.6, les résultats suivants en détail.

Sous-espaces de Fonct $(I, \mathbb{R})$.

- L'espace $C^0(I, \mathbb{R})$ des fonctions continues sur I , car la somme de deux fonctions continues est continue, le produit d'une fonction continue par une constante est continue. On voit que des résultats élémentaires du cours d'analyse sont utiles. Nous ne rappellerons pas ces résultats pour les exemples suivants.
- L'espace $C^1(I, \mathbb{R})$ des fonctions continûment dérivables sur I , c'est-à-dire dérivables et de dérivées continues sur I .
- L'espace $C^n(I, \mathbb{R})$ des fonctions n fois continûment dérivables sur I , c'est-à-dire n fois dérivables et de dérivée n -ième continue sur I .
- L'espace des fonctions admettant une primitive sur I .
- L'espace des fonctions bornées sur I .
- L'espace des fonctions f deux fois continûment dérivables sur $\mathbb{R}$ satisfaisant l'équation différentielle :

$$f'' - f = 0$$

Plus généralement, l'espace des fonctions deux fois continûment dérivables sur $\mathbb{R}$ satisfaisant une équation différentielle linéaire du second ordre sans second membre donnée.

Sous-espaces de l'espace des suites.

- L'espace c_0 des suites convergentes.
- L'espace l_∞ des suites bornées.
- L'espace l_1 des suites (u_n) telles que la somme $\sum_{n \geq 0} |u_n|$ soit finie, ce qui signifie que la suite S_N des sommes finies $\sum_{0 \leq n \leq N} |u_n|$ a une limite finie.

Sous-espaces de l'espace des polynômes. L'espace $\mathbb{R}_n[X]$ (attention à la place du n) des polynômes de degré $\leq n$.

Sous-espaces de $\mathbb{R}^n$.

- Le sous-espace engendré par un vecteur non nul est parfois appelé axe.
- Étant donnés des nombres réels $a_1, \dots, a_n$, l'espace des n -uplets $(x_1, \dots, x_n)$ vérifiant l'équation :

$$a_1x_1 + \dots + a_nx_n = 0.$$

Si les a_i ne sont pas tous nuls, cet espace est appelé hyperplan de $\mathbb{R}^n$.

- Plus généralement, on peut définir les sous-espaces vectoriels de $\mathbb{R}^n$ intersection de plusieurs hyperplans ; leurs éléments sont les n -uplets $(x_1, \dots, x_n)$ vérifiant simultanément plusieurs équations homogènes. Ceci nous permet un nouveau regard sur les systèmes linéaires sans second membre.

Proposition. Soit L un système linéaire à coefficients réels de n équations à p inconnues sans second membre :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = 0 \\ \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = 0 \end{cases}$$

L'ensemble $S(L)$ des solutions de L est un sous-espace vectoriel de $\mathbb{R}^p$. □

Espaces d'applications linéaires. Nous verrons dans des chapitres ultérieurs l'importance de l'exemple des espaces vectoriels $L(E, F)$ des applications linéaires d'un espace vectoriel dans un autre, en particulier le dual d'un espace vectoriel E que nous étudierons au chapitre 10.

5.8 COMBINAISONS LINÉAIRES ET ESPACE ENGENDRÉ

Nous allons reprendre dans le cadre général des espaces vectoriels les notions déjà abordées à plusieurs reprises.

Soit donc E un espace vectoriel.

Définition 1 : combinaison linéaire. Étant donnés des vecteurs $u_1, \dots, u_p$ de E , nous appellerons combinaison linéaire à coefficients réels de $u_1, \dots, u_p$ tout vecteur de la forme

$$\lambda_1 u_1 + \dots + \lambda_p u_p = \sum_{1 \leq i \leq p} \lambda_i u_i$$

avec $\lambda_1, \dots, \lambda_p$ réels.

Une combinaison linéaire est toujours une somme finie.

Si F est un sous-espace vectoriel d'un espace vectoriel E , F est stable par combinaison linéaire.

Définition 2 : espace engendré. L'ensemble des combinaisons linéaires à coefficients réels de $u_1, \dots, u_p$ sera noté $\text{Vect}(u_1, \dots, u_p)$ et appelé espace engendré par $u_1, \dots, u_p$.

Il ne dépend pas de l'ordre dans lequel sont donnés les vecteurs $u_1, \dots, u_p$.

L'espace engendré par un vecteur u non nul est appelé droite vectorielle engendrée par u et $\text{Vect}(u) = \{\lambda u, \lambda \in \mathbb{R}\}$.

Plus généralement, si A est une partie non vide, finie ou infinie de E , $\text{Vect}(A)$ note l'ensemble des combinaisons linéaires d'éléments de A et est appelé espace engendré par A , ce que justifie la proposition suivante.

Proposition 1 : espace Vect (A). $\text{Vect}(A)$ est un sous-espace vectoriel de E . C'est le plus petit sous-espace vectoriel de E contenant A .

Démonstration. Il est clair que $\text{Vect}(A)$ contient A . Que ce soit un sous-espace vectoriel de E se montre avec le critère de la proposition 1 de 5.6.

D'autre part, si F est un sous-espace vectoriel de E contenant A , alors F contient toute combinaison linéaire d'éléments de A , donc contient $\text{Vect}(A)$. $\text{Vect}(A)$ est donc le plus petit sous-espace vectoriel de E contenant A . $\square$

Commentaires. Cette proposition justifie qu'on pose $\text{Vect}(\emptyset) = \{0\}$ puisque $\{0\}$ est le plus petit sous-espace vectoriel de E .

Elle permet aussi de construire des sous-espaces vectoriels d'une nouvelle manière, en donnant l'ensemble des vecteurs qui les engendrent. Par exemple, on pourra parler du sous-espace vectoriel de $\mathbb{R}^4$ engendré par les vecteurs $(1, 2, -1, 3)$ et $(5, 8, -2, 0)$.

On peut aussi remarquer que $\text{Vect}(A)$ est l'intersection de la famille des sous-espaces vectoriels de E contenant A .

Les propriétés de l'espace engendré ont déjà été énoncées en 3.4 dans le cas des espaces engendrés par des vecteurs de $\mathbb{R}^n$. La principale est la suivante que nous avons déjà démontrée et commentée en 3.4.

Proposition 2. Soient $p \geq 2$ un entier et $u_1, \dots, u_p$ des vecteurs de E .

1) Soit $\lambda \neq 0$ un réel. Alors :

$$\text{Vect}(u_1, u_2, \dots, u_p) = \text{Vect}(\lambda u_1, u_2, \dots, u_p).$$

2) Soit $u'_1 = u_1 + \sum_{2 \leq k \leq p} \lambda_k u_k$. Alors :

$$\text{Vect}(u_1, u_2, \dots, u_p) = \text{Vect}(u'_1, u_2, \dots, u_p).$$

Commentaire. Cette proposition généralise la proposition 2 de 3.4 et l'on pourra s'en souvenir de la même manière : on ne change pas l'espace engendré par un ensemble de vecteurs quand on change un des vecteurs en lui ajoutant une combinaison linéaire des autres vecteurs.

Définition 3 : vecteurs colinéaires. On dit que des vecteurs u et v sont colinéaires si $u \in \text{Vect}(v)$ ou $v \in \text{Vect}(u)$, autrement dit s'il existe un réel λ tel que $u = \lambda v$ ou tel que $v = \lambda u$. Le vecteur 0 est colinéaire à tout vecteur. Comme en 3.4, proposition 1, on a le critère suivant.

Proposition 3 : critère de colinéarité. Deux vecteurs u et v sont colinéaires si et seulement s'il existe des réels λ et μ non nuls tels que :

$$\lambda u + \mu v = 0.$$

5.9 SOMME DE SOUS-ESPACES

Soient E un espace vectoriel, F et G des sous-espaces vectoriels de E .

On appelle somme de F et G le sous-espace $\text{Vect}(F \cup G)$ engendré par $F \cup G$. On le note $F + G$. Notons que l'union $F \cup G$ n'est un espace vectoriel que dans des conditions particulières (voir l'exercice 5.4.d).

On vérifie facilement que $F + G = \{u + v \mid u \in F, v \in G\}$.

On définit de même la somme d'un nombre fini de sous-espaces $E_1, \dots, E_n$ de E comme le sous-espace engendré par $E_1 \cup \dots \cup E_n$; on le note $E_1 + \dots + E_n$.

On vérifie que $E_1 + \dots + E_n = \{u_1 + \dots + u_n \mid u_1 \in E_1, \dots, u_n \in E_n\}$.

➤ Vers le chapitre 6

Il faut maintenant se donner les moyens de faire des calculs dans un espace vectoriel. C'est le but du chapitre suivant, qui introduit la notion de base d'un espace vectoriel.

Exercices

Les premiers exercices de ce chapitre ont pour but de se familiariser avec la notion de sous-espace vectoriel dans des espaces classiques.

5.1 Sous-espaces vectoriels de $\mathbb{R}^2$

Reprendre l'exercice 3.2 en déterminant, parmi les ensembles $A_1, \dots, A_7$, ceux qui sont des sous-espaces vectoriels.

5.2 Sous-espaces vectoriels de $\text{Fonct}(\mathbb{R}, \mathbb{R})$

Parmi les sous-ensembles suivants de $\text{Fonct}(\mathbb{R}, \mathbb{R})$, lesquels forment des sous-espaces vectoriels ?

- $F_1 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ est paire}\}$
- $F_2 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ est impaire}\}$
- $F_3 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid \forall x \in \mathbb{R}, f(-x) = f(x) + 2\}$
- $F_4 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid \forall x \in \mathbb{R}, f(x) \geq 0\}$
- $F_5 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid \forall x \in \mathbb{Z}, f(x) = 0\}$
- $F_6 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid \forall x \in \mathbb{Z}, f(x) = 1\}$
- $F_7 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f(1) = 0\}$
- $F_8 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ est bornée}\}$
- $F_9 = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ deux fois dérivable et } (1 - x^2)f'' - 2xf' + 6f = 0\}$
- $F_{10} = \{f: \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ continue et } \int_0^1 f(x) \sin x \, dx = 0\}$

5.3 Sous-espaces vectoriels de $\mathbb{R}[X]$

Parmi les sous-ensembles suivants de $\mathbb{R}[X]$, lesquels forment des sous-espaces vectoriels ?

- $G_1 = \{P \in \mathbb{R}[X] \mid \deg P = 7\}$
- $G_2 = \{P \in \mathbb{R}[X] \mid \deg P \leq 3\}$
- $G_3 = \{P \in \mathbb{R}[X] \mid P \text{ est pair}\}$
- $G_4 = \{P \in \mathbb{R}[X] \mid P \text{ n'a pas de terme constant}\}$
- $G_5 = \{P \in \mathbb{R}[X] \mid P(1) = P(2) = P(3) = 0\}$

5.4 Petits exercices théoriques

Soit E un espace vectoriel.

- a) Soient u et v dans E tels que $u + v = u$. Que peut-on conclure ?
- b) Soit u dans E et λ réel tels que $\lambda u = 0$. Que peut-on conclure ?
- c) Vérifier que E et $\{0\}$ sont des sous-espaces vectoriels de E .
- d) Soient F et G des sous-espaces vectoriels de E . Montrer que $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$. Dessiner des exemples et des contre-exemples dans $\mathbb{R}^3$.
- e) Soient A et B des parties de E . On suppose $A \subset B$. Montrer que $\text{Vect}(A) \subset \text{Vect}(B)$.
- f) Soit A une partie de E . Montrer que $\text{Vect}(\text{Vect}(A)) = \text{Vect}(A)$.
- g) Soient A une partie de E et F un sous-espace vectoriel de E . On suppose $\text{Vect}(A) \cap F$ non réduit à 0 ; $A \cap F$ est-il nécessairement non vide ?
- h) Soient A et B des parties de E . Montrer que $\text{Vect}(A \cup B) = \text{Vect}(A) + \text{Vect}(B)$.
- i) Soient F et G des sous-espaces vectoriels de E . Montrer que $F + G = G$ équivaut à $F \subset G$. Dessiner des exemples et des contre exemples dans $\mathbb{R}^3$.

5.5 Calculs sur les Vect

- a) On définit les fonctions $f_1 : x \mapsto e^x$, $f_2 : x \mapsto \text{sh } x$, $f_3 : x \mapsto \text{ch } x$ et $f_4 : x \mapsto e^{-x}$. Donner des présentations simples de l'espace engendré par ces quatre fonctions dans l'espace $\text{Fonct}(\mathbb{R}, \mathbb{R})$.
- b) Donner une présentation simple de l'espace F engendré par les polynômes $X - 1$, $X^2 + 3X - 11$, $X^3 - X^2$, $X^3 + 2X^2$, $X + 1$ dans $\mathbb{R}[X]$.

5.6 Espace vectoriel $\text{Fonct}(X, E)$

Soient E un espace vectoriel et X un ensemble. On définit la somme de deux fonctions $X \rightarrow E$, le produit d'une fonction $X \rightarrow E$ par un réel, comme au chapitre 1. Montrer qu'on définit sur l'ensemble $\text{Fonct}(X, E)$ des fonctions de X dans E une structure d'espace vectoriel.

Solutions

5.1 Les ensembles qui forment des sous-espaces vectoriels de $\mathbb{R}^2$ sont ceux qui sont stables par combinaisons linéaires, c'est-à-dire A_1 et A_4 .

On peut aussi dire que A_1 est un sous-espace vectoriel puisque ses éléments sont les solutions d'une équation linéaire (cf. 5.7).

Notons enfin qu'un sous-ensemble qui ne contient pas 0, comme A_2 , ne forme sûrement pas un sous-espace vectoriel.

5.2 On sait que $\text{Fonct}(\mathbb{R}, \mathbb{R})$ est un espace vectoriel. Pour montrer qu'un sous-ensemble de cet espace forme un sous-espace vectoriel, on utilise les résultats de 5.6. Pour montrer le contraire, un seul contre-exemple suffit.

- F_1 est un sous-espace vectoriel de $\text{Fonct}(\mathbb{R}, \mathbb{R})$ car :
 - F_1 est non vide : il contient, par exemple, la fonction $x \mapsto 0$;
 - la somme de deux fonctions paires f et g est une fonction paire puisque $(f+g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f+g)(x)$;
 - le produit d'une fonction paire par un réel λ est une fonction paire puisque : $(\lambda f)(-x) = \lambda f(-x) = \lambda f(x) = (\lambda f)(x)$.
- On voit, par une démonstration analogue à la précédente, que F_2 est un sous-espace vectoriel de $\text{Fonct}(\mathbb{R}, \mathbb{R})$.
- F_3, F_6 ne sont pas des sous-espaces vectoriels de $\text{Fonct}(\mathbb{R}, \mathbb{R})$ car ils ne contiennent pas la fonction nulle $x \mapsto 0$.
- F_4 n'est pas un sous-espace vectoriel de $\text{Fonct}(\mathbb{R}, \mathbb{R})$ car l'opposé de la fonction $x \mapsto 1$, qui est un élément de F_4 , est la fonction $x \mapsto -1$ qui n'est pas dans F_4 .
- Les autres F_i sont tous des sous-espaces vectoriels de $\text{Fonct}(\mathbb{R}, \mathbb{R})$. Utiliser les propriétés de la dérivation, de l'intégrale, etc., pour le vérifier.

5.3 • G_1 n'est pas un sous-espace vectoriel de $\mathbb{R}[X]$ car il ne contient pas le polynôme nul qui n'est pas de degré 7.

- G_2, G_3, G_4, G_5 sont des sous-espaces vectoriels de $\mathbb{R}[X]$.
- Pour G_3 , on peut aussi remarquer que c'est l'intersection de deux sous-espaces vectoriels de $\text{Fonct}(\mathbb{R}, \mathbb{R})$: $G_3 = \mathbb{R}[X] \cap F_1$.

5.4 a) En ajoutant $-u$ aux deux membres, on voit que $v = 0$.

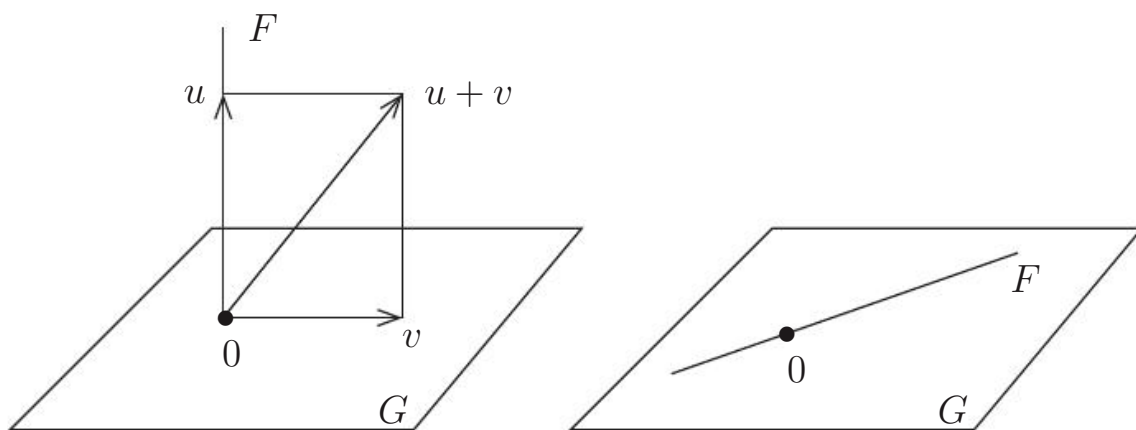
b) On peut avoir $\lambda = 0$ ou $u = 0$. Si l'un des deux est non nul, l'autre est nul.

c) Facile.

d) Si $F \subset G$, alors $F \cup G = G$ est un sous-espace vectoriel de E . Si $G \subset F$, alors $F \cup G = F$ est un sous-espace vectoriel de E . Si aucune de ces deux conditions n'est remplie, il existe un vecteur u tel que $u \in F$ et $u \notin G$ et un vecteur v tel que $v \in G$ et $v \notin F$; si $F \cup G$ était un sous-espace vectoriel, on aurait $w = u + v \in F \cup G$; si $w \in F$ alors $v = w - u$ doit appartenir à F , ce qui est faux; si $w \in G$ alors $u = w - v$ doit appartenir à G , ce qui est faux.

Le premier dessin donne un exemple de situation où $F \cup G$ n'est pas un sous-espace vectoriel, avec un vecteur u de F et un vecteur v de G dont la somme n'est pas dans $F \cup G$; le second dessin illustre le cas $F \subset G$.

e) Vect (B) est un sous-espace vectoriel qui contient B donc A . Il contient donc Vect (A) ,



qui est le plus petit sous-espace vectoriel de E contenant A .

f) Tout sous-espace vectoriel contenant A contient Vect (A) et réciproquement. Donc Vect $(\text{Vect } (A))$ et Vect (A) sont les intersections des mêmes familles de sous-espaces vectoriels et sont donc égaux. On peut aussi utiliser le fait que, si F est un sous-espace vectoriel de E , Vect $(F) = F$.

g) Prenons un exemple dans $\mathbb{R}^3$ muni de la base canonique e_1, e_2, e_3 . Posons $A = \{e_1, e_2\}$ et $F = \text{Vect } (e_1 + e_2)$. Il est clair que $A \cap F = \emptyset$ et que Vect $(A) \cap F = F$. Si on avait pris $F = \text{Vect } (e_1)$, la conclusion aurait été différente. On ne peut donc rien conclure des hypothèses.

h) Par définition, Vect $(A) + \text{Vect } (B) = \text{Vect } (\text{Vect } (A) \cup \text{Vect } (B))$. Un sous-espace vectoriel de E qui contient Vect $(A) \cup \text{Vect } (B)$ contient $A \cup B$. Réciproquement, un

Chapitre 5 • Généralités sur les espaces vectoriels

sous-espace vectoriel de E qui contient $A \cup B$ contient A et B donc contient $\text{Vect}(A)$, $\text{Vect}(B)$ donc $\text{Vect}(A) \cup \text{Vect}(B)$. La famille des sous-espaces contenant $\text{Vect}(A) \cup \text{Vect}(B)$ est donc égale à la famille des sous-espaces contenant $A \cup B$; d'où le résultat.

i) Si $F + G = G$, on a $\text{Vect}(F \cup G) \subset G$, donc $F \subset G$. La réciproque est facile. Les dessins du d) présentés plus haut illustrent les situations $F + G \supset G$ et $F + G = G$.

5.5 a) $\text{Vect}(f_1, f_2, f_3, f_4) = \text{Vect}(f_1, f_4) = \text{Vect}(f_2, f_3)$.

b) $\text{Vect}(1, X, X^2, X^3)$.

5.6 La vérification des axiomes un à un ne doit pas poser de problèmes.

6.1 INTRODUCTION

On a vu au chapitre 3 que tout vecteur de $\mathbb{R}^n$ s'exprime d'une façon unique comme combinaison linéaire des n vecteurs de la base canonique de $\mathbb{R}^n$. Cette propriété a un double avantage : d'une part, on peut exprimer n'importe quel vecteur en fonction de ces n vecteurs particuliers. D'autre part, on peut comparer du premier coup d'œil deux vecteurs u et v de $\mathbb{R}^n$: comme chacun d'eux ne s'écrit que d'une seule façon comme combinaison linéaire des vecteurs de la base canonique : $u = \sum_{1 \leq i \leq n} x_i e_i$, $v = \sum_{1 \leq i \leq n} y_i e_i$, on a $u = v$ si $x_i = y_i$ pour tout i de $1, \dots, n$ et $u \neq v$ s'il existe un i (un seul suffit) tel que $x_i \neq y_i$.

C'est pour permettre de calculer dans les espaces vectoriels que nous allons définir en toute généralité la notion de base.

6.2 FAMILLE GÉNÉRATRICE

La première propriété de la base canonique de $\mathbb{R}^n$ que nous cherchons à étendre aux espaces vectoriels est que tout vecteur de $\mathbb{R}^n$ s'exprime comme combinaison linéaire des vecteurs de la base canonique.

Soit donc E un espace vectoriel.

Définition. Une famille finie $(u_1, \dots, u_p)$ de vecteurs de E telle que $\text{Vect}(u_1, \dots, u_p) = E$ est appelée famille génératrice de E . Autrement dit, la famille $(u_1, \dots, u_p)$ est une famille génératrice de E si tout vecteur de E est une combinaison linéaire de vecteurs de cette famille.

Plus généralement, on dira qu'une famille $(x_i)_{i \in I}$ de vecteurs de E est une famille génératrice de E si tout vecteur de E est une combinaison linéaire finie de vecteurs de cette famille.

Remarque. Nous avons déjà indiqué que la distinction entre ensemble et famille de vecteurs n'était pas toujours importante. C'est le cas ici : l'espace engendré dépend de l'ensemble des vecteurs de la famille.

Espace de dimension finie. On dit que E est de dimension finie s'il existe une famille génératrice finie d'éléments de E .

Si E est de dimension finie et non réduit à $\{0\}$, il existe donc $u_1, \dots, u_p$ dans E tels que $\text{Vect}(u_1, \dots, u_p) = E$.

Exemples.

- 1) Une famille contenant les éléments d'une famille génératrice est une famille génératrice.
- 2) Dans $\mathbb{R}^n$, une famille de vecteurs de rang n est une famille génératrice (cf. 3.10).
- 3) L'espace $\mathbb{R}_n[X]$ des polynômes de degré $\leq n$ est un espace vectoriel de dimension finie : il est engendré par les polynômes $1, X, \dots, X^n$, mais l'espace $\mathbb{R}[X]$ n'est pas de dimension finie. En effet, soit une famille finie $P_1, \dots, P_p$ de polynômes de $\mathbb{R}[X]$. Il existe un entier N supérieur aux degrés de ces p polynômes. Donc $\text{Vect}(P_1, \dots, P_p) \subset \mathbb{R}_N[X]$, ce qui prouve qu'une famille finie de polynômes ne peut engendrer $\mathbb{R}[X]$.
- 4) Au chapitre 1, on a décrit l'ensemble des solutions d'une équation différentielle linéaire du second ordre homogène à coefficients constants comme l'espace engendré par deux fonctions f_1 et f_2 . On peut donc dire que la famille (f_1, f_2) est une famille génératrice de l'espace des solutions de l'équation différentielle.

6.3 FAMILLE LIBRE

Cherchons maintenant à étendre la propriété d'unicité de l'écriture dans la base canonique de $\mathbb{R}^n$.

Définition. On dit qu'une famille $(u_1, \dots, u_p)$ de vecteurs de E est une famille libre de E si tout vecteur de $\text{Vect}(u_1, \dots, u_p)$ s'exprime d'une manière unique comme combinaison linéaire de $u_1, \dots, u_p$.

On dit que la famille est liée dans le cas contraire ; le 2) de la proposition 1 ci-dessous va justifier ce vocabulaire.

Si la famille $(u_1, \dots, u_p)$ est libre, on dit aussi que les vecteurs $u_1, \dots, u_p$ sont linéairement indépendants. Si elle est liée, on dit que les vecteurs sont linéairement dépendants.

On peut généraliser cette définition à une famille infinie de vecteurs en disant qu'elle est libre si toutes ses sous-familles finies sont libres.

Notons que :

- 0 ne peut être élément d'une famille libre puisque 0 peut s'écrire $\lambda 0$ pour tout réel λ ;
- un même élément ne peut apparaître deux fois dans une famille libre ;
- la famille vide est libre ;
- une famille extraite d'une famille libre est libre ;
- l'ordre sur les vecteurs défini par la numérotation $1, \dots, n$ n'est pas important. Toute famille déduite d'une famille libre par permutation des vecteurs est encore libre.

Exemple. La famille de vecteurs de $\mathbb{R}^2$ définie par $e_1 = (1, 0)$, $e_2 = (0, 1)$ et $u = e_1 + e_2$ n'est visiblement pas libre : par exemple, $(2, -1)$ peut s'écrire $2e_1 - e_2$, $2u - 3e_2$, etc.

Proposition 1. Soit $(u_1, \dots, u_p)$ une famille finie de vecteurs d'un espace vectoriel E .

1) La famille est libre si et seulement si la seule écriture de 0 comme combinaison linéaire des u_i est $0u_1 + \dots + 0u_p$.

Cette condition s'énonce encore : pour tous scalaires $\lambda_1, \dots, \lambda_p$, l'égalité $0 = \lambda_1 u_1 + \dots + \lambda_p u_p$ implique $\lambda_1 = \dots = \lambda_p = 0$.

2) Si la famille est liée, il existe des scalaires $\lambda_1, \dots, \lambda_p$, non tous nuls, tels que $\sum_{1 \leq i \leq p} \lambda_i u_i = 0$.

3) Si la famille est liée, l'un des vecteurs est combinaison linéaire des autres.

Démonstration. 1) Dans un sens, c'est la définition de famille libre.

Pour démontrer la réciproque, supposons qu'il existe un vecteur v de $\text{Vect}(u_1, \dots, u_p)$ s'écrivant de deux façons différentes comme combinaison linéaire de $u_1, \dots, u_p$, d'une part $v = \sum_{1 \leq i \leq p} \lambda_i u_i$, d'autre part $v = \sum_{1 \leq i \leq p} \mu_i u_i$. On a donc : $0 = \sum_{1 \leq i \leq p} (\lambda_i - \mu_i) u_i$, donc $\lambda_i = \mu_i$ pour tout i , $1 \leq i \leq p$.

Le 2) est une conséquence immédiate du 1).

3) Le 2) montre qu'on peut trouver des réels non tous nuls $\lambda_1, \dots, \lambda_p$ tels que $\sum_{1 \leq i \leq p} \lambda_i u_i = 0$. Supposons, par exemple, que $\lambda_1 \neq 0$.

On a alors $u_1 = -\frac{\lambda_2}{\lambda_1}u_2 - \dots - \frac{\lambda_p}{\lambda_1}u_p$. □

Exemple.

Montrons que les vecteurs $u_1 = (3, 1, 2, 0)$, $u_2 = (0, 0, -1, 1)$ et $u_3 = (0, 0, 0, 2)$ sont linéairement indépendants dans $\mathbb{R}^4$. Il faut montrer que 0 ne peut s'écrire que d'une seule façon comme combinaison linéaire de u_1 , u_2 et u_3 . Supposons que $0 = \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3$, ce qui revient au système linéaire :

$$\begin{cases} 0 = & 3\lambda_1 \\ 0 = & \lambda_1 \\ 0 = & 2\lambda_1 - \lambda_2 \\ 0 = & \lambda_2 + 2\lambda_3 \end{cases}$$

On voit immédiatement que la seule possibilité est $\lambda_1 = \lambda_2 = \lambda_3 = 0$, ce qui prouve l'unicité de l'écriture de 0 comme combinaison linéaire de u_1 , u_2 et u_3 . Par conséquent, la famille (u_1, u_2, u_3) est libre dans $\mathbb{R}^4$. Généralisons ce dernier exemple.

Proposition 2. Dans $\mathbb{R}^n$, une famille de vecteurs $(u_1, \dots, u_p)$ triangulaire par rapport à la base canonique et sans vecteur nul est une famille libre. Il en est de même par conséquent d'une famille échelonnée sans vecteur nul.

Démonstration. Le système d'équations donné par l'égalité $0 = \lambda_1 u_1 + \dots + \lambda_p u_p$ a une forme triangulaire analogue à celle de l'exemple ci-dessus ; la seule possibilité est $\lambda_1 = \dots = \lambda_p = 0$. □

Retour sur l'algorithme du pivot de Gauss. Dans $\mathbb{R}^n$, étant donnée une famille de vecteurs $(u_1, \dots, u_p)$, la méthode du pivot de Gauss permet de construire (cf. chapitre 3) une famille échelonnée de vecteurs $(v_1, \dots, v_r)$ qui engendre le même espace et qui ne contient pas de vecteur nul. La famille $(v_1, \dots, v_r)$ est donc libre.

6.4 BASE D'UN ESPACE VECTORIEL

Nous sommes maintenant en mesure de définir la notion de « base » dans un espace vectoriel.

Bases. On dit qu'une famille de vecteurs B d'un espace vectoriel E est une base de E si c'est une famille libre et génératrice de E , c'est-à-dire si tout vecteur de E s'écrit d'une manière et d'une seule comme combinaison linéaire des vecteurs de la base B .

Commentaire. Cette notion est très importante ; comme nous l'avons déjà dit, elle permet de calculer commodément dans les espaces vectoriels. Encore faut-il en montrer l'existence. Nous allons le faire pour les espaces de dimension finie, la démonstration générale pour les espaces de dimension infinie faisant appel à l'axiome de choix (voir exercice 19.14).

Théorème 1. Soit E un espace vectoriel de dimension finie. On peut extraire de toute famille génératrice C de E une base B de E .

Démonstration. Puisque E est de dimension finie, il possède une famille génératrice finie D . Il existe une sous-famille finie C' de C engendrant les vecteurs de D ; cette famille engendre E .

- Si elle est vide, c'est que $E = \{0\}$, on pose $B = \emptyset$.
- Si elle est non vide, notons-la $(u_1, \dots, u_p)$. Considérons toutes les sous-familles de cette famille. Certaines engendrent E , d'autres pas. Choisissons-en une qui engendre E et ait le plus petit nombre n d'éléments possible ; on a $n \leq p$. Cette famille B , qu'on peut supposer formée des vecteurs $(u_1, \dots, u_n)$, est libre. En effet, si ce n'est pas le cas, l'un de ses éléments, disons u_1 , est une combinaison linéaire des autres et on a $E = \text{Vect}(u_1, \dots, u_n) = \text{Vect}(u_2, \dots, u_n)$; la famille $(u_2, \dots, u_n)$ engendre E et comporte $n - 1$ éléments, ce qui contredit la définition de n . Donc B est une base de E . $\square$

Théorème de la base incomplète. Soit E un espace vectoriel de dimension finie. Toute famille libre $C = (u_1, \dots, u_p)$ de vecteurs de E peut être complétée en une base $B = (u_1, \dots, u_p, \dots, u_n)$ de E .

Démonstration. Si la famille C engendre E , on pose $B = C$. Sinon, puisque E est de dimension finie, il existe une famille finie $(v_1, \dots, v_q)$ engendrant E . On va construire une suite de familles libres $C_0, \dots, C_q$ par récurrence.

On pose $C_0 = C$. Si on a construit des familles libres $C_1, \dots, C_k$, avec $k < q$ et $C_k = (u_1, \dots, u_p, u_{p+1}, \dots, u_s)$, on construit C_{k+1} :

- si $v_{k+1} \in \text{Vect}(C_k)$, on pose $C_{k+1} = C_k$;
- si $v_{k+1} \notin \text{Vect}(C_k)$, on pose $u_{s+1} = v_{k+1}$ et $C_{k+1} = (u_1, \dots, u_s, u_{s+1})$.

Dans ce second cas, la famille C_{k+1} est libre, car si $0 = \sum_{1 \leq i \leq s+1} \lambda_i u_i$, on voit d'abord que $\lambda_{s+1} = 0$, puis l'hypothèse de récurrence montre que $\lambda_i = 0$ pour $1 \leq i \leq s$.

De plus, C_q engendre E car $E = \text{Vect}(v_1, \dots, v_q) \subset \text{Vect}(C_q)$. Donc C_q est une base de E . $\square$

Coordonnées d'un vecteur dans une base. Soient E un espace vectoriel de dimension finie et $B = (e_1, \dots, e_n)$ une base de E . Tout vecteur u de E s'écrit de manière unique $u = x_1 e_1 + \dots + x_n e_n$ (autrement dit, il existe des réels $x_1, \dots, x_n$ uniques tels que $u = x_1 e_1 + \dots + x_n e_n$). Ces réels sont appelés les coordonnées de u dans la base B .

Ordre des vecteurs. L'ordre sur les vecteurs défini par la numérotation $1, \dots, n$ est important pour les calculs. Toute permutation des vecteurs change l'ordre des coordonnées. Par exemple, le vecteur qui s'écrit $(2, -1, 1)$ dans la base canonique (e_1, e_2, e_3) de $\mathbb{R}^3$ s'écrit $(-1, 1, 2)$ dans la base (e_2, e_3, e_1) et $(-1, 2, 1)$ dans la base (e_2, e_1, e_3) . Il s'agit de trois écritures, dans trois bases différentes, du même vecteur.

Cependant, l'ordre n'est pas toujours important : si B est une base, toute famille déduite de B par permutation des vecteurs est encore une base. Nous formerons à plusieurs reprises une base en prenant l'union des vecteurs de deux familles, sans préciser l'ordre mis sur cette union.

6.5 DIMENSION

Cette section aboutit à un résultat essentiel : toutes les bases d'un espace vectoriel de dimension finie ont le même nombre d'éléments, nombre que nous appellerons dimension de l'espace vectoriel. Pour démontrer ce théorème, nous allons utiliser l'algorithme du chapitre 3.

Proposition 1. Soient E un espace vectoriel de dimension finie et $B = (e_1, \dots, e_n)$ une base de E . Soit $m > n$. Toute famille de m vecteurs de E est liée.

Démonstration. Soit $u_1, \dots, u_m$ une famille de m vecteurs. Notons $x_{11}, \dots, x_{n1}$ les coordonnées de u_1 dans la base B , $x_{12}, \dots, x_{n2}$ les coordonnées de u_2 dans la base B , etc. La méthode du pivot (voir 3.10) peut s'appliquer. On part d'un tableau de m colonnes et de n lignes et on fait apparaître des 0 successivement dans la partie triangulaire supérieure du tableau. Comme $m > n$, l'une des colonnes au moins devient nulle à une étape de l'algorithme. Supposons que ce soit la k -ième ; alors u_k appartient à $\text{Vect}(u_1, \dots, u_{k-1})$, ce qui prouve que la famille $u_1, \dots, u_m$ est liée (on a supposé, pour simplifier l'écriture, ne pas avoir permuté les vecteurs $u_1, \dots, u_m$). $\square$

Théorème et définition. Dans un espace vectoriel E de dimension finie, toutes les bases ont le même nombre d'éléments appelé dimension de l'espace vectoriel E et noté $\dim(E)$.

Démonstration. Soient deux bases $B = (u_1, \dots, u_m)$ et $C = (v_1, \dots, v_n)$ de E .

Supposons $m > n$. Comme C est une base, la proposition 1 ci-dessus montre que B est une famille liée, ce qui est faux.

On montre de même que $m < n$ est impossible. Par conséquent, $m = n$. $\square$

Quand on travaille dans un espace vectoriel E de dimension finie n , plusieurs problèmes peuvent se poser. Le premier, c'est de déterminer une base de E . Mais, souvent, le problème n'est pas celui-là : on connaît déjà une base de E , ou au moins la dimension de E , et on cherche à montrer qu'une famille de n vecteurs est une autre base de E . On dispose du résultat suivant pour éviter « la moitié » des vérifications.

Proposition 2. *Dans un espace vectoriel E de dimension n :*

- 1) *Toute famille libre $(u_1, \dots, u_n)$ de n vecteurs de E est une base.*
- 2) *Toute famille génératrice $(u_1, \dots, u_n)$ de n vecteurs est une base.*

Démonstration. 1) La famille $(u_1, \dots, u_n)$ peut être complétée en une base $(u_1, \dots, u_{n+m})$ de E d'après le théorème de la base incomplète. Comme toutes les bases de E ont le même nombre d'éléments, c'est que $m = 0$, ce qui prouve que $(u_1, \dots, u_n)$ est une base de E .

- 2) La famille $(u_1, \dots, u_n)$ contient, d'après le théorème 1 de 6.4 une base de E . Comme toutes les bases ont le même nombre d'éléments, c'est que cette base est $(u_1, \dots, u_n)$. □

Proposition 3 : dimension d'un sous-espace.

- 1) *Un sous-espace F d'un espace vectoriel E de dimension finie est de dimension finie.*
- 2) *On a alors $\dim(F) \leq \dim(E)$.*
- 3) *Si $\dim(F) = \dim(E)$, on a $F = E$.*

Démonstration. 1) Cet énoncé semble naturel mais sa démonstration n'est pas évidente, car il n'y a aucune raison pour que des vecteurs d'une famille engendrant E soient dans F .

Supposons que E soit de dimension n . Toute famille libre finie de vecteurs de F est une famille libre finie de vecteurs de E . En tant que telle, elle peut être complétée en une base de E d'après le théorème de la base incomplète. Elle a donc au plus n éléments d'après le théorème ci-dessus. Choisissons parmi les familles libres finies de vecteurs de F une famille $(u_1, \dots, u_p)$ ayant le nombre maximal d'éléments possible. On a $p \leq n$. Si cette famille n'engendre pas F , il existe un vecteur u de F n'appartenant pas à $\text{Vect}(u_1, \dots, u_p)$. Alors $(u_1, \dots, u_p, u)$ est une famille libre de $p + 1$ vecteurs de F , contradiction ! Donc $(u_1, \dots, u_p)$ engendre F . Par conséquent, c'est une base de F .

- 2) Cela résulte de $\dim(F) = p \leq n$.
- 3) Si $p = n$ et si F est strictement contenu dans E , il existe un vecteur u de E non dans F et $(u_1, \dots, u_p, u)$ est une famille libre de E ; d'où $\dim(E) \geq n + 1$, ce qui est faux. □

La formule suivante est due à Grasmann, dans son traité de 1862.

Proposition 4 : dimension de la somme. Soit E un espace vectoriel de dimension finie et soient F et G deux sous-espaces vectoriels de E . On a :

$$\dim (F + G) = \dim (F) + \dim (G) - \dim (F \cap G).$$

Démonstration. On sait que $F, G, F \cap G$ sont de dimension finie d'après la proposition 3. Soit $B = (u_1, \dots, u_k)$ une base de $F \cap G$. On peut la compléter, d'une part, en une base $C = (u_1, \dots, u_k, v_1, \dots, v_s)$ de F , et, d'autre part, en une base $D = (u_1, \dots, u_k, w_1, \dots, w_t)$ de G .

Montrons que $C \cup D$ est une base de $F + G$ ayant $k + s + t$ éléments. Comme un vecteur de $F + G$ est somme d'un vecteur de F et d'un vecteur de G , on voit que $C \cup D$ est une famille génératrice de $F + G$. Il reste à montrer que $(u_1, \dots, u_k, v_1, \dots, v_s, w_1, \dots, w_t)$ est une famille libre de $F + G$. Si :

$$0 = \lambda_1 u_1 + \dots + \lambda_k u_k + \lambda_{k+1} v_1 + \dots + \lambda_{k+s} v_s + \lambda_{k+s+1} w_1 + \dots + \lambda_{k+s+t} w_t$$

on a une égalité entre un vecteur de F : $\lambda_1 u_1 + \dots + \lambda_k u_k + \lambda_{k+1} v_1 + \dots + \lambda_{k+s} v_s$ et un vecteur de G : $-(\lambda_{k+s+1} w_1 + \dots + \lambda_{k+s+t} w_t)$. Ces deux vecteurs sont donc dans $F \cap G$, ce qui implique $0 = \lambda_{k+1} = \dots = \lambda_{k+s} = \lambda_{k+s+1} = \dots = \lambda_{k+s+t}$; on en déduit que les autres coefficients sont aussi nuls. Donc, $\dim (F + G) = k + s + t$, $\dim (F) = k + s$, $\dim (G) = k + t$ et $\dim (F \cap G) = k$, d'où le résultat. $\square$

On peut généraliser cette formule au cas de plusieurs sous-espaces.

Définition : hyperplan. Un hyperplan d'un espace E de dimension n est un sous-espace vectoriel H de dimension $n - 1$.

Si F est un sous-espace vectoriel de E contenant un hyperplan H , alors $F = H$ ou $F = E$, d'après la proposition 3.

6.6 EXEMPLES DE BASES

Dans un problème d'algèbre linéaire, il existe en général une ou des bases mieux adaptées pour faire les calculs. Le choix de la bonne base est donc important, comme en géométrie le choix d'un repère.

Base de $\{0\}$. $B = \emptyset$ convient, donc $\dim (\{0\}) = 0$.

Familles triangulaires par rapport à une base. Dans un espace vectoriel, on dispose en général d'une base B , et on a besoin de savoir si des vecteurs dont on connaît les coordonnées par rapport à B forment eux-mêmes une base. On peut généraliser à cette situation la notion de famille triangulaire par rapport à la base canonique de $\mathbb{R}^n$ (cf. 3.6).

Définition. Soient E un espace vectoriel de dimension n et $B = (e_1, \dots, e_n)$ une base de E . Une famille $u_1, \dots, u_p$ de vecteurs de E est triangulaire par rapport à la base B si les coordonnées des vecteurs u_i dans la base B se présentent ainsi :

$$\begin{array}{ccccccc}
 u_1 & u_2 & \dots & u_r & u_{r+1} & \dots & u_p \\
 x_{1,1} & 0 & \dots & 0 & 0 & \dots & 0 \\
 x_{2,1} & x_{2,2} & \dots & 0 & \vdots & \dots & \vdots \\
 \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 x_{r1} & x_{r2} & \dots & x_{rr} & \vdots & \dots & \vdots \\
 \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
 x_{n1} & x_{n2} & \dots & x_{nr} & 0 & \dots & 0
 \end{array}$$

où les $x_{i,i}$ sont non nuls. Il peut arriver que les colonnes de 0 à droite n'existent pas.

On définit de même (voir 3.6) une famille de vecteurs échelonnée par rapport à la base B .

Proposition. Soit E un espace vectoriel et soit B une base de E .

- 1) Une famille de vecteurs de E triangulaire ou échelonnée par rapport à la base B et sans vecteurs nuls est une famille libre de E .
- 2) Si $\dim(E) = n$, une famille triangulaire de n vecteurs de E sans vecteurs nuls est une base de E .

Démonstration. 1) Le raisonnement est le même que dans la proposition 2 de 6.3.

2) Comme la famille est libre et a n éléments, la proposition 2 de 6.5 montre que c'est une base. $\square$

Bases de $\mathbb{R}^n$. On a vu que $\mathbb{R}^n$ admet pour base la base canonique $(e_1, \dots, e_n)$. On a donc $\dim(\mathbb{R}^n) = n$.

La proposition sur les familles triangulaires permet d'affirmer, par exemple, que les vecteurs $e_1, e_1 + e_2, \dots, e_1 + \dots + e_n$ forment une base de $\mathbb{R}^n$.

Espaces de fonctions. Dans le chapitre 1, nous avons vu que l'ensemble $S(E)$ des solutions d'une équation linéaire E du second ordre sans second membre à coefficients constants était un espace vectoriel engendré par deux fonctions f_1 et f_2 linéairement indépendantes. Donc (f_1, f_2) forme une base de $S(E)$ et $\dim(S(E)) = 2$.

Les espaces vectoriels de fonctions utiles en analyse sont de dimension infinie et on cherche souvent dans ces espaces des familles libres A de vecteurs telles que les fonctions qui nous intéressent puissent être approchées aussi près qu'on le souhaite (en un sens à définir) par des éléments de $\text{Vect}(A)$, c'est-à-dire par des combinaisons linéaires d'éléments de A . Soulignons que ces combinaisons linéaires sont, par définition, des sommes finies. Par exemple, en physique, les fonctions décrivant les

phénomènes sont souvent approchées par des sommes partielles de séries de Fourier, c'est-à-dire des sommes finies de la forme :

$$x \mapsto a_0 + a_1 \sin wx + b_1 \cos wx + \cdots + a_n \sin nwx + b_n \cos nwx$$

On montrera en exercice que les fonctions $x \mapsto 1$, $x \mapsto \sin wx$, $x \mapsto \cos wx$, $x \mapsto \sin 2wx$, $x \mapsto \cos 2wx$, etc. forment une famille libre (infinie), sans pouvoir dans ce chapitre aborder plus ce sujet.

Bases de $\mathbb{R}_n[X]$. La famille de polynômes $(1, X, \dots, X^n)$ forme une base de $\mathbb{R}_n[X]$, donc $\dim(\mathbb{R}_n[X]) = n + 1$.

La proposition précédente permet d'affirmer que toute famille de $n + 1$ polynômes de $\mathbb{R}_n[X]$ de degrés tous différents forme une base de $\mathbb{R}_n[X]$ car on peut la voir comme une famille triangulaire par rapport à la base $(1, X, \dots, X^n)$.

D'autres bases intéressantes de $\mathbb{R}_n[X]$ sont présentées en exercice.

6.7 RETOUR AU RANG

Revenons à la méthode du pivot de Gauss pour déterminer le rang d'une famille de vecteurs présentée dans le chapitre 3 pour des vecteurs de $\mathbb{R}^n$ dont on connaît les coordonnées par rapport à la base canonique. On peut faire de même pour une famille de vecteurs dans un espace vectoriel E de dimension finie quand on connaît les coordonnées de ces vecteurs par rapport à une base B de E .

Étant donnée une famille $(u_1, \dots, u_p)$ de vecteurs de E , la méthode du pivot de Gauss permet de déterminer une famille triangulaire ou échelonnée $(v_1, v_2, \dots, v_r)$ de vecteurs non nuls de E telle que $\text{Vect}(v_1, v_2, \dots, v_r) = \text{Vect}(u_1, \dots, u_p)$. Nous avons dit en 3.10 que la valeur de r ne dépendait pas de la façon dont on appliquait la méthode du pivot de Gauss. Nous pouvons maintenant en donner la raison. Comme la famille $(v_1, v_2, \dots, v_r)$ est triangulaire ou échelonnée et sans vecteur nul, c'est une famille libre dans E . La dimension de l'espace qu'elle engendre est donc r . On a donc $r = \dim(\text{Vect}(u_1, \dots, u_p))$, ce qui explique que r a une valeur bien déterminée.

➤ Vers le chapitre 7

Le décor commence à se mettre en place : la notion d'espace vectoriel et celle de base pour effectuer des calculs. Quand on définit une structure, on doit toujours s'intéresser aux applications compatibles avec cette structure. Nous allons donc définir les applications compatibles avec la structure d'espace vectoriel : les applications linéaires.

6.1 Vrai-faux

Les affirmations suivantes sont-elles vraies ou fausses ?

- a) Toute famille génératrice contient une base.
- b) La dimension d'un espace est le nombre de vecteurs de cet espace.
- c) Toute famille contenant une famille liée est liée.
- d) La base de $\mathbb{R}_3[X]$ est $(1, X, X^2, X^3)$.
- e) Si $E = \text{Vect}(u_1, u_2, u_3)$ et si (u_1, u_2, u_3) est une famille libre, alors $\dim(E) = 3$.
- f) $\text{Vect}(u_1, \dots, u_p) = \text{Vect}(u_1, \dots, u_{p-1})$ si et seulement si u_p est combinaison linéaire de $u_1, \dots, u_{p-1}$.
- g) Soient u, v, w trois vecteurs d'un espace vectoriel E . On suppose que deux quelconques de ces trois vecteurs ne sont pas colinéaires. Alors la famille (u, v, w) est libre.
- h) Soient p vecteurs $u_1, \dots, u_p$ d'un espace vectoriel E . On suppose qu'aucun de ces vecteurs n'est combinaison linéaire des autres. Alors la famille $(u_1, \dots, u_p)$ est libre.

6.2 Problèmes de familles libres et de familles génératrices.

Soit E un espace vectoriel.

- 1) Soit $(u_1, \dots, u_4)$ une famille libre de E .
 - a) On suppose $\dim(E) = n$. Quelle inégalité vérifie n ?
 - b) Les familles suivantes sont-elles libres :
 - i) $(u_1, u_2, 0, u_4)$;
 - ii) $(u_1, u_2 + u_3 + u_4, u_4)$;
 - iii) $(u_1 + u_2, u_3 + u_4)$;
 - iv) $(u_1 + u_2, u_2 + u_3, u_3 + u_4, u_4 + u_1)$.
- 2) Soit $u_1, \dots, u_4$ une famille génératrice de E .
 - a) On suppose $\dim(E) = n$. Quelle inégalité vérifie n ?
 - b) Les familles suivantes sont-elles génératrices :
 - i) $(u_1, u_2, u_3, 0, u_4)$;
 - ii) $(u_1, u_1 + u_2, u_3 + u_4, u_4)$;
 - iii) $(u_1 + u_2, u_3 + u_4)$;
 - iv) $(u_1 + u_2, u_2 + u_3, u_3 + u_4, u_4 + u_1)$.

6.3 Petits calculs des familles dans les $\mathbb{R}^n$.

Déterminer si les familles de vecteurs suivantes sont génératrices, libres, liées, forment des bases.

1) Dans $\mathbb{R}$, la famille (u_1, u_2) avec $u_1 = 1$ et $u_2 = -1$;

2) Dans $\mathbb{R}^2$, les familles (u_1, u_2) avec :

a) $u_1 = (1, 2)$ et $u_2 = (4, -1)$;

b) $u_1 = (1, -2)$ et $u_2 = (-4, 8)$.

Si la famille est une base, calculer les coordonnées de $e_1 = (1, 0)$ et de $e_2 = (0, 1)$ dans cette base.

3) Dans $\mathbb{R}^3$, les familles (u_1, u_2, u_3) avec :

a) $u_1 = (1, 2, 3)$, $u_2 = (4, 0, 1)$ et $u_3 = (3, 7, 9)$;

b) $u_1 = (1, 2, 3)$, $u_2 = (4, 0, 1)$ et $u_3 = (-1, 14, 19)$.

Si la famille est une base, calculer les coordonnées de $e_1 = (1, 0, 0)$ et de $e_2 = (0, 1, 0)$ dans cette base.

4) Dans $\mathbb{R}^4$, la famille (u_1, u_2, u_3, u_4) avec $u_1 = (1, 2, -1, -2)$, $u_2 = (2, 3, 0, 1)$, $u_3 = (1, 3, -1, 0)$ et $u_4 = (1, 2, 1, 4)$. Si la famille est une base, calculer les coordonnées de $v = (2, 9, -10, -13)$ dans cette base.

5) Familles de Vandermonde

a) Soient a, b deux réels distincts. Montrer que $(1, a)$ et $(1, b)$ forment une base de $\mathbb{R}^2$.

b) Soient a, b, c trois réels distincts. Montrer que $(1, a, a^2), (1, b, b^2), (1, c, c^2)$ forment une base de $\mathbb{R}^3$.

c) Soient a, b, c, d quatre réels distincts. Montrer que $(1, a, a^2, a^3), (1, b, b^2, b^3), (1, c, c^2, c^3), (1, d, d^2, d^3)$ forment une base de $\mathbb{R}^4$.

On peut généraliser au cas de n nombres réels.

6.4 Rédactions diverses

Un texte d'examen demandait de montrer qu'une famille (u_1, u_2, u_3) de vecteurs était libre. Que pensez-vous des débuts de rédactions suivants, tous authentiques ?

a) La famille (u_1, u_2, u_3) est libre si et seulement si $0 = au_1 + bu_2 + cu_3$ avec $a = b = c = 0$.

b) La famille (u_1, u_2, u_3) est libre si et seulement s'il existe une relation triviale $au_1 + bu_2 + cu_3 = 0$ avec $a = b = c = 0$.

c) Si la famille (u_1, u_2, u_3) est libre alors la combinaison linéaire $au_1 + bu_2 + cu_3 = 0$ est telle que $a = b = c = 0$.

- d)** Pour que la famille (u_1, u_2, u_3) soit libre il faut qu'il existe a, b, c tels que $au_1 + bu_2 + cu_3 = 0$.
- e)** Pour montrer que la famille (u_1, u_2, u_3) est libre, il faut montrer que $0 = au_1 + bu_2 + cu_3$ implique $a = b = c = 0$.
- f)** (u_1, u_2, u_3) est libre équivaut à $0 = au_1 + bu_2 + cu_3$.
- g)** La relation $0 = au_1 + bu_2 + cu_3$ est vraie si $a = b = c = 0$ donc la famille est libre.

6.5 Indépendance dans les espaces de fonctions

Déterminer si les familles de fonctions suivantes sont libres ou liées. On pourra utiliser des propriétés de parité, de dérivabilité, de comportement à l'infini, etc.

- a)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, la famille (f_1, f_2, f_3) , avec $f_1 : x \mapsto 1$, $f_2 : x \mapsto \cos^2 x$, $f_3 : x \mapsto \sin^2 x$.
- b)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, la famille (f_1, f_2, f_3) avec $f_1 : x \mapsto e^x$, $f_2 : x \mapsto e^{2x}$, $f_3 : x \mapsto e^{3x}$.
- c)** Dans $\text{Fonct}([0, \infty[, \mathbb{R})$, la famille (f_1, f_2, f_3) avec $f_1 : x \mapsto \ln x$, $f_2 : x \mapsto \ln^2 x$, $f_3 : x \mapsto \ln^3 x$.
- d)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, la famille (f_1, f_2, f_3) avec $f_1 : x \mapsto e^x$, $f_2 : x \mapsto e^{-x}$, $f_3 : x \mapsto \text{ch } x$.
- e)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, la famille infinie $(f_n)_{n \in \mathbb{N}}$ des fonctions $f_n : x \mapsto e^{nx}$.
- f)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, la famille (f_1, f_2, f_3) avec $f_1 : x \mapsto |x - 1|$, $f_2 : x \mapsto |x - 2|$, $f_3 : x \mapsto |x - 3|$; généraliser à la famille infinie de fonctions $(f_n)_{n \in \mathbb{N}}$ avec $f_n : x \mapsto |x - n|$.
- g)** Dans $\text{Fonct}(\mathbb{R}, \mathbb{R})$, avec $n \geq 1$ entier, les familles :
- $F_n = (f_k)_{1 \leq k \leq n}$ avec $f_k : x \mapsto \sin kx$;
 - $G_n = (g_k)_{0 \leq k \leq n}$ avec $g_k : x \mapsto \cos kx$
 - $H_n = (f_0, \dots, f_n, g_1, \dots, g_n)$;
- h)** Dans $\text{Fonct}(-2, 2[, \mathbb{R})$, la famille (f_1, f_2, f_3) , avec $f_1 : x \mapsto \frac{1}{x-2}$,
 $f_2 : x \mapsto \frac{1}{x+2}$, $f_3 : x \mapsto \frac{4x+7}{x^2-4}$.

6.6 Sommes et intersections de sous-espaces

a) Dans $\mathbb{R}^5$, on définit les sous-espaces vectoriels $F = \text{Vect}(u_1, u_2, u_3)$ et

$G = \text{Vect}(v_1, v_2, v_3)$ en posant $u_1 = (1, 5, -1, 3, -4)$, $u_2 = (1, 7, -1, 5, -4)$, $u_3 = (2, 6, 0, 4, -3)$, $v_1 = (-1, 3, -2, -1, 1)$, $v_2 = (2, -4, 3, 0, 0)$, $v_3 = (0, 4, -1, 0, 2)$. Déterminer $F \cap G$.

b) Soient F, G, H trois sous-espaces vectoriels d'un espace vectoriel E . Comparer $F \cap (G + H)$ et $(F \cap G) + (F \cap H)$.

6.7 Indépendance dans $\mathbb{R}[X]$

Montrer que les familles de polynômes suivantes sont des bases des espaces indiqués.

a) Dans $\mathbb{R}_4[X]$, la famille $(1, X - 1, (X - 1)^2, (X - 1)^3, (X - 1)^4)$; déterminer l'écriture d'un polynôme $P \in \mathbb{R}_4[X]$ dans cette base en fonction des valeurs de P et de ses dérivées successives en 1.

b) Dans $\mathbb{R}_2[X]$, la famille $(X, X(X + 1), (X + 1)^2)$; déterminer l'écriture des polynômes $1, X, X^2$ dans cette base, puis l'écriture d'un polynôme quelconque $a + bX + cX^2$.

c) Dans $\mathbb{R}_4[X]$, la famille $(P_0, \dots, P_4)$ égale à :

$$\left(1, X, \frac{X(X-1)}{2}, \frac{X(X-1)(X-2)}{6}, \frac{X(X-1)(X-2)(X-3)}{24} \right);$$

déterminer l'écriture des polynômes $1 + X + X^2 + X^3$ et $1 + X + X^2 + X^3 + X^4$ dans cette base.

d) Généraliser le **c)** au cas de $E = \mathbb{R}_n[X]$ pour n entier. Montrer que tout polynôme P de la base B trouvée est tel que $P(\mathbb{Z}) \subset \mathbb{Z}$ (autrement dit, $P(a)$ est entier pour tout entier a). Montrer que tout polynôme P de E tel que $P(\mathbb{Z}) \subset \mathbb{Z}$ est une combinaison linéaire à coefficients entiers des vecteurs de la base B .

6.8 Polynômes de Lagrange

Soit n un entier, $E = \mathbb{R}_n[X]$ et $a_1, \dots, a_{n+1}$ des nombres réels deux à deux distincts.

Pour $i = 1, \dots, n + 1$, on définit le polynôme P_i par $P_i(X) = \frac{p_i(X)}{A_i}$ avec $p_i(X) = \prod_{1 \leq j \leq n+1, j \neq i} (X - a_j)$ et $A_i = p_i(a_i)$, autrement dit :

$$P_i(X) = \frac{\prod_{1 \leq j \leq n+1, j \neq i} (X - a_j)}{\prod_{1 \leq j \leq n+1, j \neq i} (a_i - a_j)}.$$

a) Montrer que $P_1, \dots, P_{n+1}$ est une base de E .

b) Soient $b_1, \dots, b_{n+1}$ des nombres réels quelconques. Déterminer, à l'aide des vecteurs de la base précédente, un polynôme P de degré n tel que $P(a_i) = b_i$ pour $i = 1, \dots, n+1$. Le polynôme P est appelé polynôme d'interpolation de Lagrange.

c) Déterminer un polynôme $P \in \mathbb{R}_2[X]$ tel que $P(0) = 3$, $P(1) = -2$, $P(2) = 5$. Y a-t-il unicité du résultat ?

Solutions

6.1 a) Vrai : c'est le premier théorème de 6.4.

b) Faux : les deux notions sont bien distinctes. Un espace de dimension finie non nulle sur le corps $\mathbb{R}$ a une infinité de vecteurs.

c) Vrai : la relation non triviale qui lie des vecteurs de la plus petite des deux familles est vraie dans la plus grande.

d) Incorrect : on ne peut parler de « la » base de $\mathbb{R}_3[X]$ car il y en a « beaucoup » d'autres, ne fut-ce que celles formées par les multiples réels non nuls des vecteurs de cette base.

e) Vrai : la famille (u_1, u_2, u_3) est une base de E car c'est une famille libre et génératrice de E .

f) Vrai. Si u_p est combinaison linéaire de $u_1, \dots, u_{p-1}$, alors $\text{Vect}(u_1, \dots, u_p) = \text{Vect}(u_1, \dots, u_{p-1})$ d'après la proposition 2 de 5.8. Réciproquement, on a $u_p \in \text{Vect}(u_1, \dots, u_{p-1})$, ce qui signifie que u_p est combinaison linéaire de $u_1, \dots, u_{p-1}$.

g) Faux. Par exemple, si on pose $u = e_1$, $v = e_2$ et $w = e_1 + e_2$, l'hypothèse est vérifiée mais u, v et w sont liés par la relation : $w = u + v$: ils sont coplanaires. Cependant, si u, v et w sont les vecteurs de la base canonique (e_1, e_2, e_3) de $\mathbb{R}^3$, ils vérifient l'hypothèse et forment une famille libre.

h) Vrai. Supposons que $0 = \lambda_1 u_1 + \dots + \lambda_p u_p$ et que l'un des coefficients, par exemple λ_1 ne soit pas nul. Alors $u_1 = -\frac{\lambda_2}{\lambda_1} u_2 - \dots$, ce qui contredit l'hypothèse.

6.2 1) a) On a $n \geq 4$.

b) i) Non, car elle contient 0 ;

ii, iii) oui ;

On peut dans les deux cas faire un raisonnement direct ou voir ces familles comme des familles échelonnées par rapport à la base (u_1, u_2, u_3, u_4) de $\text{Vect}(u_1, u_2, u_3, u_4)$;

iv) Non, car $(u_1 + u_2) - (u_2 + u_3) + (u_3 + u_4) - (u_4 + u_1) = 0$.

2) a) $n \leq 4$.

- b) i) Oui, car elle contient la famille (u_1, u_2, u_3, u_4) .
 ii) Oui, car $\text{Vect}(u_1, u_1 + u_2, u_3 + u_4, u_4) = \text{Vect}(u_1, u_2, u_3 + u_4, u_4) = \text{Vect}(u_1, u_2, u_3, u_4)$.
 iii, iv) On ne peut rien dire : on peut donner des exemples où elles engendrent E (par exemple, si les 4 vecteurs sont colinéaires) et d'autres où elles ne l'engendrent pas.

6.3 1) La famille est génératrice mais pas libre car u_1 et u_2 sont liés par la relation $u_2 = -u_1$; comme la dimension de l'espace vectoriel $\mathbb{R}$ est 1, on ne pourra jamais trouver de famille libre de plus d'un vecteur.

2) a) Pour montrer que la première est libre, on dispose du critère de colinéarité (proposition 3.8) et de deux méthodes générales.

- i) D'abord, appliquer la méthode du pivot de Gauss à la famille (u_1, u_2) ; on pose $u'_2 = u_2 - 4u_1$:

$$\begin{array}{cc} u_1 & u'_2 \\ 1 & 0 \\ 2 & -9 \end{array}$$

La famille (u_1, u'_2) est triangulaire par rapport à la base canonique, sans vecteurs nuls, et son nombre de vecteurs est égal à la dimension de l'espace ; c'est une base de $\mathbb{R}^2$. Comme $\text{Vect}(u_1, u_2) = \text{Vect}(u_1, u'_2)$, la famille (u_1, u_2) engendre aussi $\mathbb{R}^2$; son nombre de vecteurs est égal à la dimension de l'espace, donc c'est une base de $\mathbb{R}^2$.

- ii) Montrer d'abord que la famille est libre : on écrit que $0 = \lambda_1 u_1 + \lambda_2 u_2$, ce qui conduit au système :

$$\begin{cases} \lambda_1 + 4\lambda_2 = 0 \\ 2\lambda_1 - \lambda_2 = 0 \end{cases}$$

La résolution de ce système montre que λ_1 et λ_2 sont nécessairement nuls. Par conséquent, la famille (u_1, u_2) est libre ; comme son nombre de vecteurs est égal à la dimension de l'espace, c'est une base de $\mathbb{R}^2$.

Pour calculer les coordonnées de $e_1 = (1, 0)$ et de $e_2 = (0, 1)$ dans la base (u_1, u_2) , chacune des deux méthodes précédentes peut donner le résultat. Avec la première méthode, on voit immédiatement que $e_2 = -\frac{1}{9}u'_2 = \frac{4}{9}u_1 - \frac{1}{9}u_2$, puis que $e_1 = u_1 - 2e_2 = \frac{1}{9}u_1 + \frac{2}{9}u_2$.

La seconde méthode conduit à la résolution des deux systèmes

$$\begin{cases} \lambda_1 + 4\lambda_2 = 1 \\ 2\lambda_1 - \lambda_2 = 0 \end{cases} \quad \begin{cases} \lambda_1 + 4\lambda_2 = 0 \\ 2\lambda_1 - \lambda_2 = 1 \end{cases}$$

b) La seconde famille est liée car $u_2 = -4u_1$.

Les autres questions de l'exercice 6.3 se traitent avec les mêmes arguments, mais les calculs sont souvent plus longs.

3) La méthode du pivot conduit à poser $u'_2 = u_2 - 4u_1 = (0, -8, -11)$ dans les deux cas ; dans le premier cas $u'_3 = u_3 - 3u_1 = (0, 1, 0)$; dans la base (e_1, e_3, e_2) , la famille (u_1, u'_2, u'_3) est une famille triangulaire de trois vecteurs non nuls dans un espace de dimension 3 : c'est donc une base. On a $e_2 = u'_3 = u_3 - 3u_1$. Pour exprimer e_1 , on reprend l'algorithme du pivot dans la base (e_1, e_3, e_2) . On pose successivement :

$$e'_1 = e_1 - u_1, e''_1 = e'_1 - \frac{3}{11}u'_2;$$

	u_1	u'_2	u'_3	e'_1	u_1	u'_2	u'_3	e''_1
e_1	1	0	0	0	1	0	0	0
e_3	3	-11	0	-3	3	-11	0	0
e_2	2	-8	1	-2	2	-8	1	$\frac{2}{11}$

on trouve $e''_1 = \frac{2}{11}u'_3$ et, en revenant aux vecteurs initiaux, on trouve après calculs

$$e_1 = \frac{1}{11}(-7u_1 + 3u_2 + 2u_3); \text{ on trouve aussi : } e_3 = \frac{1}{11}(28u_1 - u_2 - 8u_3).$$

Dans le second cas, on trouve : $u'_3 = u_3 + u_1 = -2u'_2$: la famille est liée par la relation $7u_1 - 2u_2 - u_3 = 0$.

4) On vérifie que la famille est une base et on trouve

$$v = -9u_1 + 5u_2 + 10u_3 - 9u_4.$$

5) Appliquer la méthode du pivot en divisant les vecteurs qui apparaissent par $b - a$, $c - a$, $d - a$, etc. pour simplifier les calculs.

6.4 Seule la formulation e) est correcte. Lisez tout de même les autres phrases pour éviter d'écrire la même chose ou d'autres semblables : la plupart semblent dire que la liberté se définit par $0 = 0u_1 + 0u_2 + 0u_3$, ce qui est toujours vrai ! La formulation c) commence par supposer ce qu'il faut démontrer.

6.5 Rappelons que 0 désigne souvent la fonction nulle $x \mapsto 0$ dans cet exercice.

a) Famille liée : $f_2 + f_3 - f_1 = 0$ car, pour tout réel x , on a $\cos^2 x + \sin^2 x = 1$.

b) Famille libre : si $0 = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$, on a, pour tout réel x :

$0 = \lambda_1 e^x + \lambda_2 e^{2x} + \lambda_3 e^{3x}$; en divisant par e^{3x} et en faisant tendre x vers $+\infty$, on voit que $\lambda_3 = 0$; on a donc, pour tout réel x : $0 = \lambda_1 e^x + \lambda_2 e^{2x}$; en divisant par e^{2x} et en faisant tendre x vers $+\infty$, on voit que $\lambda_2 = 0$; on en déduit $\lambda_1 = 0$. On peut également raisonner en prenant des limites en $-\infty$.

c) Famille libre : même méthode que ci-dessus.

d) Famille liée : on sait que $f_1 + f_2 = 2f_3$.

e) Famille libre : comme il s'agit d'une famille infinie, il faut vérifier que toute sous-famille finie est libre ; comme toute sous-famille finie est contenue dans une famille de la forme $(f_0, \dots, f_n)$ on peut se limiter à ces familles. Supposons que $0 = \lambda_0 f_0 + \dots + \lambda_n f_n$; si tous les coefficients ne sont pas nuls, il existe $k \leq n$ tel que $\lambda_k \neq 0$ et $\lambda_{k+1} = \dots = \lambda_n = 0$; en divisant par e^{kx} et en faisant tendre x vers $+\infty$, on obtient $\lambda_k = 0$: contradiction.

f) Famille libre : la fonction nulle est dérivable pour tout x ; si $0 = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ et si $\lambda_1 \neq 0$, le membre de droite n'est pas dérivable pour $x = 1$, donc $\lambda_1 = 0$. Le même argument permet de montrer que λ_2, λ_3 sont nuls, etc.

g) Pour montrer que F_n est une famille libre, on peut raisonner par récurrence sur n . Le cas $n = 1$ est clair. Supposons que F_n soit libre pour un entier n et montrons que F_{n+1} est libre. Si $0 = \lambda_1 f_1 + \dots + \lambda_{n+1} f_{n+1}$ avec $\lambda_{n+1} \neq 0$, on a, pour tout réel x :

$$0 = \lambda_1 \sin x + \lambda_2 \sin 2x + \dots + \lambda_{n+1} \sin (n+1)x$$

En dérivant deux fois et en changeant de signe, on obtient :

$$0 = \lambda_1 \sin x + 4\lambda_2 \sin 2x + \dots + (n+1)^2 \lambda_{n+1} \sin (n+1)x$$

En retranchant $(n+1)^2$ fois la première égalité de la seconde et en appliquant l'hypothèse de récurrence, on voit que $\lambda_k = 0$ pour $1 \leq k \leq n$. On en déduit $\lambda_{n+1} = 0$. Pour montrer que G_n est une famille libre, on peut raisonner de la même façon ou remarquer que si on dérive une fois $0 = \lambda_0 g_0 + \dots + \lambda_n g_n$, le résultat sur F_n montre que $\lambda_k = 0$ pour $1 \leq k \leq n$; on en déduit $\lambda_0 = 0$.

Pour montrer que H_n est une famille libre, on remarque qu'une relation de dépendance linéaire entre les éléments de H_n conduit à une égalité entre une fonction impaire (une somme de sinus) et une fonction paire (une somme de cosinus). Chacun des deux membres de l'égalité est donc nul et on utilise les résultats sur F_n et sur G_n .

Une démonstration de ces trois résultats faisant appel à des relations d'« orthogonalité » sera donnée au chapitre 16 (voir exercice 16.2).

h) Famille liée : si $0 = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$, on a, pour tout x de $] -2, 2[$:

$$0 = \lambda_1(x+2) + \lambda_2(x-2) + \lambda_3(4x+7)$$

ce qui conduit à $\lambda_1 = -\frac{15}{4}\lambda_3$ et $\lambda_2 = -\frac{1}{4}\lambda_3$.

Par exemple : $4f_3 - 15f_1 - f_2 = 0$. Les lecteurs et lectrices qui connaissent les fractions rationnelles en déduiront une décomposition en éléments simples de f_3 . La méthode d'identification n'est pas, en général, la meilleure méthode pour trouver les coefficients.

6.6 a) L'algorithme du pivot permet de calculer des familles (u_1, u'_2, u''_3) et (v_1, v'_2, v''_3) montrant que $\dim(F) = \dim(G) = 3$. Une troisième application de l'algorithme du pivot, à $(u_1, u'_2, u''_3, v_1, v'_2, v''_3)$, montre que $\dim(F \cup G) = 4$. On en déduit que $\dim(F \cap G) = 2$. Les deux colonnes nulles obtenues dans cette troisième application conduisent aux relations $v_1 + v_2 \in F$ et $v_3 - v_1 \in F$. On en déduit $F \cap G = \text{Vect}((1, 1, 1, 1, 1), (1, -1, 1, -1, 1))$.

b) Comme G et H sont contenues dans $G + H$, $F \cap G$ et $F \cap H$ sont contenus dans $F \cap (G + H)$ et leur somme l'est aussi. C'est tout ce qu'on peut dire car, par exemple, si $E = \mathbb{R}^2$, on a :

- d'une part, si $F = \text{Vect}(e_1)$, $G = \text{Vect}(e_1 + e_2)$, $H = \text{Vect}(e_1 - e_2)$, on a $G + H = \text{Vect}(e_1, e_2)$ donc $(F \cap G) + (F \cap H) = \{0\}$ et $F \cap (G + H) = \text{Vect}(e_1)$;
- d'autre part, si $F = G = H = \text{Vect}(e_1)$, on a $(F \cap G) + (F \cap H) = F \cap (G + H)$.

6.7 a) Comme on a une famille de 5 polynômes de degrés tous différents, c'est une base de $\mathbb{R}_4[X]$.

Tout polynôme de degré ≤ 4 s'écrit donc d'une façon et d'une seule comme combinaison linéaire de ces 5 polynômes. Si on écrit $P = \lambda_0 1 + \lambda_1(X - 1) + \lambda_2(X - 1)^2 + \lambda_3(X - 1)^3 + \lambda_4(X - 1)^4$, on a $P(1) = \lambda_0$. En dérivant une fois et en prenant la valeur de P' en 1, on trouve $P'(1) = \lambda_1$; en dérivant deux, trois et quatre fois et en prenant la valeur des dérivées en 1, on trouve $P''(1) = 2\lambda_2$, $P'''(1) = 6\lambda_3$, $P^{(iv)}(1) = 24\lambda_4$. On a donc $P = \sum_{0 \leq k \leq 4} \frac{(X - 1)^k}{k!} P^{(k)}(1)$.

C'est un cas particulier de la formule de Taylor pour les polynômes que nous montrerons au chapitre 13.

b) La famille est libre. En effet, si $0 = \lambda_1 X + \lambda_2 X(X + 1) + \lambda_3 (X + 1)^2$, on voit facilement que les trois coefficients sont nuls (évaluer le second membre en 0, puis en -1). Comme $\mathbb{R}_2[X]$ est de dimension 3, la famille est donc une base de $\mathbb{R}_2[X]$. Des calculs simples donnent :

$$1 = -X - X(X + 1) + (X + 1)^2, \quad X^2 = -X + X(X + 1) \text{ d'où :}$$

$$a + bX + cX^2 = (b - a - c)X + (c - a)X(X + 1) + a(X + 1)^2.$$

b) Si $0 = \sum_{0 \leq k \leq 4} \lambda_k P_k$, on montre que les coefficients sont nuls en évaluant successivement en 0, 1, 2, 3. La famille est donc une famille libre de 5 vecteurs de $\mathbb{R}_4[X]$. Comme $\mathbb{R}_4[X]$ est de dimension 5, la famille est donc une base de $\mathbb{R}_4[X]$. Pour obtenir ce résultat, on peut aussi remarquer que les polynômes de la famille sont de degrés 0, 1, 2, 3, 4.

On utilise les évaluations en 0, 1, 2, 3 pour déterminer les coordonnées des polynômes proposés dans cette base. On trouve :

$$1 + X + X^2 + X^3 = P_0 + 3P_1 + 8P_2 + 6P_3 ;$$

$$1 + X + X^2 + X^3 + X^4 = P_0 + 4P_1 + 22P_2 + 42P_3 + 24P_4.$$

d) La base B est constituée des $n + 1$ polynômes $\varepsilon_k(X) = \frac{\prod_{0 \leq i \leq k-1} (X - i)}{k!}$,

$0 \leq k \leq n$, tous de degrés différents. Pour a entier, $\varepsilon_k(a)$ est, soit nul, soit égal à un coefficient binomial (au signe près) ; c'est donc un entier. Enfin, si $P \in E$, il existe des réels $a_0, \dots, a_n$ tels que $P(X) = \sum_{0 \leq k \leq n} a_k \varepsilon_k(X)$. Supposons que P vérifie la propriété du texte. Comme $P(0) = a_0$, on voit que $a_0 \in \mathbb{Z}$; comme $P(1) = a_0 + a_1$, on voit que $a_1 \in \mathbb{Z}$; comme $P(2) = a_0 + 2a_1 + a_2$, on voit que $a_2 \in \mathbb{Z}$, etc. ; on peut écrire un raisonnement par récurrence.

6.8 a) Quand on écrit qu'une combinaison linéaire des P_i est nulle, on voit, en l'évaluant successivement en chacun des a_i , que chacun des coefficients est nul. La famille des P_i est donc libre. Comme elle a $n + 1$ éléments et que $\dim(E) = n + 1$, elle forme une base de E .

b) La même méthode donne les coefficients des P_i égaux aux b_i :

$$P = \sum_{0 \leq i \leq n} b_i P_i.$$

c) Plutôt que d'écrire un système d'équations avec les coefficients de P , on peut commencer par écrire P sur la base définie par les réels $a_i = i$ avec $i = 0, 1, 2$, ce qui donne l'unique résultat : $P = 3P_0 - 2P_1 + 5P_2$, soit, en effectuant : $P = 6X^2 - 11X + 3$. La complexité des calculs croît avec le degré.

APPLICATIONS LINÉAIRES

7

7.1 NAISSANCE DU CONCEPT

Nous avons déjà évoqué (voir 5.2) l'article que Giuseppe Peano écrit en 1888 et dans lequel il définit les espaces vectoriels et les applications linéaires entre espaces vectoriels : ce sont des applications qui ont la propriété de conserver les sommes de vecteurs et le produit par un scalaire. Il définit également la composée de deux applications linéaires et l'inverse d'une application linéaire.

Ce n'est que 30 ans plus tard que cette définition a montré tout son intérêt pour formaliser les problèmes qui se posaient dans d'autres champs des mathématiques comme l'analyse fonctionnelle que développa Stefan Banach.

Bien sûr, l'idée de linéarité était beaucoup plus ancienne et des mathématiciens comme Jean Bernoulli ou Euler en avaient déjà utilisé les propriétés pour parler de la dérivation, des différences finies ou encore des changements de repères. Mais ils n'appliquaient la linéarité qu'aux objets sur lesquels ils travaillaient. Pour parler d'application linéaire, il faut faire agir l'application linéaire, non seulement sur les objets sur lesquels on travaille, mais sur tous les objets du même type, autrement dit penser à l'espace vectoriel engendré par ces objets ou même à un espace encore plus grand.

7.2 APPLICATIONS LINÉAIRES

Définition 1 : application linéaire. Soient E et F deux espaces vectoriels sur $\mathbb{R}$. On appelle application linéaire de E dans F la donnée d'une application $f : E \rightarrow F$ telle que :

$$\text{pour tout } u \text{ et tout } v \text{ de } E : f(u + v) = f(u) + f(v).$$

$$\text{pour tout } u \text{ de } E \text{ et tout } \lambda \text{ de } \mathbb{R} : f(\lambda u) = \lambda f(u).$$

ou, de manière équivalente (vérification facile) :

$$\text{pour tout } u \text{ et tout } v \text{ de } E, \text{ tout } \lambda \text{ et tout } \mu \text{ de } \mathbb{R} :$$

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

On peut remplacer dans cette définition $\mathbb{R}$ par le corps $\mathbb{C}$ des nombres complexes ou tout autre corps K ; on pourra dire alors pour préciser que l'application est K -linéaire.

Propriétés. On a $f(0) = 0$ car pour tout u de E , on a $f(u) = f(u + 0) = f(u) + f(0)$.

La seconde propriété de la définition se généralise aux combinaisons linéaires finies

$$f(\lambda_1 u_1 + \cdots + \lambda_p u_p) = \lambda_1 f(u_1) + \cdots + \lambda_p f(u_p)$$

autrement dit :

$$f\left(\sum_{1 \leq k \leq p} \lambda_k u_k\right) = \sum_{1 \leq k \leq p} \lambda_k f(u_k)$$

Commentaire. Les applications linéaires sont définies par des applications entre ensembles qui sont compatibles avec la structure d'espace vectoriel. C'est une définition analogue qu'on va donner (cf. chapitres 11 et 12) pour les homomorphismes de groupes, d'anneaux, etc. ; on pourrait parler d'homomorphismes d'espaces vectoriels, mais le mot en usage est celui d'application linéaire.

Proposition 1 : composée d'applications linéaires. La composée de deux applications linéaires est linéaire.

Démonstration. Soient $f : E \rightarrow F$ et $g : F \rightarrow G$ deux applications linéaires. Pour tout u et tout v de E , pour tout λ de $\mathbb{R}$, on a :

$$\begin{aligned} (g \circ f)(u + v) &= g(f(u + v)) = g(f(u) + f(v)) \\ &= g(f(u)) + g(f(v)) = (g \circ f)(u) + (g \circ f)(v), \end{aligned}$$

$$(g \circ f)(\lambda u) = g(f(\lambda u)) = g(\lambda f(u)) = \lambda g(f(u)) = \lambda (g \circ f)(u). \quad \square$$

Définition 2 : somme d'applications linéaires. Soient $f : E \rightarrow F$ et $g : E \rightarrow F$ deux applications linéaires de même source et de même but. On définit leur somme $f + g : E \rightarrow F$ en posant :

$$(f + g)(u) = f(u) + g(u).$$

Proposition 2. *La somme de deux applications linéaires est linéaire.*

Démonstration. En effet, $(f + g)(u + v) = f(u + v) + g(u + v) = f(u) + f(v) + g(u) + g(v) = (f + g)(u) + (f + g)(v)$. De même, $(f + g)(\lambda u) = \lambda(f + g)(u)$. $\square$

Définition 3 : produit d'une application linéaire par un scalaire. On définit aussi le produit λf d'une application linéaire $f : E \rightarrow F$ par un scalaire λ en posant $(\lambda f)(u) = \lambda f(u)$.

Proposition 3. *Le produit d'une application linéaire par un scalaire est une application linéaire.* $\square$

Définition 4 : espace vectoriel des applications linéaires d'un espace vectoriel dans un autre. Soient E et F deux espaces vectoriels. L'ensemble $L(E, F)$ des applications linéaires de E dans F a une structure d'espace vectoriel pour les opérations de somme et de produit par un scalaire définies ci-dessus. Les vérifications sont immédiates. Nous utiliserons plus tard cette structure.

Définition 5 : endomorphisme. On appelle endomorphisme de E une application linéaire de E dans lui-même.

L'emploi du mot « endomorphisme » n'est pas obligatoire et on pourra parler indifféremment d'application linéaire de E dans E ou d'endomorphisme de E .

On note $L(E)$ l'espace vectoriel des endomorphismes de E .

Commutativité. Soient f et g deux endomorphismes de E . En général, $g \circ f$ et $f \circ g$ ne sont pas égaux, sauf cas particuliers.

7.3 EXEMPLES

Exemples généraux. Voici d'abord des exemples généraux valables dans tout espace vectoriel E .

Pour tout réel a , l'application $h_a : E \rightarrow E$ définie par $h_a(u) = au$ est linéaire. On appelle h_a homothétie de rapport a . Les homothéties étudiées dans le cadre de la géométrie affine enseignée au lycée ont un centre et un rapport ; elles agissent sur des points. Les homothéties considérées ici agissent sur des vecteurs et on doit seulement préciser leur rapport.

Pour $a = 1$, h_a est l'application identique notée $\text{id}(E)$ ou $\text{id}_E : E \rightarrow E$, qu'on abrégera en id . Pour $a = -1$, h_a est l'application $u \mapsto -u$ notée $-\text{id}$ qui peut être vue comme une symétrie.

Pour $a = 0$, h_a est l'application nulle notée 0 .

Il faut noter que, si w est un vecteur non nul de E , l'application de translation $t_w : u \mapsto u + w$ n'est pas une application linéaire car $t_w(u) + t_w(u') = u + u' + 2w \neq t_w(u + u')$. Cela ne signifie pas que ce n'est pas une application intéressante ! Son étude se fait dans le cadre de la géométrie affine.

Espaces de fonctions. Nous allons donner quelques exemples avec l'espace $E = C^\infty(\mathbb{R}, \mathbb{R})$.

- Soit a un réel. L'application $\text{ev}_a : E \rightarrow \mathbb{R}$, appelée évaluation en a et définie par $\text{ev}_a(f) = f(a)$, est linéaire :

$$\text{ev}_a(\lambda f + \mu g) = (\lambda f + \mu g)(a) = \lambda f(a) + \mu g(a) = \lambda \text{ev}_a(f) + \mu \text{ev}_a(g).$$

- L'application $D : E \rightarrow E$ définie par $D(f) = f'$ qui dérive les fonctions est linéaire, puisqu'on sait que la dérivée d'une somme de deux fonctions est la somme des dérivées, etc.

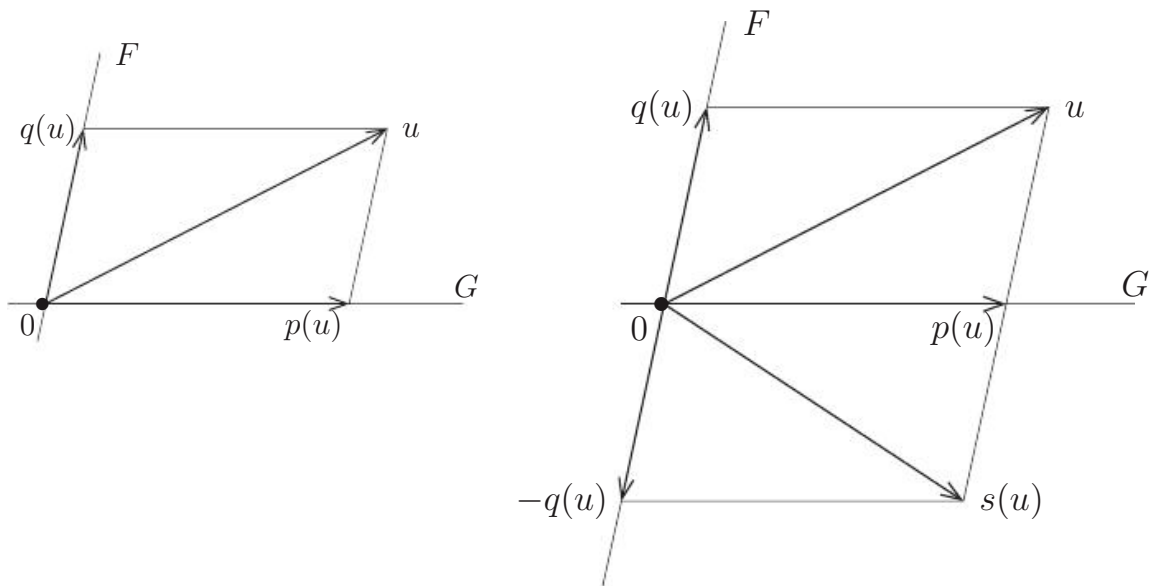
Comme la somme et la composée de deux applications linéaires, ainsi que le produit d'une application linéaire par un scalaire, sont linéaires, l'application $F = D \circ D - 3D + 2\text{id}$ est linéaire. Comme $F(f) = f'' - 3f' + 2f$, on voit qu'on peut considérer la résolution des équations différentielles linéaires à coefficients constants d'un point de vue nouveau : résoudre $f'' - 3f' + 2f = 0$, c'est trouver les f tels que $F(f) = 0$.

- Soient a et b des réels. L'application $\varphi : E \rightarrow \mathbb{R}$ définie par $\varphi(f) = \int_a^b f(t) dt$ est linéaire, d'après les propriétés de l'intégrale.
- Soit a un réel. L'application $\tau_a : E \rightarrow E$ qui associe à une fonction f la fonction $x \mapsto f(x + a)$ est une application linéaire ; le graphe de $\tau_a(f)$ est le translaté du graphe de f par le vecteur $(-a, 0)$.
- Soit encore a un réel. On appelle différence finie l'application linéaire $\Delta_a = \tau_a - \text{id}$, autrement dit $\Delta_a(f)$ est la fonction $x \mapsto f(x + a) - f(x)$.

Espaces $\mathbb{R}^n$. Voici quelques exemples classiques d'applications linéaires $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ ou $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ que nous généraliserons à des espaces quelconques. Les figures employées dans ces cas particuliers peuvent servir à se représenter ces généralisations.

- Projection $p : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ parallèlement à une droite vectorielle F sur une droite vectorielle G .

Dans le dessin (figure de gauche ci-dessous), $p(u)$ appartient à G et $u - p(u) = q(u)$ est dans F . Ici, $q(u)$ est l'image de u par la projection sur F parallèlement à G . On a $p + q = \text{id}$.



Quand on choisit une base (e_1, e_2) de $\mathbb{R}^2$ avec $e_1 \in G$ et $e_2 \in F$ alors $p : (x, y) \mapsto (x, 0)$, $q : (x, y) \mapsto (0, y)$.

On définit de même une projection dans $\mathbb{R}^3$ sur une droite vectorielle parallèlement à un plan vectoriel ou une projection sur un plan vectoriel parallèlement à une droite vectorielle. Par exemple : $(x, y, z) \mapsto (x, 0, 0)$, $(x, y, z) \mapsto (x, 0, z)$.

- Symétrie s dans $\mathbb{R}^2$ par rapport à une droite vectorielle G parallèlement à une droite vectorielle F .

Dans le dessin (figure de droite ci-dessus), on a représenté le symétrique $s(u)$ de u par rapport à G parallèlement à F . On a $s(u) = p(u) - q(u)$ et $s(u) - u = 2(p(u) - u)$.

Quand on choisit une base (e_1, e_2) de $\mathbb{R}^2$ avec $e_1 \in G$ et $e_2 \in F$ alors $s : (x, y) \mapsto (x, -y)$.

On définit de même une symétrie dans $\mathbb{R}^3$ par rapport à une droite vectorielle parallèlement à un plan vectoriel ou une symétrie par rapport à un plan vectoriel parallèlement à une droite vectorielle. Par exemple : $(x, y, z) \mapsto (x, -y, -z)$, $(x, y, z) \mapsto (x, -y, z)$.

7.4 PROPRIÉTÉ UNIVERSELLE

Dans les exemples précédents, les applications linéaires ont une définition globale sur l'espace tout entier. La proposition suivante, qu'on appellera propriété universelle d'un espace vectoriel muni d'une base, permet de construire une application linéaire en donnant simplement les images des vecteurs d'une base de l'espace de départ.

Proposition : propriété universelle des espaces vectoriels munis de base. Soit E un espace vectoriel de dimension finie n et soit $B = (e_1, \dots, e_n)$ une base de E .

Pour tout espace vectoriel F et toute famille $(u_1, \dots, u_n)$ de n vecteurs de F , il existe une unique application linéaire $l : E \rightarrow F$ telle que $l(e_i) = u_i$ pour $1 \leq i \leq n$.

Démonstration. Soit F un espace vectoriel et $(u_1, \dots, u_n)$ une famille de n vecteurs de F .

Soit u un vecteur de E . Il s'écrit de manière unique $u = \sum_{1 \leq i \leq n} x_i e_i$. On a nécessairement $l(u) = \sum_{1 \leq i \leq n} x_i u_i$, ce qui permet de définir une application $l : E \rightarrow F$. Mais l est-elle linéaire ? Vérifions-le. Si $v = \sum_{1 \leq i \leq n} y_i e_i$ est un autre vecteur de E , on a :

$$\begin{aligned} l(u+v) &= l\left(\sum_{1 \leq i \leq n} x_i e_i + \sum_{1 \leq i \leq n} y_i e_i\right) = l\left(\sum_{1 \leq i \leq n} (x_i + y_i) e_i\right) \\ &= \sum_{1 \leq i \leq n} (x_i + y_i) u_i = \sum_{1 \leq i \leq n} x_i u_i + \sum_{1 \leq i \leq n} y_i u_i = l(u) + l(v). \end{aligned}$$

et de même, on vérifie : $l(\lambda u) = \lambda l(u)$. □

Exemple. Notons $e_1, \dots, e_n$ les vecteurs de la base canonique de $\mathbb{R}^n$. La donnée de n vecteurs $u_1, \dots, u_n$ d'un espace vectoriel E équivaut donc à la donnée d'une application linéaire $l : \mathbb{R}^n \rightarrow E$ telle que $l(e_i) = u_i$ pour $1 \leq i \leq n$.

Application linéaire définie par un vecteur. En particulier, la donnée d'une application linéaire $\mathbb{R} \rightarrow E$ équivaut à la donnée du vecteur u image de 1. On notera $l_u : \mathbb{R} \rightarrow E$ cette application linéaire ; on a donc $l_u(\lambda) = \lambda u$ pour tout réel λ et, en particulier, $l_u(1) = u$. Si $f : E \rightarrow F$ est une application linéaire, le composé de f avec l_u est l'application linéaire de $\mathbb{R}$ dans F définie par $f(u)$. On a donc $l_{f(u)}(1) = f(u) = f(l_u(1)) = (f \circ l_u)(1)$. On peut donc écrire : $f \circ l_u = l_{f(u)}$, autrement dit, le diagramme suivant :

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{l_u} & E \\ & \searrow l_{f(u)} & \downarrow f \\ & & F \end{array}$$

est commutatif.

7.5 NOYAU D'UNE APPLICATION LINÉAIRE

Définition : noyau d'une application linéaire. Soient E et F deux espaces vectoriels et $f: E \rightarrow F$ une application linéaire. On appelle noyau de f et on note $\ker(f)$ (la notation vient de l'allemand *Kern* : noyau) l'ensemble des vecteurs de E dont l'image est le vecteur nul de F , autrement dit :

$$\ker(f) = \{u \mid u \in E \text{ et } f(u) = 0\}$$

Proposition 1. *Le noyau d'une application linéaire $f: E \rightarrow F$ est un sous-espace vectoriel de E .*

Démonstration. Soient u et u' dans le noyau de f et λ et μ deux réels. Il faut montrer que $\lambda u + \mu u'$ est dans $\ker(f)$, ce qui résulte de :

$$f(\lambda u + \mu u') = \lambda f(u) + \mu f(u') = 0 + 0 = 0.$$

□

Exemples.

➤ Posons $E = C^\infty(\mathbb{R}, \mathbb{R})$ et considérons l'équation différentielle :

$$f'' - 3f' + 2f = 0 \tag{E}$$

comme en 7.3. L'ensemble des solutions $S(E)$ est $\ker(F)$.

➤ Prenons maintenant l'ensemble E des suites de nombres réels et considérons la relation

$$u_{n+2} - u_{n+1} - u_n = 0. \tag{R}$$

L'ensemble $S(R)$ des suites vérifiant (R) est $\ker(F)$ en définissant l'application linéaire $F: (u_n)_{n \in \mathbb{N}} \mapsto (u_{n+2} - u_{n+1} - u_n)_{n \in \mathbb{N}}$.

➤ L'exemple des systèmes linéaires sera étudié au paragraphe 7.9.

Applications linéaires injectives. Rappelons qu'une application entre ensembles $f: A \rightarrow B$ est dite injective si elle « conserve les différences », c'est-à-dire si, pour tous x et x' de A distincts, on a $f(x) \neq f(x')$. On sait que, pour les démonstrations, il est souvent plus pratique de montrer que, pour tout x et x' de A , si $f(x) = f(x')$, alors $x = x'$.

On dispose d'un critère beaucoup plus simple pour montrer qu'une application linéaire est injective.

Proposition 2 : critère d'injectivité. *Soit $f: E \rightarrow F$ une application linéaire. L'application linéaire f est injective si et seulement si $\ker(f) = \{0\}$.*

Démonstration. Supposons f injective. On sait que $f(0) = 0$; aucun autre élément de E ne peut avoir pour image 0, donc $\ker(f) = \{0\}$.

Réciproquement, si u et u' sont deux vecteurs ayant la même image : $f(u) = f(u')$, on a : $f(u - u') = 0$, donc $u - u' \in \ker(f) = \{0\}$, donc $u = u'$. $\square$

Proposition 3. Soit $f : E \rightarrow F$ une application linéaire injective. L'image d'une famille libre $(u_1, \dots, u_p)$ de E , en particulier l'image d'une base de E , est une famille libre de F .

Démonstration. Il s'agit de montrer que la famille $(f(u_1), \dots, f(u_p))$ est libre. Supposons que 0 s'écrive sous la forme $0 = \lambda_1 f(u_1) + \dots + \lambda_p f(u_p)$. L'application f étant linéaire, on a : $0 = f(\lambda_1 u_1 + \dots + \lambda_p u_p)$. Comme le noyau de f est $\{0\}$, on a $\lambda_1 u_1 + \dots + \lambda_p u_p = 0$; la famille $(u_1, \dots, u_p)$ étant libre, les λ_i , $1 \leq i \leq p$ sont tous nuls, ce qui permet de conclure. $\square$

7.6 IMAGE D'UNE APPLICATION LINÉAIRE

Définition : image d'une application linéaire. Soient E et F deux espaces vectoriels et $f : E \rightarrow F$ une application linéaire. On appelle image de f et on note $\text{Im}(f)$ ou $f(E)$ l'ensemble des vecteurs de F images d'au moins un vecteur de E , autrement dit :

$$\text{Im}(f) = f(E) = \{v \in F \text{ tels qu'il existe } u \in E \text{ tel que } f(u) = v\}$$

Proposition 1 : image d'un sous-espace. Soient E et F deux espaces vectoriels et $f : E \rightarrow F$ une application linéaire. L'image par f d'un sous-espace vectoriel G de E est un sous-espace vectoriel de F noté $f(G)$. En particulier, $\text{Im}(f) = f(E)$ est un sous-espace vectoriel de F .

Démonstration. Soient v et v' dans l'image de G et λ et μ deux réels. Il existe des éléments u et u' de G , tels que $f(u) = v$ et $f(u') = v'$. Il faut montrer que $\lambda v + \mu v'$ est dans $f(G)$, ce qui résulte de :

$$\lambda v + \mu v' = \lambda f(u) + \mu f(u') = f(\lambda u + \mu u').$$

$\square$

Sous-espaces stables. Soient E un espace vectoriel, F un sous-espace de E , f un endomorphisme de E . On dit que F est stable par f si $f(F) \subset F$.

Proposition 2 : image de l'espace engendré. Soient E et F des espaces vectoriels et $f : E \rightarrow F$ une application linéaire.

1) Soit $(u_1, \dots, u_n)$ une famille de vecteurs de E . Alors :

$$f(\text{Vect}(u_1, \dots, u_n)) = \text{Vect}(f(u_1), \dots, f(u_n)).$$

2) Soit $(u_1, \dots, u_n)$ une famille génératrice de vecteurs de E , par exemple une base de E . Alors $\text{Im}(f) = \text{Vect}(f(u_1), \dots, f(u_n))$.

Démonstration. 1) Si $v \in f(\text{Vect}(u_1, \dots, u_n))$, v est de la forme $v = f(\lambda_1 u_1 + \dots + \lambda_n u_n)$, donc $v = \lambda_1 f(u_1) + \dots + \lambda_n f(u_n)$, ce qui prouve que $v \in \text{Vect}(f(u_1), \dots, f(u_n))$.

Réciproquement, si v est dans $\text{Vect}(f(u_1), \dots, f(u_n))$, v est de la forme $v = \lambda_1 f(u_1) + \dots + \lambda_n f(u_n)$, d'où $v = f(\lambda_1 u_1 + \dots + \lambda_n u_n)$; donc v est dans $f(\text{Vect}(u_1, \dots, u_n))$.

Le 2) est une conséquence directe du 1). $\square$

Rang d'une application linéaire. Le rang d'une application linéaire est par définition la dimension de $\text{Im}(f)$. On le notera $\text{rg}(f)$. Pour déterminer le rang de $f: E \rightarrow F$, il suffit, d'après la proposition précédente, de connaître une base de F et d'appliquer l'algorithme du chapitre 3 aux vecteurs $f(u_1), \dots, f(u_n)$ images par f d'une famille génératrice de E .

Exemple. Considérons la projection $p: \mathbb{R}^5 \rightarrow \mathbb{R}^5$ définie par $(x, y, z, s, t) \mapsto (x, 0, z, 0, t)$. La famille (e_1, e_3, e_5) est de rang 3 et engendre $\text{Im}(f)$, donc $\text{rg}(f) = 3$.

7.7 LE THÉORÈME DU RANG OU DES DIMENSIONS

Théorème du rang ou des dimensions. Soient E un espace vectoriel de dimension finie, F un espace vectoriel et $f: E \rightarrow F$ une application linéaire. Alors $\ker(f)$ et $\text{Im}(f)$ sont des espaces vectoriels de dimension finie et :

$$\dim(\text{Im}(f)) + \dim(\ker(f)) = \dim(E).$$

Démonstration. Comme $\ker(f)$ est un sous-espace vectoriel de l'espace de dimension finie E , il est de dimension finie.

Posons $\dim(E) = n$ et soit $(u_1, \dots, u_r)$ une base de $\ker(f)$. On peut compléter cette base en une base $(u_1, \dots, u_r, \dots, u_n)$ de E . Montrons que $B = (f(u_{r+1}), \dots, f(u_n))$ est une base de $\text{Im}(f)$.

On sait que B est une famille génératrice de $\text{Im}(f)$.

B est aussi une famille libre de $\text{Im}(f)$, car si :

$$0 = \lambda_{r+1} f(u_{r+1}) + \dots + \lambda_n f(u_n),$$

on a :

$$f(\lambda_{r+1}u_{r+1} + \dots + \lambda_n u_n) = 0,$$

donc :

$$\lambda_{r+1}u_{r+1} + \dots + \lambda_n u_n \in \ker(f).$$

Ce vecteur étant dans $\ker(f)$, s'écrit $\lambda_1 u_1 + \dots + \lambda_r u_r$. On a donc :

$$\lambda_{r+1}u_{r+1} + \dots + \lambda_n u_n = \lambda_1 u_1 + \dots + \lambda_r u_r.$$

Tous les coefficients de cette égalité sont nuls car $(u_1, \dots, u_r, \dots, u_n)$ est une base de E , en particulier $\lambda_{r+1}, \dots, \lambda_n$ sont nuls ; donc B est une famille libre de $\text{Im}(f)$.

Par conséquent, B est une base de $\text{Im}(f)$. Comme elle a $n - r$ éléments et que $\dim(\ker(f)) = r$, on conclut. $\square$

7.8 RÉOLUTION D'UNE ÉQUATION LINÉAIRE

Proposition. Soit $f: E \rightarrow F$ une application linéaire et soit $b \in F$. La résolution de l'équation linéaire

$$f(u) = b$$

se ramène :

- 1) soit à montrer que b n'est pas dans $\text{Im}(f)$ et alors l'équation n'a aucune solution ;
- 2) soit, d'une part, à trouver une solution particulière $s \in E$ de l'équation $f(u) = b$ et, d'autre part, à déterminer le noyau $\ker(f)$ de f , autrement dit à résoudre l'équation $f(u) = 0$. Alors, l'ensemble S des solutions de l'équation $f(u) = b$ s'écrit :

$$S = s + \ker(f) = \{s + v \mid v \in \ker(f)\}.$$

Démonstration. Soit $u \in s + \ker(f)$; u est de la forme $s + v$ avec $v \in \ker(f)$ et on a $f(u) = f(s + v) = f(s) + f(v) = b$, donc $u \in S$.

Réciproquement, soit $u \in S$; on a $f(u - s) = f(u) - f(s) = b - b = 0$, donc $u - s \in \ker(f)$, soit $u \in s + \ker(f)$. $\square$

Cette proposition importante unifie les résultats vus dans les premiers chapitres. Ici encore, on peut dire que la solution générale de l'équation avec second membre $f(u) = b$ s'obtient comme somme d'une solution particulière s de cette équation (si elle existe) et de la solution générale de l'équation sans second membre $f(u) = 0$.

On remarquera également que cette proposition ne fait aucune hypothèse sur les dimensions de E et de F et vaut donc pour des espaces de dimension infinie.

Le principe de superposition (le nom vient de la superposition de solutions dans des problèmes de physique) de 1.6 et 2.3 s'explique dans ce cadre général : pour résoudre une équation linéaire $f(x) = b$ avec $b = b_1 + b_2$, on peut, grâce à la linéarité, scinder la difficulté en cherchant x_1, x_2 tels que $f(x_1) = b_1$ et $f(x_2) = b_2$; alors $x_1 + x_2$ est solution de l'équation initiale.

7.9 RÉOLUTION D'UN SYSTÈME LINÉAIRE

Applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^n$. Notons $(e_1, \dots, e_p)$ la base canonique de $\mathbb{R}^p$ et $(\varepsilon_1, \dots, \varepsilon_n)$ la base canonique de $\mathbb{R}^n$.

Donnons-nous des vecteurs $v_1, \dots, v_p$ de $\mathbb{R}^n$: $v_i = \sum_{1 \leq k \leq n} a_{ki} \varepsilon_k$ pour $1 \leq i \leq p$. La propriété universelle nous assure de l'existence d'une application linéaire unique $f: \mathbb{R}^p \rightarrow \mathbb{R}^n$ telle que $f(e_i) = v_i$ pour $1 \leq i \leq p$.

Soit $u = x_1 e_1 + \dots + x_p e_p$ un vecteur de $\mathbb{R}^p$. Son image par f est $f(u) = f(x_1 e_1 + \dots + x_p e_p) = x_1 v_1 + \dots + x_p v_p$. La première coordonnée de $f(u)$ est $a_{11}x_1 + \dots + a_{1p}x_p$, la seconde coordonnée de $f(u)$ est $a_{21}x_1 + \dots + a_{2p}x_p$, etc.

Réciproquement, une application $f: \mathbb{R}^p \rightarrow \mathbb{R}^n$ telle que $f(x_1, \dots, x_p)$ soit un vecteur de $\mathbb{R}^n$ de la forme $(a_{11}x_1 + \dots + a_{1p}x_p, a_{21}x_1 + \dots + a_{2p}x_p, \dots)$ est une application linéaire : c'est l'application linéaire définie par $f(e_i) = \sum_{1 \leq k \leq n} a_{ki} \varepsilon_k$ pour $1 \leq i \leq p$.

Par exemple, l'application $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $f(x, y, z) = (2x - 3y + z, y - 5z, 7x + y - z)$ est une application linéaire et $f(e_1) = (2, 0, 7), f(e_2) = (-3, 1, 1), f(e_3) = (1, -5, -1)$.

Noyau et solutions du système homogène. Avec les notations précédentes, dire que u appartient au noyau de f équivaut à dire que les n coordonnées de $f(u)$ sont nulles, autrement dit, la recherche du noyau de f équivaut à la recherche de l'ensemble des solutions du système linéaire L sans second membre :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = 0 & l_1 \\ \vdots & \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = 0 & l_n \end{cases}$$

Dans ce système, pour $1 \leq k \leq p$, les coefficients de la k -ième inconnue de L sont les coordonnées du vecteur v_k dans la base canonique de $\mathbb{R}^n$.

Le système L équivaut donc à l'unique équation

$$f(u) = 0$$

et l'ensemble des solutions de L est $S(L) = \ker(f)$.

Cas particuliers. C'est le moment de voir comment interpréter les cas $n = 0$ ou $p = 0$. L'application linéaire f devient :

- si $p = 0$, l'application $\{0\} \rightarrow \mathbb{R}^n$, donc $S(L) = \{0\}$;
- si $n = 0$, l'application $\mathbb{R}^p \rightarrow \{0\}$, donc $S(L) = \mathbb{R}^p$;
- en particulier, si $n = p = 0$, c'est-à-dire si le système n'a aucune équation et aucune inconnue, $S(L) = \{0\}$.

Égalité des deux notions de rang. Si r note le rang du système L , on sait, d'après 4.7, que $S(L)$ est un espace engendré par $n - r$ vecteurs linéairement indépendants. Comme $S(L) = \ker(f)$, on a :

$$\dim(\ker(f)) = n - r,$$

ce qui prouve que le nombre $r = n - \dim(\ker(f))$ ne dépend pas de la méthode de résolution du système L comme on l'avait annoncé au chapitre 4.

D'autre part, l'image de f est $\text{Vect}(u_1, \dots, u_p)$. On a défini le rang de la famille $(u_1, \dots, u_p)$ au chapitre 3 et on sait que c'est $\dim(\text{Vect}(u_1, \dots, u_p))$. Ce rang est donc $\dim(\text{Im}(f)) = n - \dim(\ker(f)) = r$.

Les notions de rang d'un système linéaire et de rang d'une famille de vecteurs sont donc enfin toutes deux bien définies et reliées.

Résolution d'un système linéaire. L'« abondance » des coefficients nous a empêché jusqu'ici de voir un système linéaire avec second membre sous la forme :

$$f(u) = b \tag{L'}$$

ce qui permet de lui associer le système homogène sans second membre

$$f(u) = 0 \tag{L}$$

dont les solutions forment l'espace vectoriel $\ker(f)$. Les résultats généraux sur les solutions d'une équation linéaire vus au paragraphe précédent s'appliquent aux systèmes linéaires et on retrouve les résultats du chapitre 4 :

- si $b \notin \text{Im}(f)$, alors $S(L') = \emptyset$;
- si $b \in \text{Im}(f)$, il existe donc au moins un élément s de E tel que $f(s) = b$ et on a $S(L') = s + \ker(f) = s + S(L)$.

Équation d'un hyperplan. Une application linéaire $f : \mathbb{R}^p \rightarrow \mathbb{R}$ est de la forme $(x_1, \dots, x_p) \mapsto a_1x_1 + \dots + a_px_p$. Si les coefficients ne sont pas tous nuls, $\dim(\text{Im}(f)) = 1$, donc $\dim(\ker(f)) = p - 1$. Le noyau de f est un hyperplan de $\mathbb{R}^p$. Il est défini par l'équation $a_1x_1 + \dots + a_px_p = 0$.

On peut voir la résolution d'un système linéaire comme la détermination de l'intersection de la famille d'hyperplans définis par les équations du système.

7.10 ISOMORPHISMES

Définition. Soient E et F deux espaces vectoriels. Un isomorphisme d'espaces vectoriels est une application linéaire $f: E \rightarrow F$ telle qu'il existe une application linéaire $g: F \rightarrow E$ inverse de f , c'est-à-dire telle que $g \circ f = \text{id}(E)$ et que $f \circ g = \text{id}(F)$. L'inverse d'un isomorphisme f est noté f^{-1} .

Un isomorphisme d'un espace vectoriel dans lui-même est appelé automorphisme de E .

Les isomorphismes sont toujours intéressants puisque leur existence assure que les mêmes propriétés (relatives à la structure considérée) sont vérifiées dans la source et le but de l'isomorphisme.

Proposition 1. *Le composé de deux isomorphismes $f: E \rightarrow F$ et $g: F \rightarrow G$ est un isomorphisme $g \circ f: E \rightarrow G$ et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.* $\square$

Les automorphismes de E forment donc un groupe ; nous l'étudierons un peu dans le chapitre 11.

Pour montrer qu'une application linéaire est un isomorphisme, deux propositions vont nous permettre de réduire les vérifications.

Proposition 2. *Soit $f: E \rightarrow F$ une application linéaire.*

L'application f est un isomorphisme de E sur F si et seulement si f est bijective.

Démonstration. Si f est un isomorphisme d'inverse g , $f \circ g = \text{id}(F)$ montre que f est surjective et $g \circ f = \text{id}(E)$ montre que f est injective.

Réciproquement, si f est bijective, elle possède, en tant qu'application de l'ensemble E dans l'ensemble F , une application réciproque $g: F \rightarrow E$. Il reste à prouver que cette application g est linéaire.

Soient v et v' deux éléments quelconques de F . Il existe u et u' dans E tels que $f(u) = v$ et $f(u') = v'$. On a $g(v + v') = g(f(u) + f(u')) = g(f(u + u')) = u + u' = g(v) + g(v')$. On montre de même que $g(\lambda v) = \lambda g(v)$. $\square$

Proposition 3. *Soient E et F des espaces vectoriels de même dimension n et $f: E \rightarrow F$ une application linéaire.*

- 1) *Si f est injective, alors f est un isomorphisme.*
- 2) *Si f est surjective, alors f est un isomorphisme.*
- 3) *S'il existe une application linéaire $g: F \rightarrow E$ telle que $g \circ f = \text{id}_E$, alors f est un isomorphisme.*
- 4) *S'il existe une application linéaire $g: F \rightarrow E$ telle que $f \circ g = \text{id}_F$, alors f est un isomorphisme.*

Démonstration. 1) Si f est injective, on a $\ker(f) = \{0\}$ et le théorème de la dimension donne : $\dim(\operatorname{Im}(f)) = n = \dim(F)$, d'où $\operatorname{Im}(f) = F$, donc f est surjective.

2) Si f est surjective, $\dim(\operatorname{Im}(f)) = n = \dim(F)$, donc $\dim(\ker(f)) = 0$. Par conséquent, $\ker(f) = \{0\}$, ce qui prouve que f est injective.

Dans les deux cas, f étant injective et surjective est bijective et la proposition précédente montre que c'est un isomorphisme.

3) La condition implique que f est injective et on applique le 1).

4) La condition implique que f est surjective et on applique le 2). □

Enfin, le résultat suivant classifie complètement les espaces vectoriels réels de dimension finie.

Proposition 4 : classification des espaces vectoriels réels de dimension finie.

1) Deux espaces vectoriels de dimensions finies sont isomorphes si et seulement s'ils ont la même dimension.

2) Soit E un espace vectoriel de dimension n . Alors E est isomorphe à $\mathbb{R}^n$.

Démonstration.

1) Supposons que $f : E \rightarrow F$ est un isomorphisme ; comme f est injective, $\ker(f) = \{0\}$; comme f est surjective, $\operatorname{Im}(f) = F$. Le théorème de la dimension donne $\dim(E) = \dim(F)$.

Réciproquement, notons n la dimension commune de E et de F , $(e_1, \dots, e_n)$ une base de E , $(\varepsilon_1, \dots, \varepsilon_n)$ une base de F . La propriété universelle montre qu'il existe une application linéaire $f : E \rightarrow F$ définie par $f(e_j) = \varepsilon_j$. Comme $\operatorname{Im}(f)$ est engendré par les images des vecteurs de base, f est surjective ; c'est donc un isomorphisme.

2) Conséquence immédiate du 1). □

Commentaire.

1) Les propriétés d'un espace vectoriel de dimension n seront donc celles de $\mathbb{R}^n$ et ne nécessitent pas d'étude supplémentaire.

2) Cette proposition ressemble à la propriété des ensembles finis : s'ils ont même nombre d'éléments, ils sont isomorphes, au sens d'isomorphisme entre ensembles, c'est-à-dire qu'il existe une bijection entre eux.

Voici pour terminer deux propositions dont nous laissons la démonstration en exercice.

Proposition 5. Soit $f: E \rightarrow F$ un isomorphisme d'espaces vectoriels. Pour tout sous-espace G de E , la restriction de f à G , notée $f|_G$, est un isomorphisme de G sur $f(G)$.

Proposition 6. Soit $f: E \rightarrow F$ une application linéaire, $\varphi: E' \rightarrow E$, $\psi: F \rightarrow F'$ des isomorphismes d'espaces vectoriels. On a $\text{rg}(f) = \text{rg}(f \circ \varphi) = \text{rg}(\psi \circ f)$.

➤ Vers le chapitre 8

Nous avons dit que pour calculer dans un espace vectoriel de dimension finie, il fallait introduire la notion de base. Maintenant que nous avons défini les applications linéaires et présenté quelques-unes de leurs propriétés, nous allons introduire les matrices pour pouvoir faire des calculs sur les applications linéaires.

Exercices

7.1 Linéarité

a) Parmi les applications suivantes, indiquer celles qui sont linéaires.

- $f_1: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $(x, y, z) \mapsto (x + 2y + 3z, 2y - z, x + z)$;
- $f_2: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $(x, y, z) \mapsto (0, 2y - z, 0)$;
- $f_3: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $(x, y, z) \mapsto (x + 2y + 3, 2y - z, x + z)$;
- $f_4: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $(x, y, z) \mapsto (x + 2y + 3z, 2yz, x + z)$;
- $f_5: \mathbb{R}^3 \rightarrow \mathbb{R}$ définie par $(x, y, z) \mapsto x + 2y + 3z$;
- $f_6: \mathbb{R} \rightarrow \mathbb{R}$ définie par $x \mapsto x^2 + 2x$.

b) Même question, en posant $E = C^\infty(\mathbb{R}, \mathbb{R})$.

- $\varphi_1: E \rightarrow E$ définie par $f \mapsto 3f'' + 8f' + 5f$;
- $\varphi_2: E \rightarrow E$ définie par $f \mapsto f'' - 2xf' + 6f$;
- $\varphi_3: E \rightarrow E$ définie par $f \mapsto ff'$;
- $\varphi_4: E \rightarrow E$ définie par $f \mapsto (x \mapsto f(x) - \int_0^x (x-t)f(t) dt)$;
- $\varphi_5: E \rightarrow \mathbb{R}$ définie par $f \mapsto f(0) - 2f(3)$.

7.2 Existence d'applications linéaires, noyau, image dans les $\mathbb{R}^n$

1) On note (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$.

- a) Déterminer l'image des vecteurs de la base canonique de $\mathbb{R}^3$ par les applications f_1, f_2, f_3 de l'exercice 7.1.
- b) Existe-t-il un endomorphisme $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ vérifiant les conditions suivantes :
 - $g(e_1) = e_2 + e_3, g(e_2) = e_1, g(e_3) = e_2 - e_3$;
 - $g(e_1 + e_2) = e_3, g(e_2 + e_3) = e_1, g(e_3 - e_1) = e_2$.

Si oui, indiquer l'image d'un vecteur quelconque (x, y, z) de $\mathbb{R}^3$.

- 2) Décrire sous forme de Vect, en précisant leurs dimensions, le noyau et l'image de l'application linéaire $f : \mathbb{R}^4 \rightarrow \mathbb{R}^5$ définie par $f(e_1) = (1, 1, 1, 2, 5)$, $f(e_2) = (2, 1, 0, 3, 4)$, $f(e_3) = (-1, 0, -1, 4, 7)$, $f(e_4) = (-9, -2, 1, -1, 9)$ où (e_1, e_2, e_3, e_4) note la base canonique de $\mathbb{R}^4$.
- 3) On note (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. On définit l'endomorphisme $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ par $f(e_1) = 3e_1 + 4e_2 - 2e_3$, $f(e_2) = -e_1 - e_2 + e_3$, $f(e_3) = e_1 + 2e_2$ et on pose $g = f - \text{id}$.

a) Calculer $f(x, y, z)$ et $g(x, y, z)$.

b) Déterminer les noyaux et images de f et de g et les comparer.

c) Trouver une base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ de $\mathbb{R}^3$ telle que $f(\varepsilon_1) = 0, f(\varepsilon_2) = \varepsilon_2, f(\varepsilon_3) = \varepsilon_3$.

- 4) Soient E un espace vectoriel de dimension 4, de base (e_1, e_2, e_3, e_4) et F un espace vectoriel de dimension 5, de base $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5)$. Construire, si cela est possible, des applications linéaires $f : E \rightarrow F$ vérifiant :

a) $\text{Im}(f) = \text{Vect}(\varepsilon_2, \varepsilon_4)$;

b) $\text{Im}(f) = \{0\}$;

c) $\text{Im}(f) = F$;

d) $\ker(f) = \text{Vect}(e_2 + e_3, e_4)$ et $\text{Im}(f) = \text{Vect}(\varepsilon_2, \varepsilon_3)$.

7.3 Noyau, image dans les espaces de polynômes

- 1) On note $\Delta : \mathbb{R}[X] \rightarrow \mathbb{R}[X]$ l'application linéaire définie par $\Delta(P) = P(X + 1) - P(X)$.

a) Décrire sous forme de Vect, en précisant leurs dimensions, le noyau et l'image de la restriction Δ_4 de Δ à $\mathbb{R}_4[X]$.

- b) Même question en remplaçant 4 par un entier n quelconque. En déduire que Δ est surjective.
- c) Soit Q un polynôme de $\mathbb{R}[X]$. Montrer que l'équation $\Delta(P) = Q$ a toujours des solutions. Comment peut-on déduire l'ensemble des solutions de cette équation de l'une d'entre elles ?
- 2) On note $f: \mathbb{R}_4[X] \rightarrow \mathbb{R}_4[X]$ l'application linéaire définie par $f(P) = (X-1)P' - P$.
- a) Calculer $f(aX^4 + bX^3 + cX^2 + dX + e)$. En déduire $\ker(f)$.
- b) L'équation $f(P) = Q$ a-t-elle des solutions dans $\mathbb{R}_4[X]$ pour tout Q de $\mathbb{R}_4[X]$?
- c) Calculer $f((X-1)^k)$ pour $k = 0, 1, 2, 3, 4$. En déduire une caractérisation des polynômes Q pour lesquels l'équation $f(P) = Q$ a des solutions.
- d) Résoudre $(X-1)P' - P = X^2 - 2X + 2$.

7.4 Noyau, image dans les espaces de fonctions

On pose $E = C^0(\mathbb{R}, \mathbb{R})$. On considère l'application $\varphi: E \rightarrow E$ définie par $\varphi(f) = xf$.

- 1) Vérifier que φ est une application linéaire.
- 2) Déterminer $\ker(\varphi)$.
- 3) L'application φ est-elle surjective ?
- 4) Caractériser les éléments de $\text{Im}(\varphi)$.

7.5 Géométrie dans le plan et l'espace

On se place dans le plan ou l'espace de la géométrie du secondaire, considérés comme des espaces vectoriels, munis d'une base orthonormée : (e_1, e_2) ou (e_1, e_2, e_3) (l'usage du secondaire est plutôt de noter les vecteurs de base $\vec{i}, \vec{j}$, et $\vec{k}$). L'étude générale de l'orthogonalité sera abordée au chapitre 16.

- 1) Donner l'action sur les vecteurs de base des applications linéaires suivantes :
 - a) rotation ρ d'angle θ dans le plan ;
 - b) symétrie orthogonale σ dans le plan par rapport à l'axe défini par le vecteur $(\cos \theta, \sin \theta)$.
- 3) Déterminer l'image d'un vecteur $u = (x, y, z)$ de l'espace par les applications linéaires suivantes :
 - d) rotation directe ρ d'angle θ autour de l'axe défini par e_3 ;
 - e) rotation directe ρ' d'angle θ autour de l'axe défini par e_1 .

7.6 Petits exercices théoriques

- 1) Soient E, F et G trois espaces vectoriels et soient $f, g : E \rightarrow F, h : F \rightarrow G$ des applications linéaires. Montrer que $h \circ (f + g) = h \circ f + h \circ g$ et que, pour tout réel a , on a : $(ah) \circ f = h \circ (af) = a(h \circ f)$.
- 2) Soit $f : E \rightarrow F$ une application linéaire entre espaces de dimensions finies.
 - a) Montrer que si f est injective, alors il existe une application linéaire $g : F \rightarrow E$ telle que $g \circ f = \text{Id}_E$.
 - b) Montrer que si f est surjective, alors il existe une application linéaire $g : F \rightarrow E$ telle que $f \circ g = \text{Id}_F$.
- 3) Donner un exemple d'applications linéaires $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ et $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ vérifiant $g \circ f = \text{id}_{\mathbb{R}^2}$.
- 4) Soient E un espace vectoriel, f et g deux endomorphismes de E tels que $f \circ g = g \circ f$. Montrer que $\ker(f)$ et $\text{im}(f)$ sont stables par g .
- 5) Avec les notations de la section 7.4, vérifier que $l_u + l_{u'} = l_{u+u'}$ et que $\lambda l_u = l_{\lambda u}$.

7.7 Différence finie

On pose $E = \text{Fonct}(\mathbb{R}, \mathbb{R})$ et on considère l'application $\Delta : E \rightarrow E$ définie par $\Delta(f)(x) = f(x+1) - f(x)$.

- 1) Vérifier que Δ est une application linéaire.
- 2) Déterminer $\ker(\Delta)$.
- 3) Déterminer $\ker(\Delta^2)$ en utilisant le résultat précédent.
- 4) Déterminer $\ker(\Delta^n)$ pour tout entier $n \geq 1$.

7.8 Noyaux itérés

- 1) Soit E un espace vectoriel et $f : E \rightarrow E$ un endomorphisme de E .
 - a) Montrer que $\ker(f) \subset \ker(f^2)$ et plus généralement, que $\ker(f^n) \subset \ker(f^{n+1})$.
 - b) Montrer que s'il existe un entier s tel que $\ker(f^s) = \ker(f^{s+1})$, alors $\ker(f^s) = \ker(f^{s+n})$ pour tout entier $n \geq 0$.
 - c) Montrer que si E est de dimension finie, il existe un entier s tel que $\ker(f^s) = \ker(f^{s+n})$ pour tout entier $n \geq 0$.

- 2) On suppose maintenant que $E = C^\infty(\mathbb{R}, \mathbb{R})$ et on note D l'application $f \mapsto f'$. On se propose de montrer, en raisonnant par l'absurde, qu'il n'existe pas d'endomorphisme L de E et d'entier $n \geq 2$ tels que $D = L^n$. Supposons l'existence de L et n .
- Déterminer $\ker(D)$.
 - Montrer que $\ker(L)$ n'est pas réduit à $\{0\}$.
 - En déduire $\ker(L^k) = \ker(D)$ pour tout entier $k \geq 1$.
 - Obtenir une contradiction en considérant $\ker(D^2)$.

Solutions

7.1 a) f_1, f_2, f_5 sont linéaires : on peut le vérifier pour chacune des trois applications à partir des axiomes ou le voir comme une application de 7.9.

f_3, f_4, f_6 ne sont pas linéaires : on peut remarquer, par exemple, que $f_3(0) = (3, 0, 0) \neq 0$, $f_4(2u) \neq 2f_4(u)$ en général, $f_6(2) = 8 \neq 2f_6(1)$.

b) φ_1, φ_2 sont linéaires. On peut le vérifier directement ou remarquer qu'on a des sommes, des produits par des scalaires ou des fonctions données de l'application linéaire de dérivation D ou de l'identité.

φ_4 est linéaire d'après les propriétés de linéarité de l'intégrale.

φ_5 est linéaire : le voir directement ou remarquer que $\varphi_5 = \text{ev}_0 - 2 \text{ev}_3$ où ev_a désigne l'évaluation en a .

φ_3 n'est pas linéaire car, par exemple : $\varphi_3(2f) = 4\varphi_3(f) \neq 2\varphi_3(f)$ pour $f' \neq 0$.

7.2 1) a) $f_1(e_1) = f_1((1, 0, 0)) = (1, 0, 1)$, etc.

b) • oui, d'après la propriété universelle ; $g(x, y, z) = (y, x + z, x - z)$.

• non, car $e_2 = g(e_3 - e_1) = g((e_2 + e_3) - (e_1 + e_2)) = g(e_2 + e_3) - g(e_1 + e_2) = e_1 - e_3$ est impossible.

2) Nous utilisons les calculs faits dans l'exercice 3.6 e). $\text{Im}(f)$ est engendré par les images des vecteurs de base. On a déterminé le rang de ce système par l'algorithme du pivot : c'est 3. On en déduit que le noyau de f est de dimension 1, par le théorème de la dimension. L'algorithme du pivot nous a montré que $\text{Im}(f)$ est engendré par $u_1 = f(e_1)$, $u_2 = f(e_2)$, $u_3 = f(e_3)$ et que $u_4 = f(e_4) = 2u_3 - 5u_2 + 3u_1$. Donc $f(e_4 - 2e_3 + 5e_2 - 3e_1) = 0$ et $\ker(f) = \text{Vect}(e_4 - 2e_3 + 5e_2 - 3e_1)$.

3) **a)** $f(x, y, z) = xf(e_1) + yf(e_2) + zf(e_3)$, donc

$$f(x, y, z) = (3x - y + z, 4x - y + 2z, -2x + y).$$

$$g(x, y, z) = (2x - y + z, 4x - 2y + 2z, -2x + y - z).$$

- b)** Pour trouver les noyaux de f et g , on résout des systèmes linéaires. On trouve $\ker(f) = \text{Vect}(1, 2, -1)$ et $\ker(g) = \text{Vect}((-1, -1, 1), (0, 1, 1))$.
- Pour trouver les images de f et de g , on applique l'algorithme du pivot aux familles de vecteurs $(f(e_1), f(e_2), f(e_3))$ et $(g(e_1), g(e_2), g(e_3))$; on trouve $\text{Im}(f) = \ker(g)$ et $\text{Im}(g) = \ker(f)$.
- c)** Les calculs précédents conduisent à $\varepsilon_1 = (1, 2, -1)$, $\varepsilon_2 = (-1, -1, 1)$, $\varepsilon_3 = (1, 2, 0)$.
- 4) a)** Prendre, par exemple, $f(e_1) = \varepsilon_2$, $f(e_2) = \varepsilon_4$ et $f(e_3) = f(e_4) = 0$.
- b)** Prendre les images des e_i nulles.
- c)** Impossible : d'après le théorème de la dimension, on a : $\dim(\text{Im}(f)) \leq \dim(E) = 4$.
- d)** Prendre $f(e_1) = \varepsilon_2$, $f(e_2) = \varepsilon_3$ et $f(e_3) = -\varepsilon_3$, $f(e_4) = 0$ et vérifier que les conditions sont bien remplies.

7.3 1) a) Comme $\Delta(X^k) = (X + 1)^k - X^k$ est un polynôme de degré $k - 1$, $\text{Im}(\Delta_4)$ est engendré par la famille triangulaire $\Delta(X^4) = 4X^3 + 6X^2 + 4X + 1$, $\Delta(X^3) = 3X^2 + 3X + 1$, $\Delta(X^2) = 2X + 1$, $\Delta(X) = 1$, $\Delta(1) = 0$, donc $\text{Im}(\Delta_4) = \mathbb{R}_3[X]$ et $\text{rg}(\Delta_4) = 4$; comme $\mathbb{R}_4[X]$ est de dimension 5, $\ker(\Delta_4)$ est de dimension 1 ; on voit que $\ker(\Delta_4)$ est le sous-espace des polynômes constants.

- b)** On trouve de même que $\text{Im}(\Delta_{n+1}) = \mathbb{R}_n[X]$. Par conséquent, tout polynôme de degré n est image par Δ d'un polynôme de degré $n + 1$, ce qui prouve que Δ est surjective.
- c)** L'équation $\Delta(P) = Q$ est une équation linéaire. Comme Δ est surjectif, elle admet une solution S . L'ensemble des solutions est de la forme $S + \ker(\Delta)$, autrement dit, c'est l'ensemble des polynômes égaux à S à une constante près. On remarquera qu'on ne s'est pas préoccupé de calculer effectivement S .
- 2) a)** $f(aX^4 + bX^3 + cX^2 + dX + e) = 3aX^4 + (2b - 4a)X^3 + (c - 3b)X^2 - 2cX - d - e$ donc $\ker(f) = \text{Vect}(X - 1)$.
- b)** Comme la dimension du noyau est 1, l'image de f a pour dimension 4, ce qui prouve que f n'est pas surjective, donc l'équation $f(P) = Q$ n'a pas toujours de solution.
- c)** $f((X - 1)^k) = (k - 1)(X - 1)^k$. On a donc $\text{Im}(f) = \text{Vect}((X - 1)^4, (X - 1)^3, (X - 1)^2, 1)$.
Comme $((X - 1)^4, (X - 1)^3, (X - 1)^2, X - 1, 1)$ est une base de $\mathbb{R}_4[X]$, tout polynôme Q dont l'écriture dans cette base n'a pas de terme en $X - 1$ est dans l'image de f et l'équation $f(P) = Q$ a des solutions.
- d)** $X^2 - 2X + 2 = (X - 1)^2 + 1 = f((X - 1)^2 - 1)$, donc l'ensemble des solutions est $X^2 - 2X + 2 + \text{Vect}(X - 1)$.

7.4 2) Si $f \in \ker(\varphi)$, on a $xf(x) = 0$ pour tout x réel, donc $f(x) = 0$ pour tout x non nul. Comme f est continue, cela implique $f(0) = 0$, donc $f = 0$; d'où $\ker(\varphi) = \{0\}$; φ est injective.

3) La fonction $x \mapsto 1$ n'est pas dans l'image de φ , car $xf(x) = 1$ est impossible pour $x = 0$.

4) L'équation $xf = g$ définit $f(x)$ pour tout x non nul par $f(x) = \frac{g(x)}{x}$. Pour que f existe, il faut donc qu'on puisse prolonger cette définition par continuité en 0, c'est-à-dire que g soit dérivable en 0. Cette condition assure l'existence de f .

7.5 1) a) $\rho(e_1) = (\cos \theta, \sin \theta)$, $\rho(e_2) = (-\sin \theta, \cos \theta)$.

b) $\sigma(e_1) = (\cos 2\theta, \sin 2\theta)$, $\sigma(e_2) = (\sin 2\theta, -\cos 2\theta)$.

2) On calcule d'abord les images des vecteurs de base :

$\rho(e_1) = (\cos \theta, \sin \theta, 0)$, $\rho(e_2) = (-\sin \theta, \cos \theta, 0)$, $\rho(e_3) = e_3$;

$\rho'(e_1) = e_1$, $\rho'(e_2) = (0, \cos \theta, \sin \theta)$, $\rho'(e_3) = (0, -\sin \theta, \cos \theta)$.

On a $\rho(x, y, z) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta, z)$ et un résultat analogue pour ρ' .

7.6 1) L'application des définitions donne les égalités :

$$[h \circ (f + g)](u) = h((f + g)(u)) = h(f(u) + g(u)) = h(f(u)) + h(g(u))$$

$$= (h \circ f)(u) + (h \circ g)(u) = [h \circ f + h \circ g](u) \quad , \text{ etc.}$$

2) a) Soit $(e_1, \dots, e_n)$ une base de E . Si f est injective, $(f(e_1), \dots, f(e_n))$ est une famille libre de F qu'on peut compléter en une base $(f(e_1), \dots, f(e_n), \varepsilon_{n+1}, \dots, \varepsilon_{n+l})$ de F . On peut alors définir g par les images des vecteurs de cette base : $g(f(e_1)) = e_1, \dots, g(f(e_n)) = e_n$, $g(\varepsilon_{n+1}) = 0, \dots, g(\varepsilon_{n+l}) = 0$.

b) Soit $(\varepsilon_1, \dots, \varepsilon_n)$ une base de F . Si f est surjective, il existe $e_1, \dots, e_n$ tels que $f(e_i) = \varepsilon_i$ pour $1 \leq i \leq n$. On définit g en posant $g(\varepsilon_i) = e_i$ pour $1 \leq i \leq n$.

3) Prendre, par exemple, pour f l'application $(x, y) \mapsto (x, y, 0)$ de $\mathbb{R}^2$ dans $\mathbb{R}^3$ et pour g l'application $(x, y, z) \mapsto (x, y)$ de $\mathbb{R}^3$ dans $\mathbb{R}^2$.

4) Si $u \in \ker(f)$, on a $f(g(u)) = g(f(u)) = 0$, donc $g(u) \in \ker(f)$. Si $v = f(u) \in \text{im}(f)$, on a $g(v) = g(f(u)) = f(g(u)) \in \text{im}(f)$.

5) Les deux termes des égalités donnent la même image pour 1.

7.7 2) $\ker(\Delta)$ est l'ensemble des fonctions périodiques de période 1.

3) $\ker(\Delta^2)$ est l'ensemble des fonctions dont l'image par Δ est dans $\ker(\Delta)$. Il faut donc résoudre les équations linéaires de la forme $\Delta(f) = \varphi$ où φ est une fonc-

tion périodique de période 1. On connaît déjà le noyau de Δ , il reste à trouver une solution particulière de l'équation : on peut trouver $x \mapsto x\varphi(x)$ (ou aussi $x \mapsto E(x)\varphi(x)$ où $E(x)$ note la partie entière de x).

On peut dire que les éléments de $\ker(\Delta^2)$ sont de la forme $x \mapsto x\varphi_1(x) + \varphi_0(x)$ où φ_1 et φ_0 sont des fonctions périodiques de période 1.

- 4) On montre alors que $\ker(\Delta^n)$ est l'ensemble des fonctions de la forme $x \mapsto \sum_{0 \leq k \leq n-1} x^k \varphi_k(x)$ où les φ_k sont des fonctions périodiques de période 1.

7.8 1) b) Supposons que $\ker(f^s) = \ker(f^{s+1})$.

Si $u \in \ker(f^{s+2})$, on a $f(u) \in \ker(f^{s+1}) = \ker(f^s)$ donc $u \in \ker(f^{s+1})$, etc.

- c) Considérons la suite des dimensions des espaces $\ker(f^s)$. C'est une suite croissante d'entiers bornée par la dimension de E . On ne peut donc avoir $\dim(\ker(f^s)) < \dim(\ker(f^{s+1}))$ pour tout s entier.

- 2) a) $\ker(D)$ est le sous-espace Vect (1) des fonctions constantes.

- b) Si $\ker(L) = \{0\}$, alors L est injective et $D = L^n$ aussi.

Comme $\ker(L) \subset \ker(D)$, que la dimension de $\ker(D)$ est 1 et celle de $\ker(L) \geq 1$, on a $\ker(L) = \ker(D)$. Comme $\ker(L) \subset \ker(L^2) \subset \ker(D)$, on a $\ker(L) = \ker(L^2)$, donc $\ker(L^k) = \ker(D)$ pour tout entier $k \geq 1$ comme ci-dessus.

- c) On aurait donc $\ker(D) = \ker(L^{2n}) = \ker(D^2)$. Mais $\ker(D^2)$ est l'ensemble des fonctions affines ; il est donc différent de $\ker(D)$.

8.1 MATRICE D'UNE APPLICATION LINÉAIRE

8.1.1 Définition de la matrice d'une application linéaire

Soient E un espace vectoriel de dimension finie $p \geq 1$ et F un espace vectoriel de dimension finie $n \geq 1$. On note $B = (e_1, \dots, e_p)$ une base de E et $C = (\varepsilon_1, \dots, \varepsilon_n)$ une base de F .

On a vu qu'une application linéaire était déterminée par les images des vecteurs de base (voir la propriété universelle du chapitre précédent). Donc, si $f: E \rightarrow F$ est une application linéaire, elle est déterminée par les vecteurs $f(e_1), \dots, f(e_p)$.

Chacun de ces vecteurs est défini par ses coordonnées dans la base C . La notation de ces coordonnées nécessite un double indice. On notera a_{ij} la i -ième coordonnée dans la base C de l'image $f(e_j)$ du j -ième vecteur de la base B . Ces np coefficients sont regroupés dans un tableau de n lignes et de p colonnes, appelé matrice de l'application linéaire f par rapport aux bases B et C qu'on présente de la façon suivante :

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

Exemple. Si $B = (e_1, e_2, e_3)$ est la base canonique de $\mathbb{R}^3$, et si $C = (\varepsilon_1, \varepsilon_2)$ est la base canonique de $\mathbb{R}^2$, l'application linéaire $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ définie par $f(e_1) = 2\varepsilon_1 - \varepsilon_2$, $f(e_2) = 5\varepsilon_2$, $f(e_3) = -\varepsilon_1 + 3\varepsilon_2$ a pour matrice :

$$\begin{pmatrix} 2 & 0 & -1 \\ -1 & 5 & 3 \end{pmatrix}$$

Le slogan à retenir est :

On met en colonnes les images des vecteurs de la base de « départ » repérés dans la base « d'arrivée ».

Pour ne pas se tromper, les débutants pourront, pour s'habituer à ce que signifie une matrice d'application linéaire, écrire en haut des colonnes les transformés des vecteurs de la base de départ et à gauche ou à droite des lignes les vecteurs de la base d'arrivée, ce qui donne pour l'exemple précédent :

$$\begin{array}{ccc} & f(e_1) & f(e_2) & f(e_3) \\ \varepsilon_1 & \begin{pmatrix} 2 & 0 & -1 \end{pmatrix} & & \\ \varepsilon_2 & \begin{pmatrix} -1 & 5 & 3 \end{pmatrix} & & \end{array}$$

La matrice de f dépend des bases B et C et on la notera $M(f, B, C)$, ce qu'on abrégiera en $M(f)$ quand il n'y aura pas d'ambiguïté sur les bases. Le plus souvent, on utilisera une seule lettre pour désigner une matrice ; on dira, par exemple : la matrice A de f par rapport aux bases B et C . Pour indiquer les coefficients de A , on écrira $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ en supprimant la mention des bornes de variation des indices i et j dès que cela ne créera pas de confusion.

Ordre des vecteurs de base. On voit tout de suite que la matrice d'une application linéaire ne sera pas la même suivant l'ordre dans lequel on prend les vecteurs de B et l'ordre dans lequel on prend les vecteurs de C . Dans l'exemple précédent, posons $B' = (e_3, e_1, e_2)$ et $C' = (\varepsilon_2, \varepsilon_1)$. On a :

$$M(f, B', C') = \begin{pmatrix} f(e_3) & f(e_1) & f(e_2) \\ 3 & -1 & 5 \\ -1 & 2 & 0 \end{pmatrix} \begin{array}{l} \varepsilon_2 \\ \varepsilon_1 \end{array}$$

autrement dit :

$$M(f, B', C') = \begin{pmatrix} 3 & -1 & 5 \\ -1 & 2 & 0 \end{pmatrix}$$

8.1.2 Matrice d'un vecteur

Soient E un espace vectoriel de dimension p muni d'une base $B = (e_1, \dots, e_p)$ et u un vecteur de E . On a vu en 7.4 que u définit une application linéaire $l_u : \mathbb{R} \rightarrow E$

telle que $l_u(1) = u$. La matrice $U = M(l_u, \text{can}, B)$ est la matrice formée par les coordonnées $(u_1, \dots, u_p)$ de u dans la base B :

$$U = \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$$

Image d'un vecteur par une application linéaire. Conservons les notations précédentes et soit maintenant F un espace vectoriel de dimension n muni d'une base $C = (\varepsilon_1, \dots, \varepsilon_n)$ et $f: E \rightarrow F$ une application linéaire. On pose $A = M(f, B, C) = (a_{ij})$.

Comment trouver avec U et A la matrice $V = M(l_{f(u)}, \text{can}, C)$? On vient de voir que V est la matrice colonne formée par les coordonnées $(v_1, \dots, v_n)$ de $v = f(u)$ dans la base C .

Comme :

$$\begin{aligned} f(u) &= f\left(\sum_{1 \leq j \leq p} u_j e_j\right) = \sum_{1 \leq j \leq p} u_j f(e_j) \\ &= \sum_{1 \leq j \leq p} u_j \left(\sum_{1 \leq i \leq n} a_{ij} \varepsilon_i\right) = \sum_{1 \leq i \leq n} \left(\sum_{1 \leq j \leq p} a_{ij} u_j\right) \varepsilon_i, \end{aligned}$$

on a $v_i = \sum_{1 \leq j \leq p} a_{ij} u_j$.

On voit que v_i s'obtient en prenant les termes de la i -ième ligne de A et en les multipliant par les termes de U de même rang puis en faisant la somme de ces produits.

Ce que nous venons de décrire est un cas particulier du produit de deux matrices que nous présenterons en 8.6. Nous écrirons $V = AU$:

$$V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = AU = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$$

Par exemple, l'image du vecteur $(2, 3)$ de $\mathbb{R}^2$ par l'application linéaire $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par la matrice

$$A = \begin{pmatrix} 1 & -1 \\ 3 & -2 \\ -2 & 2 \end{pmatrix}$$

est définie par la matrice

$$V = \begin{pmatrix} 1 & -1 \\ 3 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

où les trois termes de V s'obtiennent par les calculs :

$$-1 = 1 \times 2 + (-1) \times 3,$$

$$0 = 3 \times 2 + (-2) \times 3,$$

$$2 = (-2) \times 2 + 2 \times 3.$$

8.2 MATRICES ET APPLICATIONS LINÉAIRES

Définition 1 : matrice. Une matrice à coefficients réels de $n \geq 1$ lignes et $p \geq 1$ colonnes est un tableau rectangulaire de np nombres réels $(a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

Le terme général a_{ij} de cette matrice est situé à l'intersection de la i -ème ligne et de la j -ième colonne. Une matrice à n lignes et p colonnes est dite de type (n, p) .

Les matrices à coefficients dans un autre corps se définissent de la même façon.

On note M_{np} l'ensemble des matrices à coefficients réels de n lignes et p colonnes. Si on veut préciser le corps K dans lequel sont pris les coefficients, on écrira $M_{np}(K)$. Si $n = p$, on utilisera les notations M_n ou $M_n(K)$.

On conviendra de ne pas mettre de parenthèses si $n = p = 1$.

Une dernière remarque : on peut envisager des matrices sans lignes ou sans colonnes (cas $n = 0$ ou $p = 0$) pour certains cas particuliers.

Application linéaire associée à une matrice. Nous venons en 8.1 d'associer une matrice à une application linéaire.

Soit maintenant une matrice de np coefficients : $A = (a_{ij})$. Soient E un espace vectoriel de dimension finie p muni d'une base $B = (e_1, \dots, e_p)$ et F un espace vectoriel de dimension finie n muni d'une base $C = (\varepsilon_1, \dots, \varepsilon_n)$.

On peut associer à A une application linéaire $f : E \rightarrow F$ de la façon suivante : les images des vecteurs de la base B de E sont définies comme les vecteurs de F dont les coordonnées dans la base C sont les colonnes successives de la matrice, c'est-à-dire que $f(e_j) = \sum_{1 \leq i \leq n} a_{ij} \varepsilon_i$ pour $1 \leq j \leq p$.

En particulier, on peut définir une application linéaire $\mathbb{R}^p \rightarrow \mathbb{R}^n$ dont la matrice par rapport aux bases canoniques soit (a_{ij}) .

La correspondance entre matrices et applications linéaires est donc la suivante.

Proposition : correspondance entre matrices et applications linéaires. Soient E un espace vectoriel de dimension finie p muni d'une base B et F un espace vectoriel de dimension finie n muni d'une base C . L'application Φ définie par $\Phi(f) = M(f, B, C)$ est une bijection de l'ensemble $L(E, F)$ des applications linéaires de E dans F sur l'ensemble M_{np} .

Nous allons définir la somme de deux matrices, le produit d'une matrice par un scalaire et voir que la bijection Φ est compatible avec ces deux opérations : elle transforme la somme d'applications linéaires en somme des matrices correspondantes, etc.

Définition 2 : somme de matrices. Soient $A = (a_{ij})$ et $A' = (a'_{ij})$ deux matrices de M_{np} . On définit la somme des matrices A et A' en posant $A + A' = (a_{ij} + a'_{ij})$.

C'est la matrice dont les coefficients sont les sommes des coefficients de mêmes indices de A et de A' .

Matrice de la somme. Soient E un espace vectoriel de dimension finie p muni d'une base $B = (e_1, \dots, e_p)$ et F un espace vectoriel de dimension finie n muni d'une base C . Soient $f, f' : E \rightarrow F$ des applications linéaires. La matrice de $f + f'$ par rapport aux bases B et C est : $M(f + f', B, C) = M(f, B, C) + M(f', B, C)$.

La justification de cette propriété est simple : les colonnes de $M(f + f', B, C)$ sont données par les coordonnées de $(f + f')(e_i)$ dans la base C ; comme $(f + f')(e_i) = f(e_i) + f'(e_i)$, ce sont les sommes des coordonnées de $f(e_i)$ et de $f'(e_i)$ dans la base C .

Définition 3 : produit d'une matrice par un scalaire. Soient $A = (a_{ij})$ une matrice de M_{np} et λ un scalaire. On définit la matrice λA en posant $\lambda A = (\lambda a_{ij})$. C'est la matrice A dont *tous* les coefficients ont été multipliés par λ .

Matrice de λf . Soient E un espace vectoriel de dimension finie p muni d'une base $B = (e_1, \dots, e_p)$ et F un espace vectoriel de dimension finie n muni d'une base C . Soient $f : E \rightarrow F$ une application linéaire. La matrice de λf par rapport aux bases B et C est : $M(\lambda f, B, C) = \lambda M(f, B, C)$.

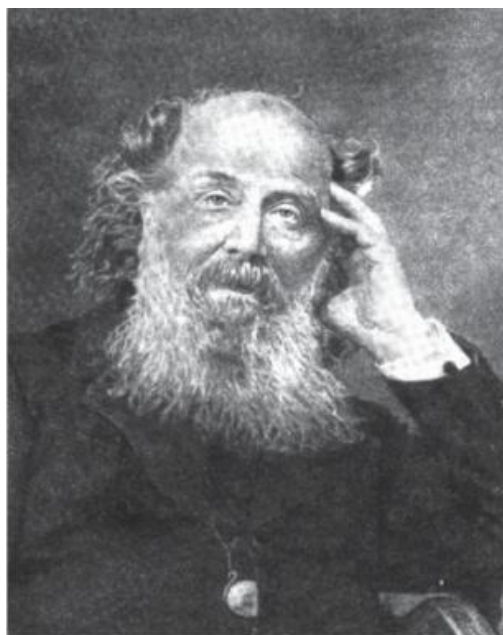
En effet, les colonnes de $M(\lambda f, B, C)$ sont données par les coordonnées des $(\lambda f)(e_i) = \lambda f(e_i)$ dans la base C .

Remarque. Les propriétés précédentes permettent de munir M_{np} d'une structure d'espace vectoriel avec la somme des matrices et le produit d'une matrice par un scalaire. On peut alors dire que l'application $\Phi : L(E, F) \rightarrow M_{np}$ est un isomorphisme d'espaces vectoriels.

8.3 UN PEU D'HISTOIRE

Nous avons vu (voir 4.1) comment Liu Hui résolvait un système linéaire en calculant sur le tableau des coefficients. C'est peut-être la première idée de l'intérêt des tableaux de nombres. L'étude des systèmes linéaires au XVIII^e siècle conduit à développer le calcul des déterminants (voir chapitre 14). La complexité technique des calculs fascine les mathématiciens pendant un siècle. La notation matricielle comme tableau rectangulaire ou carré de nombres, apparaît au XIX^e siècle : Gauss l'emploie lorsqu'il substitue des combinaisons linéaires de termes dans d'autres, obtenant la formule du produit de deux matrices carrées d'ordre 3, et indique que son résultat s'étend à des dimensions quelconques, sans préciser.

Le développement des calculs avec les déterminants conduit aussi à la notion de matrice. C'est James Sylvester (1814-1897) qui invente le mot en 1850 (en anglais : matrix) pour désigner des tableaux de nombres, et c'est Arthur Cayley (1821-1895) qui publie, en 1858, un article considéré comme fondateur.



James Joseph Sylvester (1814-1897)

Iouchkiévitch et Kolmogorov – *Mathematika 19 vovieka*, Moskva, 1978

Apparition du mot *Matrix*

For this purpose we must commence, not with a square, but with an oblong arrangement of terms consisting, suppose, of m lines and n columns. This will not in itself represent a determinant, but is, as it were, a Matrix out of which we may form various systems of determinants by fixing upon a number p , and selecting at will p lines and p columns, the squares corresponding to which may be termed determinants of the p th order.

Sylvester J. – Addition to the articles *On a new class of theorems* and *On Pascal theorem*, in *Philosophical Magazine*, 1850.

Les notations de Cayley à cette époque sont très proches de nos usages actuels : il utilise des parenthèses pour la première ligne et des traits droits pour les suivantes :

$$\begin{array}{l} (a, \quad b, \quad c) \\ |a', \quad b', \quad c'| \\ |a'', \quad b'', \quad c''| \end{array}$$

ce qui lui permet d'écrire un système linéaire de trois équations à trois inconnues x, y, z :

$$(X, Y, Z) = \begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix} (x, y, z)$$

Beaucoup de choses sont définies dans cet article, comme la somme, le produit de deux matrices, l'inverse d'une matrice, les puissances, etc. (voir aussi le chapitre 15). Cayley ne se pose pas de problèmes de notation générale des coefficients d'une matrice, n'envisageant que des matrices de petites tailles où ses $a, b, c, a', \dots$ lui suffisent sans avoir à utiliser des indices.

Quelques années plus tard, en 1867, Edmond Laguerre (1834-1886) expose les mêmes idées avec une vision encore plus proche de la nôtre. Mais le calcul matriciel était encore peu connu des non mathématiciens dans les années 1920, par exemple du physicien Werner Heisenberg (1901-1976). C'est un de ses collègues qui lui a expliqué que les calculs qu'il conduisait pour fonder la mécanique quantique se comprenaient en terme de matrices.

8.4 MATRICES PARTICULIÈRES

Nous allons rencontrer différents types de matrices qu'il faut maintenant décrire.

Matrices carrées. Lorsque $p = n$, la matrice M a autant de lignes que de colonnes et on dit que c'est une matrice carrée d'ordre n (on dit aussi de taille n) :

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Matrices diagonales. Une matrice carrée dont les termes se trouvant en dehors de la diagonale sont nuls est appelée matrice diagonale. Si $(a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ est une matrice diagonale, on a donc $a_{ij} = 0$ si $i \neq j$:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Matrice unité. On appelle matrice unité d'ordre n et on note I_n la matrice diagonale telle que $a_{ii} = 1$ pour $1 \leq i \leq n$. Par exemple,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Chapitre 8 • Matrices

I_n est la matrice de l'application identique d'un espace vectoriel E de dimension n dans n'importe quelle base de E (on oublie souvent l'indice n).

Matrices triangulaires. Une matrice carrée dont tous les termes au-dessous de la diagonale sont nuls est appelée matrice triangulaire supérieure. Si (a_{ij}) est une matrice triangulaire supérieure d'ordre n , on a donc $a_{ij} = 0$ si $i > j$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

De même, les matrices triangulaires inférieures sont telles que $a_{ij} = 0$ si $i < j$.

Matrices lignes, matrices colonnes. Si $n = 1$, la matrice possède une seule ligne et est appelée matrice ligne :

$$(a_{11} \quad \dots \quad a_{1p})$$

Si $p = 1$, la matrice possède une seule colonne et est appelée matrice colonne :

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$$

Base de M_{np} On notera E_{ij} la matrice de M_{np} dont tous les coefficients sont nuls sauf celui de la ligne i et de la colonne j qui vaut 1 ; les dimensions de la matrice E_{ij} sont supposées être données par le contexte. Par exemple, dans les matrices carrées d'ordre 3, on a :

$$E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Une matrice $A = (a_{ij})$ s'écrit donc $A = \sum_{ij} a_{ij} E_{ij}$ de manière unique. Les np matrices E_{ij} , $1 \leq i \leq n$, $1 \leq j \leq p$, de M_{np} forment une base de cet espace vectoriel. En particulier, les n^2 matrices E_{ij} , $1 \leq i, j \leq n$ forment une base de M_n .

Matrices de transvection. Une matrice carrée d'ordre n de la forme $T_{ij}(\alpha) = I + \alpha E_{ij}$, avec $j \neq i$, est appelée matrice de transvection. Par exemple, si $n = 3$:

$$T_{23}(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{pmatrix}$$

Matrices de transposition. Une matrice carrée d'ordre n de la forme $P_{ij} = I + E_{ij} + E_{ji} - E_{ii} - E_{jj}$, avec $j \neq i$, est appelée matrice de transposition. On peut aussi décrire cette matrice comme la matrice I_n dans laquelle les lignes d'indice i et j ont été échangées. Par exemple, si $n = 3$:

$$p_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Matrices symétriques. Une matrice carrée (a_{ij}) d'ordre n est dite symétrique si $a_{ij} = a_{ji}$ pour tout i et tout j de $\{1, \dots, n\}$.

8.5 EXEMPLES

Nous présentons quelques exemples de matrices d'applications linéaires entre espaces $\mathbb{R}^n$. Il ne faut pas oublier le slogan énoncé en 8.1 : *mettre en colonnes les images des vecteurs de base*.

Matrice d'une rotation de $\mathbb{R}^2$. On suppose le plan muni d'une base orthonormée (e_1, e_2) au sens de la géométrie du lycée. Notons $\rho(\theta) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la rotation d'angle θ . En calculant les coordonnées des transformés de e_1 et de e_2 par cette rotation, on trouve :

$$M(\rho(\theta)) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

On remarquera que la matrice de $\rho(\theta)$ ne dépend pas du choix de la base orthonormée.

Matrice d'une symétrie dans $\mathbb{R}^2$. On suppose $\mathbb{R}^2$ muni d'une base (e_1, e_2) . Notons $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la symétrie par rapport à l'axe défini par e_1 parallèlement à e_2 . On a :

$$M(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Matrice d'une projection dans $\mathbb{R}^2$. On suppose $\mathbb{R}^2$ muni d'une base (e_1, e_2) . Notons $\pi_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la projection sur l'axe défini par e_1 parallèlement à e_2 . On a :

$$M(\pi_1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Matrice d'une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une application linéaire. Choisissons une base $B = (e_1, \dots, e_n)$ dans $\mathbb{R}^n$ et la base définie par le vecteur 1 dans $\mathbb{R}$. L'application f est déterminée par les images des vecteurs de la base B qui sont des réels. Donc $M(f)$ est une matrice ligne. Si $f(e_i) = a_i$ pour $1 \leq i \leq n$, on a :

$$M(f) = (a_1 \cdots a_n)$$

Matrice d'une application de $\mathbb{R}$ dans $\mathbb{R}^n$. Voir 8.1.2.

8.6 MATRICE DE LA COMPOSÉE

Ce paragraphe va généraliser le calcul de l'image d'un vecteur vu en 8.1.2.

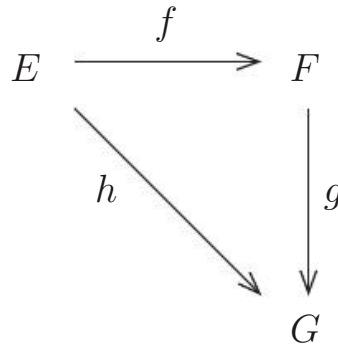
On considère trois espaces vectoriels de dimensions finies :

- E de dimension m avec une base $B = (e_1, \dots, e_m)$,
- F de dimension n avec une base $C = (\varepsilon_1, \dots, \varepsilon_n)$,
- G de dimension p avec une base $D = (\eta_1, \dots, \eta_p)$,

et des applications linéaires $f: E \rightarrow F$ et $g: F \rightarrow G$.

On a vu que l'application composée $h = g \circ f$ est linéaire.

On peut représenter la situation dans le diagramme commutatif suivant :



On pose :

- $R = M(f, B, C) = (r_{jk})_{1 \leq j \leq n, 1 \leq k \leq m}$,
- $S = M(g, C, D) = (s_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$,
- $T = M(h, B, D) = (t_{ik})_{1 \leq i \leq p, 1 \leq k \leq m}$

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & & \vdots \\ r_{n1} & \cdots & r_{nm} \end{pmatrix} \quad S = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & & \vdots \\ s_{p1} & \cdots & s_{pn} \end{pmatrix}$$

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1m} \\ \vdots & & \vdots \\ t_{p1} & \cdots & t_{pm} \end{pmatrix}$$

Quelle relation y a-t-il entre les matrices R , S et T ? Pour le savoir, il suffit de calculer la i -ème coordonnée t_{ik} dans la base D de l'image par h du k -ième vecteur de la base B .

On a :

$$h(e_k) = g(f(e_k)) = g\left(\sum_{1 \leq j \leq n} r_{jk} \varepsilon_j\right)$$

par définition de la matrice R , donc :

$$h(e_k) = \sum_{1 \leq j \leq n} r_{jk} g(\varepsilon_j)$$

par linéarité de g , d'où :

$$h(e_k) = \sum_{1 \leq j \leq n} r_{jk} \left(\sum_{1 \leq i \leq p} s_{ij} \eta_i \right).$$

par définition de la matrice S .

Comme on a aussi : $h(e_k) = \sum_{1 \leq i \leq p} t_{ik} \eta_i$ par définition de la matrice T on trouve finalement :

$$t_{ik} = \sum_{1 \leq j \leq n} s_{ij} r_{jk}$$

puisque la décomposition de $h(e_k)$ sur la base D est unique.

Cette formule est vraiment la formule importante et nous allons expliquer comment l'utiliser. Isolons la i -ième ligne de S et la k -ième colonne de R :

$$\begin{array}{cccc} & & & r_{1k} \\ & & & \vdots \\ s_{i1} & \cdots & s_{in} & \\ & & & \vdots \\ & & & r_{nk} \end{array}$$

On voit que le terme t_{ik} se présente comme la somme des produits des termes de même rang de la ligne et de la colonne. Pour calculer complètement T , il faut calculer mp sommes de produits de ce genre.

Produit de deux matrices. La définition est faite pour correspondre à la composition d'applications linéaires. Mais historiquement, le produit de matrices a été inventé avant que la notion de matrice soit définie. C'était en 1812 et il s'agissait de multiplier des déterminants. Augustin Cauchy et Jacques Binet (d'origine rennaise) donnèrent indépendamment les formules.

On se donne des matrices $R = (r_{jk})_{1 \leq j \leq n, 1 \leq k \leq m}$ et $S = (s_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$, le nombre n de colonnes de S étant égal au nombre de lignes de R . On définit le produit noté SR des matrices S et R comme la matrice $T = (t_{ik})_{1 \leq i \leq p, 1 \leq k \leq m}$ définie par les formules :

$$t_{ik} = \sum_{1 \leq j \leq n} s_{ij} r_{jk}$$

On remarquera que le produit SR de deux matrices S et R n'est défini que lorsque le nombre de colonnes de S est égal au nombre de lignes de R .

Si $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ note l'application linéaire associée à R et $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ l'application linéaire associée à S , les bases choisies dans $\mathbb{R}^m$, $\mathbb{R}^n$, $\mathbb{R}^p$ étant les bases canoniques, le produit SR est donc la matrice de $g \circ f$.

Nous venons de montrer que la matrice de la composée de deux applications linéaires est le produit des matrices de chacune de ces applications, ce qu'on va énoncer plus précisément.

Proposition. *Une fois fixées les bases B, C, D des espaces vectoriels de dimensions finies E, F et G , pour toute application linéaire $f : E \rightarrow F$ et toute application linéaire $g : F \rightarrow G$, on a $M(g \circ f, B, D) = M(g, C, D)M(f, B, C)$. $\square$*

Exemple.

$$\begin{pmatrix} 2 & -3 & 5 \\ a & 0 & 3 \\ -1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 0 \\ -2 & a & 3 \\ 3 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 23 & -9-3a & -4 \\ a+9 & -2a-3 & 3 \\ 7 & 2a-2 & 10 \end{pmatrix}$$

Le premier coefficient de la première ligne du produit des deux matrices s'obtient en considérant la première ligne de la première matrice : $(2, -3, 5)$, la première colonne de la seconde matrice : $(1, -2, 3)$, en les multipliant termes à termes et en faisant la somme : $2 + 6 + 15 = 23$; on procède de même pour calculer les autres coefficients.

Les calculs proposés en exercice peuvent être faits souvent de tête, mais on peut adopter la disposition suivante, qui indique automatiquement quelle ligne et quelle colonne doivent être considérées pour calculer les coefficients du produit. Pour calculer le produit SR de deux matrices, on écrit S , puis on décale R au-dessus de la ligne ; ainsi les coefficients de SR à calculer se trouvent à l'intersection de la ligne de S et de la colonne de R qu'il faut considérer :

$$S = \begin{pmatrix} 2 & -3 & 5 \\ a & 0 & 3 \\ -1 & 2 & 4 \end{pmatrix} \cdots \begin{matrix} \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{matrix} \quad SR = \begin{pmatrix} 23 & -9-3a & -4 \\ a+9 & -2a-3 & 3 \\ 7 & 2a-2 & 10 \end{pmatrix}$$

Image d'un vecteur par une application linéaire. On peut maintenant s'assurer que le calcul de l'image d'un vecteur par une application linéaire fait en 8.1.2 est bien un produit de matrices.

Interprétation matricielle des systèmes linéaires. Soit L un système linéaire :

$$\begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = b_1 & l_1 \\ \vdots & \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = b_n & l_n \end{cases}$$

où les coefficients a_{ij} et b_i sont des réels.

Nous avons vu au chapitre précédent que ce système pouvait s'écrire sous la forme $f(u) = b$ où $f: \mathbb{R}^p \rightarrow \mathbb{R}^n$ est une application linéaire et où $\mathbb{R}^p, \mathbb{R}^n$ sont munis des bases canoniques. Comme la matrice de f est la matrice $A = (a_{ij})$, on peut écrire le système linéaire L sous forme matricielle : $AX = B$ où

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

8.7 PROPRIÉTÉS DU PRODUIT

Matrice unité. Pour toute matrice M ayant n colonnes, on a $MI_n = M$.

Pour toute matrice M ayant n lignes, on a $I_n M = M$.

La première égalité correspond à l'égalité $f \circ \text{id}_E = f$ et la seconde à $\text{id}_F \circ f = f$.

Associativité. Soient E, F et G des espaces vectoriels de dimensions finies et $f: E \rightarrow F, g: F \rightarrow G$ et $h: G \rightarrow H$ des applications linéaires. On sait que la composition d'applications est associative, c'est-à-dire que $h \circ (g \circ f) = (h \circ g) \circ f$. On en déduit que $M(h)(M(g)M(f)) = (M(h)M(g))M(f)$.

Cela suffit pour prouver que le produit de matrices, quand il est défini, est associatif, puisque toute matrice peut être interprétée comme matrice d'une application linéaire.

Compatibilités. Les relations suivantes se déduisent des relations correspondantes pour les applications linéaires (quand les produits ont un sens) :

$$R(R_1 + R_2) = RR_1 + RR_2$$

$$(R_1 + R_2)R = R_1R + R_2R$$

$$R(aS) = (aR)S = a(RS)$$

Inverse d'une matrice. Soit R une matrice carrée d'ordre n . On dit que R est inversible s'il existe une matrice carrée S d'ordre n telle que $SR = RS = I_n$. Si une telle matrice existe, elle est unique car si S' vérifie les mêmes relations que S , on a $S' = S' I_n = S' RS = I_n S = S$. On dit alors que S est la matrice inverse ou l'inverse de R et on pose $S = R^{-1}$. Alors S est inversible et $S^{-1} = R$.

Si R n'est pas carrée, il ne peut exister de matrice S vérifiant $SR = RS = I_n$. Il n'y a donc pas lieu de chercher à inverser R .

Inverse d'un produit. Soient R et S deux matrices carrées d'ordre n inversibles. Alors RS est une matrice inversible et son inverse est $(RS)^{-1} = S^{-1} R^{-1}$. En effet, $S^{-1} R^{-1} RS = S^{-1} IS = I$ et $RSS^{-1} R^{-1} = I$. De même, SR est inversible d'inverse $R^{-1} S^{-1}$.

Matrice de l'inverse. Soit $f : E \rightarrow E$ un automorphisme (voir 7.10) d'espace vectoriel de dimension finie n et soit B une base de E . Notons R et S les matrices de f et de f^{-1} par rapport à B . Comme $f \circ f^{-1} = \text{id}_E$ et que $f^{-1} \circ f = \text{id}_E$, on a $RS = I_n$ et $SR = I_n$. La matrice S est donc l'inverse de R pour la multiplication des matrices.

Proposition. Soit R une matrice carrée d'ordre n . S'il existe une matrice S telle que $SR = I_n$ ou $RS = I_n$, alors R est inversible et $S = R^{-1}$.

Démonstration. Soit $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ les applications linéaires de matrices R et S par rapport à la base canonique. Si $SR = I_n$, on a $g \circ f = \text{id}$, donc f est injective. D'après 7.10, f est donc un isomorphisme d'inverse g . Par conséquent, R est inversible d'inverse S . De même, si $RS = I_n$ on a $f \circ g = \text{id}$, donc f est surjective, donc f est un isomorphisme d'inverse g d'après 7.10, etc. $\square$

Puissances. Pour toute matrice carrée A , on définit $A^2 = AA$, $A^3 = A^2 A = A A^2, \dots$, $A^{n+1} = AA^n = A^n A$ pour tout entier $n \geq 1$ et, quand A est inversible, $A^{-n} = (A^{-1})^n = (A^n)^{-1}$. La propriété habituelle des puissances est vérifiée : pour tous entiers m, n , on a $A^{m+n} = A^m A^n = A^n A^m$.

Commutativité. Étant donnés deux endomorphismes $f, g : E \rightarrow E$, on a vu que $f \circ g$ était en général différent de $g \circ f$. Par conséquent, le produit de matrices n'est pas commutatif, sauf cas particuliers comme, par exemple : I_n commute avec toute matrice carrée d'ordre n , et les puissances d'une même matrice commutent entre elles. Deux matrices diagonales commutent entre elles.

8.8 CALCUL DE L'INVERSE D'UNE MATRICE

Inverse d'une matrice diagonale. L'inverse d'une matrice (a_{ij}) diagonale d'ordre n avec $a_{ii} \neq 0$ pour $1 \leq i \leq n$ est la matrice (a_{ij}^{-1}) . Par exemple, si $a, b, c \neq 0$, l'inverse de

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

est:

$$A^{-1} = \begin{pmatrix} a^{-1} & 0 & 0 \\ 0 & b^{-1} & 0 \\ 0 & 0 & c^{-1} \end{pmatrix}$$

Inverse d'une matrice triangulaire. Montrons qu'une matrice (a_{ij}) triangulaire supérieure d'ordre n , avec $a_{ii} \neq 0$ pour $1 \leq i \leq n$, est inversible. Le cas des matrices triangulaires inférieures est analogue. Pour cela, cherchons l'inverse sous la forme d'une matrice (α_{ij}) triangulaire supérieure. On doit avoir :

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Les coefficients de la première ligne vérifient donc :

$$\begin{cases} \alpha_{11}a_{11} = 1 \\ \alpha_{11}a_{12} + \alpha_{12}a_{22} = 0 \\ \dots \\ \alpha_{11}a_{1n} + \dots + \alpha_{1n}a_{nn} = 0 \end{cases}$$

et on peut les déterminer de proche en proche, puisque les a_{ii} sont tous non nuls. Il en est de même pour les autres lignes.

Multiplication à gauche par une matrice de transvection. Soit A une matrice ayant n lignes et p colonnes. Que se passe-t-il quand on multiplie A à gauche par une matrice de transvection $T_{ij}(\alpha)$ d'ordre n : les lignes du produit obtenu sont les mêmes que celles de A , sauf la i -ième ligne qui est remplacée par la somme de la i -ième ligne de A et de la j -ième ligne de A multipliée par α . Par exemple, en multipliant à gauche par $T_{23}(3)$, on a :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c & d \\ d' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ a' + 3a'' & b' + 3b'' & c' + 3c'' & d' + 3d'' \\ a'' & b'' & c'' & d'' \end{pmatrix}$$

Multiplication à droite par une matrice de transvection. De la même façon, la multiplication à droite par $T_{ij}(\alpha)$ a pour effet d'ajouter à la j -ième colonne de A la i -ième colonne multipliée par α . Par exemple :

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & c + 3b \\ a' & b' & c' + 3b' \\ a'' & b'' & c'' + 3b'' \end{pmatrix}$$

Multiplication à gauche par une matrice de transposition. Que se passe-t-il quand on multiplie maintenant A à gauche par une matrice de transposition P_{ij} d'ordre n ? Les lignes du produit obtenu sont les mêmes que celles de A , sauf la i -ième et la j -ième ligne de A qui sont échangées. Par exemple, en multipliant à gauche par P_{23} , on a :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a & b & c & d \\ d' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ a'' & b'' & c'' & d'' \\ a' & b' & c' & d' \end{pmatrix}$$

Inverses de matrices de transvection et de transposition. L'inverse de la matrice $T_{ij}(\alpha)$ est $T_{ij}(-\alpha)$: en effet, on a $T_{ij}(\alpha) \circ T_{ij}(-\alpha) = I_n$ puisque, dans ce produit, la i -ième ligne de $T_{ij}(-\alpha)$ est remplacée par la somme de la i -ième ligne de $T_{ij}(-\alpha)$ et de la j -ième ligne de $T_{ij}(-\alpha)$ multipliée par α .

L'inverse de P_{ij} est P_{ij} elle-même, puisque le produit par P_{ij} à gauche échange les lignes d'indices i et j de P_{ij} .

Cas général : calcul de l'inverse par résolution d'un système linéaire. Soit $A = (a_{ij})$ une matrice carrée d'ordre n . On peut la considérer comme la matrice d'un endomorphisme f de $\mathbb{R}^n$ muni de la base canonique $(e_1, \dots, e_n)$.

Si A est inversible, pour tout n -uplet $(y_1, \dots, y_n)$ de matrice Y , il existe un n -uplet $(x_1, \dots, x_n)$ de matrice X tel que $AX = Y$. Cette égalité matricielle définit un système linéaire. La résolution de ce système par la méthode du pivot de Gauss conduit à une relation de la forme $X = A_1 Y$ où A_1 est une matrice carrée d'ordre n car les différentes opérations de la méthode du pivot donnent pour les membres de droite des combinaisons linéaires des y_i . On a donc $Y = AX = AA_1 Y$ pour tout $(y_1, \dots, y_n)$ de $\mathbb{R}^n$. Comme cette relation est vraie pour tous les vecteurs de base, on en déduit $AA_1 = I_n$, ce qui prouve que A est inversible d'inverse A_1 .

Le plus souvent, on ne sait pas *a priori* si la matrice A est inversible et c'est la résolution du système $AX = Y$ qui montre si le système est de rang n ou non, c'est-à-dire si A est inversible ou non.

Exemple de calcul d'inverse. Considérons la matrice

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 3 & -1 & 1 \\ -2 & 1 & -3 \end{pmatrix}$$

Le système linéaire associé à l'égalité $AX = Y$ s'écrit :

$$\begin{cases} x + 3y - z = y_1 & l_1 \\ 3x - y + z = y_2 & l_2 \\ -2x + y - 3z = y_3 & l_3 \end{cases}$$

dont les premiers membres sont ceux de l'exemple 1 de 4.8. Les mêmes calculs que dans cet exemple : $l'_2 = l_2 - 3l_1$, $l'_3 = l_3 + 2l_1$, $l''_3 = l'_3 + \frac{7}{10}l'_2$ montrent que A est inversible. En continuant les calculs, on trouve :

$$\begin{cases} x = \frac{1}{11}(y_1 + 4y_2 + y_3) \\ y = \frac{1}{22}(7y_1 - 5y_2 - 4y_3) \\ z = \frac{1}{22}(y_1 - 7y_2 - 10y_3) \end{cases}$$

Par conséquent :

$$A^{-1} = \frac{1}{22} \begin{pmatrix} 2 & 8 & 2 \\ 7 & -5 & -4 \\ 1 & -7 & -10 \end{pmatrix}$$

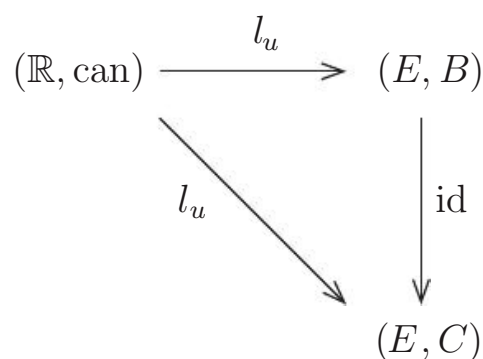
ce qu'un logiciel comme *Sage* donne instantanément. Si on fait les calculs à la main, il vaut mieux vérifier que le résultat trouvé multiplié par A donne I_n ; on n'a pas trop à se fatiguer : d'après la proposition de 8.7, un seul côté suffit.

Interprétation matricielle. La résolution du système $AX = Y$ utilise, nous l'avons dit au chapitre 4, uniquement les deux opérations de transposition de deux lignes ou d'ajout de a fois une ligne à une autre ligne. Ces deux opérations correspondent à des multiplications à gauche par des matrices de transposition ou de transvection. La résolution du système linéaire $AX = Y$ correspond donc à une suite de multiplications à gauche par ces matrices. Notons S le produit de ces dernières. D'après les calculs de 4.8, SA est une matrice diagonale D dont les coefficients de la diagonale sont non nuls. On a $SA = D$ donc $A^{-1} = D^{-1} S$. La matrice inverse de A s'écrit donc comme un produit de matrices de transvection, de transposition et d'une matrice diagonale. Notons qu'on peut calculer S en effectuant sur les lignes de I exactement les mêmes opérations qu'on effectue sur les lignes de A pour la ramener à la forme diagonale.

8.9 CHANGEMENT DE BASE

Changement de base pour un vecteur. Quand on parle d'un vecteur d'un espace vectoriel, on parle d'un objet précis : le vecteur u de l'espace E . Mais quand on veut faire des calculs avec une base B de E , on calcule avec les coordonnées du vecteur u dans cette base B . Si on veut conduire les calculs dans une base C , le vecteur u n'a pas changé, mais on doit calculer avec les coordonnées de u dans la base C . Quelle est la relation entre ces coordonnées ?

Considérons le diagramme suivant :



où l'application identique est marquée avec une flèche pour indiquer l'espace de départ et l'espace d'arrivée et où les bases pour chaque espace sont indiquées dans les parenthèses.

La commutativité du diagramme précédent donne l'égalité matricielle :

$$M(l_u, \text{can}, C) = M(\text{id}, B, C)M(l_u, \text{can}, B).$$

Les matrices $U = M(l_u, \text{can}, B)$ et $V = M(l_u, \text{can}, C)$ sont des matrices colonnes, dont les coefficients sont les coordonnées du vecteur u dans les bases B et C respectivement. La matrice $P = M(\text{id}, B, C)$ est la matrice de l'identité, mais attention, ce n'est pas une raison pour que ce soit la matrice unité I_n : ses colonnes sont les images des vecteurs de la base B par l'identité, donc les vecteurs de la base B eux-mêmes, exprimés dans la base C . On a donc

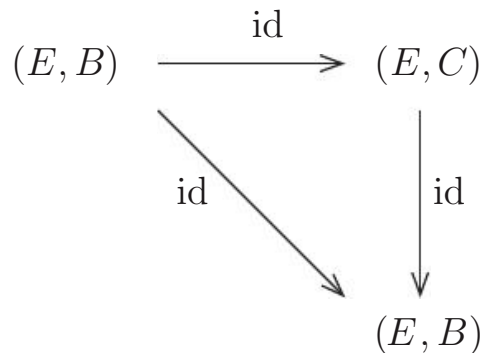
$$V = PU.$$

Le rôle joué par la matrice P nous conduit à la définition suivante.

Matrice de passage. Soit E un espace vectoriel de dimension finie muni de deux bases B et C . Les matrices de l'identité $M(\text{id}, B, C)$ et $M(\text{id}, C, B)$ sont appelées matrices de passage. La lettre P est souvent utilisée pour noter une matrice de passage.

L'usage est d'appeler matrice de passage de la base C à la base B (dans cet ordre) la matrice $M(\text{id}, B, C)$, et matrice de passage de la base B à la base C la matrice $M(\text{id}, C, B)$. La matrice $M(\text{id}, B, C)$ a pour colonnes les vecteurs de la base B exprimés dans la base C et la matrice $M(\text{id}, C, B)$ a pour colonnes les vecteurs de la base C exprimés dans la base B .

Relation entre matrices de passage. Le diagramme évidemment commutatif suivant :



donne l'égalité matricielle :

$$M(\text{id}, C, B)M(\text{id}, B, C) = M(\text{id}, B, B) = I_n.$$

Par conséquent, les deux matrices de passage $P = M(\text{id}, C, B)$ et $Q = M(\text{id}, B, C)$ sont inverses l'une de l'autre.

Exemple de changement de base pour une application linéaire. Dans $\mathbb{R}^2$ muni de sa base canonique (e_1, e_2) , posons $u = (3, 1)$ et $v = (5, 2)$. Les vecteurs u et v forment une base B de $\mathbb{R}^2$. Considérons alors l'application linéaire de $\mathbb{R}^2$ dans lui-même définie par :

$$f(u) = 2u,$$

$$f(v) = -v.$$

Par rapport à la base B , f a pour matrice une matrice diagonale :

$$A' = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

Déterminons maintenant la matrice de f par rapport à la base canonique. Comme :

$$u = 3e_1 + e_2$$

$$v = 5e_1 + 2e_2$$

on a :

$$e_1 = 2u - v$$

$$e_2 = 3v - 5u.$$

On en déduit :

$$f(e_1) = f(2u - v) = 2f(u) - f(v) = 4u + v = (17, 6)$$

$$f(e_2) = f(3v - 5u) = 3f(v) - 5f(u) = -3v - 10u = (-45, -16).$$

La matrice de f par rapport à la base canonique est donc :

$$A = \begin{pmatrix} 17 & -45 \\ 6 & -16 \end{pmatrix}$$

Il est clair que la base B est plus adaptée aux calculs sur l'application f . Montrons-le sur deux exemples.

1) La matrice de f^5 dans la base B est facile à calculer :

$$A'^5 = \begin{pmatrix} 32 & 0 \\ 0 & -1 \end{pmatrix}$$

mais il faut du temps pour calculer dans la base canonique :

$$A^5 = \begin{pmatrix} 197 & -495 \\ 66 & -166 \end{pmatrix}$$

2) La matrice de f^{-1} dans la base B est :

$$A'^{-1} = \begin{pmatrix} 0,5 & 0 \\ 0 & -1 \end{pmatrix}$$

alors que l'inverse de A est :

$$A^{-1} = \begin{pmatrix} 8 & -22,5 \\ 3 & -8,5 \end{pmatrix}$$

Formules de changement de base. Pour terminer cet exemple, nous allons montrer que A' et A sont liées par une relation matricielle.

Pour cela, considérons les matrices de passages $P = M(\text{id}_{\mathbb{R}^2}, B, \text{can})$ et $Q = M(\text{id}_{\mathbb{R}^2}, \text{can}, B)$. On a :

$$P = \begin{matrix} & \begin{matrix} u & v \end{matrix} \\ \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} & \begin{matrix} e_1 \\ e_2 \end{matrix} \end{matrix} \quad Q = \begin{matrix} & \begin{matrix} e_1 & e_2 \end{matrix} \\ \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} & \begin{matrix} u \\ v \end{matrix} \end{matrix}$$

Comme cela a déjà été souligné, il faut bien voir (voir le slogan) que, dans ces deux matrices, les colonnes sont formées des images des vecteurs d'une base par l'application linéaire identique, c'est-à-dire eux-mêmes, mais exprimés dans une autre base. Quand on connaît les expressions de u et v dans la base canonique, la matrice P est immédiate. On a vu que $Q = P^{-1}$, ce qui permet de calculer Q connaissant P .

Considérons alors le diagramme :

$$\begin{array}{ccc} (\mathbb{R}^2, \text{can}) & \xrightarrow{f} & (\mathbb{R}^2, \text{can}) \\ \text{id}_{\mathbb{R}^2} \uparrow & & \downarrow \text{id}_{\mathbb{R}^2} \\ (\mathbb{R}^2, B) & \xrightarrow{f} & (\mathbb{R}^2, B) \end{array}$$

Dans ce diagramme, on a deux chemins possibles en partant du coin en bas à gauche : le premier en partant horizontalement donne f , et le second, en passant par le haut, donne $\text{id}_{\mathbb{R}^2} \circ f \circ \text{id}_{\mathbb{R}^2}$. La matrice correspondant au premier chemin est A' , et celle correspondant au second est $QAP = P^{-1}AP$. Comme $\text{id}_{\mathbb{R}^2} \circ f \circ \text{id}_{\mathbb{R}^2} = f$, le diagramme est commutatif et on a donc la formule :

$$A' = P^{-1}AP.$$

En multipliant les deux termes de cette formule à gauche par P et à droite par P^{-1} , on obtient :

$$A = PA'P^{-1}.$$

Cette seconde formule se déduit aussi du diagramme commutatif :

$$\begin{array}{ccc} (\mathbb{R}^2, B) & \xrightarrow{f} & (\mathbb{R}^2, B) \\ \text{id}_{\mathbb{R}^2} \uparrow & & \downarrow \text{id}_{\mathbb{R}^2} \\ (\mathbb{R}^2, \text{can}) & \xrightarrow{f} & (\mathbb{R}^2, \text{can}) \end{array}$$

Nous venons d'établir, sur un exemple, deux formules essentielles.

Base adaptée. La recherche d'une base bien adaptée au problème qu'on étudie est un thème important de l'algèbre linéaire. Si une base n'est pas bien adaptée, une application linéaire très simple peut avoir une matrice où il est impossible de voir cette simplicité. Nous verrons plus tard, au chapitre 15, comment trouver de bons changements de base.

Changement de base : cas général. Il n'y a pas grand-chose à faire pour passer au cas général. On considère deux espaces vectoriels de dimension finie E et F , deux bases B et B' dans E , deux bases C et C' dans F et une application linéaire $f: E \rightarrow F$. On pose $A = M(f, B, C)$, $A' = M(f, B', C')$, $P = M(\text{id}_E, B', B)$ et $Q = M(\text{id}_F, C', C)$. Le diagramme commutatif :

$$\begin{array}{ccc} (E, B) & \xrightarrow{f} & (F, C) \\ \text{id}_E \uparrow & & \downarrow \text{id}_F \\ (E, B') & \xrightarrow{f} & (F, C') \end{array}$$

donne la formule :

$$A' = Q^{-1}AP.$$

On peut même mentionner les matrices dans le diagramme pour ne pas se tromper :

$$\begin{array}{ccc} (E, B) & \xrightarrow{f, A} & (F, C) \\ \text{id}_E, P \uparrow & & \downarrow \text{id}_F, Q^{-1} \\ (E, B') & \xrightarrow{f, A'} & (F, C') \end{array}$$

ce qui donne encore la formule :

$$A' = Q^{-1}AP.$$

Souvent, $E = F$, $B = C$ et $B' = C'$. La situation est celle du diagramme suivant :

$$\begin{array}{ccc} (E, B) & \xrightarrow{f} & (E, B) \\ \text{id}_E \uparrow & & \downarrow \text{id}_E \\ (E, B') & \xrightarrow{f} & (E, B') \end{array}$$

Dans ce cas, $Q = P$ et on retrouve les formules de l'exemple :

$$\begin{aligned} A' &= P^{-1}AP, \\ A &= PA'P^{-1}. \end{aligned}$$

On en déduit :

$$A^n = PA'^n P^{-1},$$

formule encore valable pour tout n de $\mathbb{Z}$ quand f est inversible, car alors A et A' sont inversibles et :

$$A^{-1} = PA'^{-1}P^{-1}.$$

Nous avons déjà dit que ces formules sont essentielles. Pour s'en souvenir, il est préférable de se souvenir du diagramme dont elles proviennent plutôt que d'apprendre par cœur des formules où le risque d'erreur n'est pas nul.

Matrices semblables. Deux matrices carrées A et A' d'ordre n sont dites semblables s'il existe une matrice carrée P d'ordre n inversible telle que $A' = P^{-1}AP$. La relation de similitude entre matrices est une relation d'équivalence.

De façon équivalente, A et A' sont semblables, s'il existe une application linéaire $f: E \rightarrow E$ et des bases B et B' de E telles que $A = M(f, B, B)$, $A' = M(f, B', B')$; en effet, si on pose $P = M(\text{id}, B', B)$, on a $A' = P^{-1}AP$; réciproquement, si $f: E \rightarrow E$ est une application linéaire de matrice A par rapport à une base B , sa matrice sera A' par rapport à la base B' dont les vecteurs sont donnés dans la base B par les colonnes de P .

8.10 RANG ET TRACE

Rang d'une matrice. Nous avons défini au chapitre précédent le rang d'une application linéaire $f: E \rightarrow F$: c'est la dimension de $\text{Im}(f)$. Pour toute base B de E et toute base C de F , c'est donc la dimension de l'espace vectoriel engendré par les vecteurs colonnes de $M(f, B, C)$.

Définition. Soit M une matrice de n lignes et de p colonnes. On appelle rang de M et on note $\text{rg}(M)$, le rang de la famille des p vecteurs colonnes de M .

Le calcul du rang d'une matrice peut donc s'effectuer en appliquant la méthode du pivot aux vecteurs colonnes de la matrice.

Propriétés du rang. Soit M une matrice. Pour toute application linéaire $f: E \rightarrow F$, toute base B de E , toute base C de F , si $M = M(f, B, C)$, alors $\text{rg}(f) = \text{rg}(M)$.

Deux matrices semblables ont même rang, autrement dit, le rang est invariant par changement de base.

Trace d'une matrice. Soit $A = (a_{ij})$ une matrice carrée d'ordre n . On appelle trace de A , et on note $\text{Tr}(A)$, la somme des coefficients diagonaux de A :

$$\text{Tr}(A) = \sum_{1 \leq i \leq n} a_{ii}$$

Propriétés de la trace. Soit $B = (b_{ij})$ une seconde matrice carrée d'ordre n . Un calcul simple montre que :

$$\text{Tr}(AB) = \text{Tr}(BA)$$

En effet, les deux membres sont égaux à $\sum_{1 \leq i, j \leq n} a_{ij}b_{ji}$.

On en déduit que, si deux matrices A et A' sont semblables, elles ont la même trace, puisque, si $A' = P^{-1}AP$, on a $\text{Tr}(A') = \text{Tr}(P^{-1}AP) = \text{Tr}(APP^{-1}) = \text{Tr}(A)$. Comme le rang, la trace est invariante par changement de base.

Enfin, si f est un endomorphisme d'un espace vectoriel E de dimension finie, les matrices de f dans les différentes bases de E sont semblables. Elles ont la même trace et cette valeur commune est appelée trace de l'application linéaire f et notée $\text{Tr}(f)$.

Exemples.

- 1) La trace de I_n est n . Si E est un espace de dimension finie n , la trace de l'application identique de E est n .
- 2) La trace de la symétrie $(x, y) \mapsto (x, -y)$ de $\mathbb{R}^2$ est nulle.
- 3) La trace de la rotation d'angle θ dans le plan vectoriel $\mathbb{R}^2$ est $2 \cos \theta$.
- 4) La trace de la rotation d'angle θ autour de l'axe des z dans l'espace $\mathbb{R}^3$ est $1 + 2 \cos \theta$.

► Vers le chapitre 9

Maintenant que nous avons mis au point des outils pour les calculs, nous revenons aux espaces vectoriels pour étudier un peu mieux comment ils sont faits et comment en construire d'autres à l'aide des notions de produit et de somme d'espaces vectoriels.

Exercices

8.1 On considère les matrices suivantes :

$$A = \begin{pmatrix} 0 & 2 \\ 1 & 4 \\ 3 & -2 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 & -1 & 0 & 7 \\ 0 & 1 & -2 & 1 & 2 \\ 4 & 0 & 1 & -1 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 0 & 3 & -1 \end{pmatrix}$$

- a)** Pour chacune de ces matrices, décrire l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ associée : images des vecteurs de la base canonique, image d'un vecteur quelconque.
- b)** Peut-on former les produits ABC , CBA , BAC ? Si oui, les calculer de deux manières pour vérifier, sur ce cas particulier, l'associativité du produit de matrices.

8.2 Petites équations

Trouver les matrices carrées d'ordre 2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

telles que :

a) $A^2 = A$;

b) $A^2 = I_2$;

c) $AB = BA$, avec $B = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$

8.3 Calculs de puissances

Soit a un nombre réel non nul et soit

$$A = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$$

Calculer A^n pour tout n de $\mathbb{Z}$.

8.4 Calculs d'inverses

Calculer les inverses des matrices suivantes.

a)

$$A_1 = \begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix} \quad A_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

b)

$$A_3 = \begin{pmatrix} 1 & 5 & -3 \\ 2 & 11 & 1 \\ 2 & 9 & -11 \end{pmatrix} \quad A_4 = \begin{pmatrix} 1 & 5 & -3 \\ 2 & 11 & 1 \\ 1 & 4 & -10 \end{pmatrix}$$

c) Utilisation d'une relation

On pose

$$A = \begin{pmatrix} 2 & 5 & -3 \\ 2 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix}$$

Montrer que A vérifie la relation $A^3 - 2A^2 - 5A - 24I = 0$. En déduire l'inverse de A .

8.5 Matrices nilpotentes

On appelle *nilpotentes* les matrices carrées dont une puissance est nulle.

a) On pose

$$N_1 = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$$

Calculer les puissances de N_1 ; N_1 est-elle inversible ?

b) On pose :

$$N_2 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

Calculer les puissances de N_2 ; N_2 est-elle inversible ?

c) Montrer qu'une matrice triangulaire supérieure est nilpotente si et seulement si ses coefficients diagonaux sont tous nuls.

d) Soit N une matrice carrée d'ordre n nilpotente. Montrer que $I + N$ est une matrice inversible et donner son inverse ; on pourra s'inspirer de la formule $1 - x^n = (1 - x)(1 + x + \dots + x^{n-1})$.

e) Calculer les inverses de $I + N_1$, $I + N_2$ et de :

$$A = \begin{pmatrix} 1 & a & a^2 & a^3 \\ 0 & 1 & a & a^2 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

avec la méthode générale de calcul de l'inverse et avec la formule précédente.

8.6 Changements de base

1) On considère une application linéaire f de matrice

$$A = \begin{pmatrix} 9 & -6 & 10 \\ -5 & 2 & -5 \\ -12 & 6 & -13 \end{pmatrix}$$

par rapport à la base canonique de $\mathbb{R}^3$.

- a)** Montrer que les vecteurs $u_1 = (2, -1, -2)$, $u_2 = (1, 0, -1)$ et $u_3 = (-2, 1, 3)$ forment une base $B = (u_1, u_2, u_3)$ de $\mathbb{R}^3$.
 - b)** Calculer les matrices de passage $P = M(\text{id}, B, \text{can})$ et $Q = M(\text{id}, \text{can}, B)$.
 - c)** Calculer la matrice de f dans la base B .
- 2)** On note $\Delta : \mathbb{R}_4[X] \rightarrow \mathbb{R}_4[X]$ l'application définie par $\Delta(P)(X) = P(X+1) - P(X)$.
- a)** Déterminer la matrice A de Δ dans la base $(1, X, X^2, X^3, X^4)$.
 - b)** Déterminer la matrice A' de Δ dans la base $(1, X, X(X-1), X(X-1)(X-2), X(X-1)(X-2)(X-3))$.
 - c)** Déterminer les matrices de passage entre les deux bases de $\mathbb{R}_4[X]$. Quelle est la relation entre A et A' ?

3) a) Montrer que les deux matrices triangulaires :

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

sont semblables, ainsi que les matrices :

$$C = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

b) Montrer que les matrices AB et BA sont semblables si l'une d'entre elle est inversible. Donner des exemples de matrices carrées non semblables.

8.7 Application des changements de base

a) On considère la matrice

$$A = \begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix}$$

comme matrice d'un endomorphisme f de $\mathbb{R}^2$ muni de la base canonique. Quelle est la matrice A' de f par rapport à la base (u_1, u_2) définie par $u_1 = (3, 1)$, $u_2 = (2, 1)$? Quelle est la relation entre A et A' ?

b) On note R la relation de récurrence $u_{n+2} = 5u_{n+1} - 6u_n$. On pose $u_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$.

Exprimer U_{n+1} en fonction de U_n . Déterminer $S(R)$ à l'aide du changement de base précédent.

c) On note E l'équation différentielle $f'' - 5f' + 6f = 0$. Exprimer $\begin{pmatrix} f'' \\ f' \end{pmatrix}$ en fonction de $\begin{pmatrix} f' \\ f \end{pmatrix}$. Déterminer $S(E)$ à l'aide du changement de base précédent.

8.8 Multiplication rapide de matrices

Comment mesurer l'intérêt d'un algorithme : en comptant le nombre d'opérations nécessitées par son exécution. Pour les produits de matrices, les additions de nombres sont bien moins coûteuses que les multiplications d'entiers et on ne les compte pas. Il y a *a priori* n^3 multiplications pour calculer le produit de deux matrices. De nouvelles

méthodes ont été proposées depuis 1969 (la première par Volker Strassen, né en 1936), intéressantes quand n est suffisamment grand, mais elles demandent à être évaluées : diminuer le nombre de multiplications dans un produit de matrices est intéressant, mais différentes difficultés peuvent rendre une méthode sans intérêt, comme des constantes intervenant dans le calcul de la complexité qui peuvent se révéler bien trop grandes ou la nécessité de stockages intermédiaires de gros volumes d'informations.

Dans le même ordre d'idées, un algorithme de multiplication rapide de polynômes a été mis au point en 1960 par Anatole Karatsuba (1937-2008) ; il réduit le nombre de multiplications nécessaires pour calculer un produit de deux polynômes de degrés n de n^2 à $n^{1,6}$ environ.

1) Cas de la multiplication de matrices carrées d'ordre 2

Posons $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, $C = AB = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$.

a) Déterminer le nombre d'additions (ou soustractions) et le nombre de multiplications à effectuer pour calculer C , puis pour calculer le produit de deux matrices de taille n , avec la méthode présentée en 8.6.

b) On pose

$$\begin{aligned} E &= (a_{11} + a_{22})(b_{11} + b_{22}), F = (a_{21} + a_{22})b_{11}, \\ G &= a_{11}(b_{12} - b_{22}), H = a_{22}(b_{21} - b_{11}), J = (a_{11} + a_{12})b_{22}, \\ K &= (a_{21} - a_{11})(b_{11} + b_{12}), L = (a_{12} - a_{22})(b_{21} + b_{22}). \end{aligned}$$

Vérifier que

$$\begin{aligned} c_{11} &= E + H - J + L, c_{12} = G + J, \\ c_{21} &= F + H, c_{22} = E + G - F + K. \end{aligned}$$

c) Déterminer le nombre d'addition et le nombre de multiplication à effectuer pour calculer C avec cette seconde méthode.

2) Cas de la multiplication de matrices carrées d'ordre 2^k

Les formules précédentes peuvent s'appliquer au produit de deux matrices de taille 2^k en les décomposant en blocs de quatre matrices de taille 2^{k-1} (c'est diviser pour régner, dit-on), les matrices de ces produits pouvant elles-mêmes être décomposées en produit de quatre matrices de taille 2^{k-2} , etc. (les formules ne posent pas de problème de commutativité). Déterminer le nombre d'additions u_k et le nombre de multiplications v_k à effectuer pour un produit de deux matrices carrées de taille 2^k avec cette seconde méthode. Comparer le nombre de multiplications à celui de la méthode de 8.6. pour $k = 10$.

3) Cas de la multiplication de matrices carrées de taille quelconque

Expliquer comment se ramener au cas précédent.

Solutions

8.1 a) Notons $(e_1, e_2, e_3, e_4, e_5), (e'_1, e'_2, e'_3, e'_4), (e''_1, e''_2, e''_3)$ et (e'''_1, e'''_2) les bases canoniques de $\mathbb{R}^5, \mathbb{R}^4, \mathbb{R}^3, \mathbb{R}^2$ respectivement.

A, B et C sont les matrices des applications $f: \mathbb{R}^2 \rightarrow \mathbb{R}^5, g: \mathbb{R}^5 \rightarrow \mathbb{R}^3, h: \mathbb{R}^4 \rightarrow \mathbb{R}^2$, en définissant ces applications par les images des vecteurs de base et en définissant ces images comme les colonnes de ces matrices. Par exemple, $h(e'_1) = e'''_1 + 2e'''_2, g(e_3) = -e''_1 - 2e''_2 + e''_3$, etc. On trouve aussi : $f(x, y) = (2y, x + 4y, 3x - 2y, 5x, x + y)$, etc.

b) Compte tenu des dimensions, la seule composée possible est $g \circ f \circ h$, ce qui conduit à calculer $B(AC)$ et $(BA)C$. On doit trouver le même résultat :

$$\begin{pmatrix} 44 & 12 & 51 & -1 \\ 22 & 4 & 28 & -4 \\ 13 & -2 & 22 & -10 \end{pmatrix}$$

8.2 a) $A^2 = A$ donne un système de quatre équations :

$$\begin{cases} a^2 + bc = a \\ b(a + d) = b \\ c(a + d) = c \\ bc + d^2 = d \end{cases}$$

Le système se résout en étudiant d'abord trois cas particuliers : $b = c = 0$; $b = 0, c \neq 0$; $b \neq 0, c = 0$. Quand on n'est pas dans ces cas, on peut se ramener aux équations $a^2 + bc = a$ et $a + d = 1$: les solutions sont de la forme :

$$A = \begin{pmatrix} a & b \\ \frac{a - a^2}{b} & 1 - a \end{pmatrix}.$$

b) $A^2 = I_2$ donne un système de quatre équations :

$$\begin{cases} a^2 + bc = 1 \\ b(a + d) = 0 \\ c(a + d) = 0 \\ bc + d^2 = 1 \end{cases}$$

qui se résout en distinguant des cas particuliers comme ci-dessus. On trouve : $b = 0, c = 0, a = \pm 1, d = \pm 1$; $b = 0, c \neq 0, a = -d = \pm 1$; $a = -d, bc = 1 - a^2$, etc.

c) $AB = BA$ donne un système de 4 équations qui montre que les matrices commutant

avec B sont de la forme : $\begin{pmatrix} b + d & b \\ -b & d \end{pmatrix}$

8.3 Par récurrence et après calcul de l'inverse de A , on trouve que, pour tout n de

$$\mathbb{Z} : A^n = \begin{pmatrix} a^n & na^{n-1} \\ 0 & a^n \end{pmatrix}.$$

8.4 a) $A_1^{-1} = \frac{1}{22} \begin{pmatrix} 5 & 3 \\ -4 & 2 \end{pmatrix}$

Pour $ad - bc \neq 0$, A_2 est inversible et $A_2^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

b) $A_3^{-1} = \begin{pmatrix} -65 & 14 & 19 \\ 12 & -5/2 & -7/2 \\ -2 & 1/2 & 1/2 \end{pmatrix}.$

La matrice A_4 n'est pas inversible.

c) La relation s'écrit : $A \left(\frac{1}{24} (A^2 - 2A - 5I) \right) = I$, ce qui donne

$$A^{-1} = \frac{1}{24} (A^2 - 2A - 5I) \text{ d'où, après calculs, } A^{-1} = \frac{1}{24} \begin{pmatrix} -1 & 5 & 8 \\ 4 & 4 & -8 \\ -2 & 10 & -8 \end{pmatrix}.$$

Une erreur fréquente est d'oublier le I en factorisant A dans $-5A$.

8.5 a) et b) $N_1^2 = 0$, $N_2^3 = 0$ montrent que N_1 et N_2 ne sont pas inversibles.

c) Soit N une matrice triangulaire supérieure d'ordre n .

Si les termes sur la diagonale sont tous nuls, on a $N^n = 0$; en effet, si f est une application linéaire de matrice N par rapport à la base $(e_1, \dots, e_n)$, on a $f(e_k) \in \text{Vect}(e_1, \dots, e_{k-1})$ pour tout $k \geq 1$, etc.

S'il existe i tel que $a_{ii} \neq 0$, a_{ii}^n est le terme de mêmes indices de N^n , donc N n'est pas nilpotente.

d) Comme N commute avec I et avec les puissances de N , on a :

$$(I - N)(I + N + N^2 + \dots + N^{n-1}) = I - N^n = I ; \text{ donc l'inverse de } I + N \text{ est } I - N + N^2 + \dots + (-1)^{n-1} N^{n-1}.$$

e) $(I + N_1)^{-1} = I - N_1 = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$

$$(I + N_2)^{-1} = I - N_2 + N_2^2 = \begin{pmatrix} 1 & -a & -b + ac \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}$$

$$A^{-1} = I - N + N^2 - N^3 = \begin{pmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -a & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ avec } N = \begin{pmatrix} 0 & a & a^2 & a^3 \\ 0 & 0 & a & a^2 \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

8.6 1) b) Les colonnes de la matrice P sont données par les vecteurs de B ; pour obtenir Q , il faut inverser P . On trouve :

$$P = \begin{pmatrix} 2 & 1 & -2 \\ -1 & 0 & 1 \\ -2 & -1 & 3 \end{pmatrix} \quad Q = P^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

c) On trouve $P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ en n'oubliant pas le conseil d'écrire le dia-

gramme pour retrouver la formule.

2) a) $A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

b) $A' = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

c) On a $P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 2 & -6 \\ 0 & 0 & 1 & -3 & 11 \\ 0 & 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

et $A' = P^{-1}AP$.

3) a) Pour les matrices A et B , on peut chercher une matrice triangulaire supérieure P telle que $AP = PB$ et résoudre le système linéaire que doivent satisfaire les coefficients de cette matrice ; on trouve la famille des matrices de la forme

$$P = \begin{pmatrix} a & b & c & d \\ 0 & 2a & 3a+2b & 4a+3b+2c \\ 0 & 0 & 4a & 12a+4b \\ 0 & 0 & 0 & 8a \end{pmatrix} \text{ avec } a, b, c, d \in \mathbb{R}.$$

Si C est la matrice de l'application linéaire f dans la base $(e_1, \dots, e_n)$, D est la matrice de f dans la base $(e_n, e_{n-1}, \dots, e_1)$.

b) Si A est inversible, on a $AB = A(BA)A^{-1}$. Dans $M_n(\mathbb{R})$, les matrices αI , $\alpha \in \mathbb{R}$, qui commutent avec toutes les matrices, ne sont semblables qu'à elle-mêmes. Deux matrices de rang différent ne sont pas semblables entre elles.

8.7 a) On pose $P = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$ d'où $P^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$ et $A' = P^{-1}AP = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ (faire le diagramme).

b) La relation s'écrit $U_{n+1} = AU_n$ d'où $U_{n+1} = PA'P^{-1}U_n$, $P^{-1}U_{n+1} = A'P^{-1}U_n$. En posant $V_n = P^{-1}U_n = \begin{pmatrix} v_n \\ w_n \end{pmatrix}$, on voit que la suite V_n vérifie $V_{n+1} = A'V_n$; par conséquent, les suites (v_n) et (w_n) vérifient $v_{n+1} = 3v_n$ et $w_{n+1} = 2w_n$ d'où $V_n = \begin{pmatrix} c3^n \\ d2^n \end{pmatrix}$ avec c et d réels. Comme $U_n = PV_n$, on obtient $u_n = c3^n + d2^n$, le résultat attendu.

c) Posons $F = \begin{pmatrix} f' \\ f \end{pmatrix}$ et $F' = \begin{pmatrix} f'' \\ f' \end{pmatrix}$. L'équation différentielle s'écrit $F' = AF$ d'où $F' = PA'P^{-1}F$, $P^{-1}F' = A'P^{-1}F$. En posant $G = P^{-1}F = \begin{pmatrix} g \\ h \end{pmatrix}$, on voit que G vérifie

$G' = A'G$; par conséquent, on a $g' = 3g$ et $h' = 2h$, d'où $G = \begin{pmatrix} x \mapsto ce^{3x} \\ x \mapsto de^{2x} \end{pmatrix}$ avec c et d réels. Comme $F = PG$, on obtient $f: x \mapsto ce^{3x} + de^{2x}$. Comme on connaît l'ensemble des solutions des équations différentielles satisfaites par g et h , on est sûr d'avoir obtenu l'ensemble des solutions de l'équation différentielle proposée.

8.8 1) a) On a :

$$c_{11} = a_{11}b_{11} + a_{12}b_{21},$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22},$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21},$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22}.$$

Le nombre d'additions dans ce calcul est 4 et le nombre de multiplications est 8 ; dans le cas général, ces nombres sont respectivement $n^2(n-1)$ et n^3 .

c) Une fois les vérifications du **b)** achevées, on constate que la nouvelle méthode demande 18 additions et 7 multiplications.

2) La décomposition de la matrice en 4 blocs montre que $v_k = 7v_{k-1}$, d'où $v_k = 7^k = n^{\log_2 7}$. Comme $\log_2 7 \approx 2,807$, on a un gain sur l'exposant. La multiplication de deux matrices de taille $n = 2^{10} = 1024$, demande près de 282 millions de multiplications contre plus d'un milliard avec la méthode de 8.6.

La formule de récurrence pour le nombre d'additions est $u_k = 7u_{k-1} + (9/2)4^k$. C'est une relation de récurrence linéaire qu'on peut résoudre avec les méthodes du chapitre 2 : $u_k = 6 \times 7^k - 6 \times 4^k = 7^k (6 - (4/7)^k)$. On voit que $u_k < 6 \times 7^k$; et on sait que les additions sont moins coûteuses en temps que les multiplications.

3) Pour des matrices carrées dont la taille n n'est pas une puissance de 2, on se ramène à des matrices carrées de taille la plus petite puissance de 2 supérieure à n en complétant les matrices par des 0.

SOMMES DIRECTES, PRODUITS, QUOTIENTS

9

9.1 EXEMPLES

Exemple 1. Considérons dans l'espace vectoriel $\mathbb{R}^3$, un sous-espace vectoriel F de dimension 1 et un sous-espace vectoriel G de dimension 2. Supposons que F ne soit pas contenu dans G . Dans ce cas, on a $F \cap G = \{0\}$; si e_1 note un vecteur non nul de F , e_2, e_3 des vecteurs non colinéaires de G , il est clair que (e_1, e_2, e_3) est une base de $\mathbb{R}^3$. On a donc $F + G = \mathbb{R}^3$.

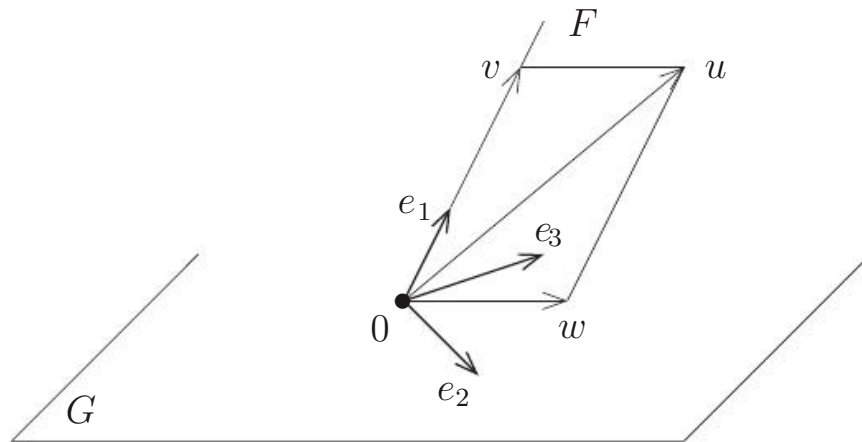
Soit u un vecteur de $\mathbb{R}^3$. Il existe une unique écriture de u dans la base (e_1, e_2, e_3) : $u = xe_1 + ye_2 + ze_3$. Comme xe_1 est un vecteur de F et $ye_2 + ze_3$ un vecteur de G , on peut donc écrire $u = v + w$ avec $v \in F$ et $w \in G$; cette écriture est également unique. Son intérêt est qu'elle donne une expression des vecteurs de l'espace $\mathbb{R}^3$ en fonction des vecteurs de deux sous-espaces sans référence à des bases de ceux-ci.

On peut utiliser le dessin ci-après pour représenter la situation (voir page suivante).

Cette façon de voir un espace vectoriel comme décomposé en somme de sous-espaces est le premier sujet de ce chapitre.

Exemple 2. Considérons l'espace vectoriel E des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Notons G le sous-espace vectoriel de E dont les éléments sont les fonctions paires et H le sous-espace vectoriel de E dont les éléments sont les fonctions impaires. Il est clair que $G \cap H = \{0\}$, où 0 représente la fonction nulle.

Montrons que toute fonction f de E admet une écriture unique de la forme $f = g + h$, avec $g \in G$ et $h \in H$.



Si g et h existent, on a

$$f(x) = g(x) + h(x) \text{ et } f(-x) = g(-x) + h(-x) = g(x) - h(x),$$

donc nécessairement $g(x) = \frac{f(x) + f(-x)}{2}$ et $h(x) = \frac{f(x) - f(-x)}{2}$.

Il est facile de vérifier que les deux fonctions ainsi définies conviennent.

Dans ces deux exemples, on a une décomposition de l'espace en deux sous-espaces tels que $F + G = E$ et $F \cap G = \{0\}$.

9.2 DÉCOMPOSITION EN SOMME DIRECTE

Définition : somme directe, supplémentaire. Soit E un espace vectoriel et soient F et G deux sous-espaces vectoriels de E . On dit que E est somme directe de F et de G , et on écrit $E = F \oplus G$ (ou $E = G \oplus F$), si tout vecteur u de E admet une écriture unique de la forme $u = v + w$ avec $v \in F$ et $w \in G$. Dans cette situation, on dit que F et G sont des sous-espaces supplémentaires de E ou que F est un supplémentaire de G dans E .

Proposition 1 : critère pour les sommes directes. Avec les notations précédentes, la condition $E = F \oplus G$ équivaut aux deux conditions suivantes : $E = F + G$ et $F \cap G = \{0\}$.

Démonstration. Supposons d'abord que $E = F \oplus G$. Comme il existe, pour tout u de E des vecteurs v de F et w de G tels que $u = v + w$, on a $E = F + G$. Si $u \in F \cap G$, on peut écrire $u = u + 0 = 0 + u$; dans les deux écritures, u est la somme d'un vecteur de F et d'un vecteur de G ; la condition d'unicité impose $u = 0$.

Montrons la réciproque. Soit $u \in E$. Comme $E = F + G$, il existe $v \in F$ et $w \in G$ tels que $u = v + w$. Supposons qu'on ait aussi $u = v_1 + w_1$ avec $v_1 \in F$ et $w_1 \in G$. On a $v - v_1 = w_1 - w$; la condition $F \cap G = \{0\}$ montre que $v - v_1 = w_1 - w = 0$, d'où l'unicité. $\square$

Proposition 2 : base d'une somme directe. Soit E un espace vectoriel de dimension finie et soient F et G deux sous-espaces vectoriels de E tels que $E = F \oplus G$.

- 1) Soit B une base de F et C une base de G . Alors $B \cup C$ est une base de E .
- 2) $\dim(E) = \dim(F) + \dim(G)$.

Démonstration. On a $\text{Vect}(B \cup C) = \text{Vect}(\text{Vect}(B) \cup \text{Vect}(C)) = \text{Vect}(F \cup G) = F + G = E$, ce qui prouve que $B \cup C$ est une famille génératrice de E . Il est facile de vérifier que $B \cup C$ est une famille libre de E en utilisant $F \cap G = \{0\}$; d'où le 1). Le 2) s'en déduit immédiatement. $\square$

Proposition 3 : existence de supplémentaires. Soient E un espace vectoriel de dimension finie et F un sous-espace de E . Alors F admet un supplémentaire dans E .

Démonstration. Notons B une base de F . D'après le théorème de la base incomplète, on peut compléter B en une base $B \cup C$ de E . Il est facile de vérifier que $\text{Vect}(C)$ est un supplémentaire de F . $\square$

Si $\dim(E) = n$, le supplémentaire d'un sous-espace de dimension k est de dimension $n - k$. Les sous-espaces $\{0\}$ et E sont supplémentaires dans E .

On remarquera qu'on peut, sauf cas particulier, choisir C de multiples façons. Il existe donc de multiples possibilités de choisir un supplémentaire d'un sous-espace vectoriel de E . Par exemple, dans $\mathbb{R}^3$, les supplémentaires du plan engendré par les vecteurs e_1 et e_2 de la base canonique sont les droites vectorielles engendrées par des vecteurs dont la troisième composante n'est pas nulle.

On remarquera également qu'on ne peut simplifier des égalités entre sommes directes. Si $F \oplus G = F \oplus H$, cela n'entraîne pas, en général, $G = H$.

Le dessin du paragraphe 1 peut servir d'image pour penser une somme directe même dans un espace de dimension supérieure.

La proposition 3 subsiste en dimension infinie ; la démonstration s'appuie sur l'axiome de choix (voir exercice 19.14.2).

9.3 SOMMES DIRECTES FINIES

On peut généraliser ce qui précède au cas de n sous-espaces F_i , $1 \leq i \leq n$, d'un espace vectoriel E , ce qui nous sera particulièrement utile dans le chapitre 15.

Définition : sommes directes finies. On dira que E est somme directe (finie) des sous-espaces F_i , $1 \leq i \leq n$, si tout vecteur u de E s'écrit de manière unique $u = u_1 + \dots + u_n$ avec $u_i \in F_i$ pour $1 \leq i \leq n$. On écrira $E = \bigoplus_{1 \leq i \leq n} F_i$ ou $E = F_1 \oplus \dots \oplus F_n$.

Pour $1 \leq i \leq n$, notons E_i le sous-espace vectoriel de E somme de la famille $(F_j)_{1 \leq j \leq n, j \neq i}$. Par exemple, $E_2 = F_1 + F_3 + \dots + F_n$.

Proposition 1 : critère pour les sommes directes finies. Avec les notations précédentes, la condition $E = \bigoplus_{1 \leq i \leq n} F_i$ équivaut aux conditions suivantes : $E = F_1 + \dots + F_n$ et, pour $1 \leq i \leq n$, $E_i \cap F_i = \{0\}$.

Démonstration. Supposons d'abord que $E = \bigoplus_{1 \leq i \leq n} F_i$. Comme il existe, pour tout u de E des vecteurs u_i de F_i , $1 \leq i \leq n$, tels que $u = u_1 + \dots + u_n$, on a $E = F_1 + \dots + F_n$. Pour montrer la seconde condition, une permutation des indices permet de ramener les différents cas au cas $i = 1$. Si $u \in E_1 \cap F_1$, on a $u \in E_1$ et $u \in F_1$; $u \in E_1$ signifie qu'il existe des vecteurs u_i de F_i , $2 \leq i \leq n$, tels que $u = u_2 + \dots + u_n$, $u \in F_1$ signifie que u s'écrit $u = u + 0 + \dots + 0$. La comparaison des écritures montre que $u = 0$.

Montrons la réciproque. Soit $u \in E$. Comme $E = F_1 + \dots + F_n$, il existe des $u_i \in F_i$, $1 \leq i \leq n$ tels que $u = u_1 + \dots + u_n$. Supposons qu'il existe des $v_i \in F_i$, $1 \leq i \leq n$ tels que $u = v_1 + \dots + v_n$. En faisant la différence de ces deux égalités, on voit que $u_i - v_i$ est, pour $1 \leq i \leq n$, à la fois dans E_i et dans F_i , ce qui implique $u_i - v_i = 0$ et l'unicité. $\square$

On démontrera par récurrence les généralisations de la proposition 2 de 9.2 suivantes.

Proposition 2 : base d'une somme directe finie. Soit E un espace vectoriel de dimension finie et soient F_i , $1 \leq i \leq n$, des sous-espaces vectoriels de E tels que $E = \bigoplus_{1 \leq i \leq n} F_i$.

- 1) Soient B_i des bases des F_i , $1 \leq i \leq n$. Alors $B_1 \cup \dots \cup B_n$ est une base de E .
- 2) $\dim(E) = \dim(F_1) + \dots + \dim(F_n)$.

9.4 PRODUIT DE DEUX ESPACES VECTORIELS

Définition : produit d'espaces vectoriels. Soient E_1 et E_2 deux espaces vectoriels. L'ensemble $E_1 \times E_2$ est l'ensemble des couples (u, v) formés par un vecteur u_1 de E_1 et un vecteur u_2 de E_2 . On peut munir cet ensemble d'une structure d'espace vectoriel en définissant la somme de deux couples et le produit d'un couple par un réel par les formules :

$$(u_1, u_2) + (u'_1, u'_2) = (u_1 + u'_1, u_2 + u'_2)$$

$$\lambda(u_1, u_2) = (\lambda u_1, \lambda u_2)$$

Les vérifications sont immédiates.

Définition d'applications. Les applications $p_1 : E_1 \times E_2 \rightarrow E_1$ et $p_2 : E_1 \times E_2 \rightarrow E_2$ définies par : $p_1(u_1, u_2) = u_1$ et $p_2(u_1, u_2) = u_2$ sont clairement linéaires et surjectives ;

leurs noyaux sont $\ker(p_1) = \{(0, u_2) \mid u_2 \in E_2\}$ qui est isomorphe à E_2 et $\ker(p_2) = \{(0, u_1) \mid u_1 \in E_1\}$ qui est isomorphe à E_1 .

Les applications $j_1 : E_1 \rightarrow E_1 \times E_2$ et $j_2 : E_2 \rightarrow E_1 \times E_2$ définies par : $j_1(u_1) = (u_1, 0)$ et $j_2(u_2) = (0, u_2)$ sont clairement linéaires et injectives ; leurs images sont $j_1(E_1) = \ker(p_2)$ et $j_2(E_2) = \ker(p_1)$.

Notons enfin que $p_1 \circ j_1 = \text{id}_{E_1}$ et que $p_2 \circ j_2 = \text{id}_{E_2}$.

Proposition 1 : $E_1 \times E_2$ comme somme directe.

1) Avec les notations précédentes, $E_1 \times E_2 = j_1(E_1) \oplus j_2(E_2)$.

2) On a $\dim(E_1 \times E_2) = \dim(E_1) + \dim(E_2)$.

Démonstration. 1) Soit (u_1, u_2) un vecteur de $E_1 \times E_2$. Comme $(u_1, u_2) = (u_1, 0) + (0, u_2) = j_1(u_1) + j_2(u_2)$, on a $E_1 \times E_2 = j_1(E_1) + j_2(E_2)$. Pour conclure, il suffit de vérifier que $j_1(E_1) \cap j_2(E_2) = \{0\}$, ce qui est évident.

2) Cela résulte de la seconde proposition de 9.2. □

Nous venons de voir que l'espace vectoriel $E_1 \times E_2$ peut être considéré de deux façons différentes : comme somme directe et comme produit. Ces deux points de vue conduisent à des propriétés différentes.

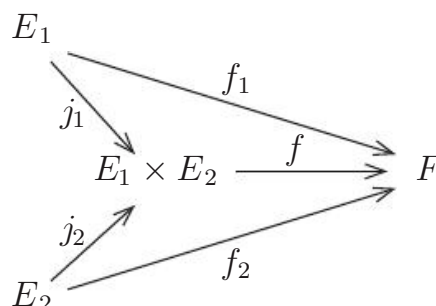
Proposition 2 : propriété universelle de $E_1 \times E_2$ comme somme directe. Soient E_1 et E_2 deux espaces vectoriels et soient $j_1 : E_1 \rightarrow E_1 \times E_2$ et $j_2 : E_2 \rightarrow E_1 \times E_2$ définies comme ci-dessus. Soit F un espace vectoriel et soient $f_1 : E_1 \rightarrow F$ et $f_2 : E_2 \rightarrow F$ deux applications linéaires. Il existe une application linéaire unique $f : E_1 \times E_2 \rightarrow F$ telle que $f \circ j_1 = f_1$ et $f \circ j_2 = f_2$.

Démonstration. Si f existe, on a nécessairement, pour tout $u_1 \in E_1$ et tout $u_2 \in E_2$ $f(u_1, u_2) = f[j_1(u_1) + j_2(u_2)] = f(j_1(u_1)) + f(j_2(u_2)) = f_1(u_1) + f_2(u_2)$.

Cette condition définit entièrement f . Il reste à vérifier que le f ainsi défini est bien linéaire et que $f \circ j_1 = f_1$ et $f \circ j_2 = f_2$. Cela se fait sans problème. □

Cette proposition permet de construire des applications linéaires de source $E_1 \times E_2$ à partir d'applications linéaires de sources E_1 et E_2 et de même but.

Le diagramme commutatif suivant permet de mieux mémoriser la propriété universelle de la somme directe :

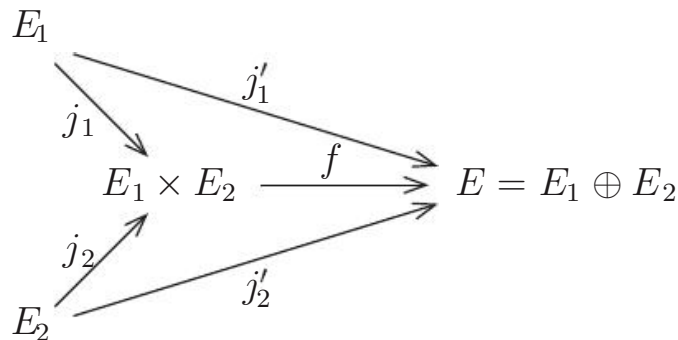


Dans ce contexte, l'espace $E_1 \times E_2$ muni des applications j_1 et j_2 est appelé somme directe des espaces E_1 et E_2 .

Inversement, un espace qui admet une décomposition en somme directe est isomorphe à un produit d'après la proposition suivante.

Proposition 3 : somme directe et produit. Soient E un espace vectoriel et E_1 et E_2 deux sous-espaces vectoriels supplémentaires de E . Alors $E_1 \times E_2$ est isomorphe à E .

Démonstration. Notons $j'_1 : E_1 \rightarrow E$ et $j'_2 : E_2 \rightarrow E$ les applications linéaires définies par les inclusions de E_1 et E_2 dans E . La propriété universelle précédente permet de construire une application linéaire $f : E_1 \times E_2 \rightarrow E$ telle que le diagramme suivant commute.



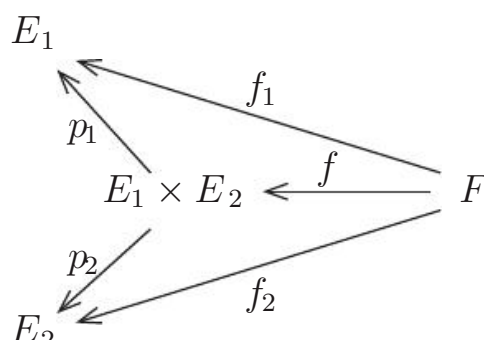
On a $f(u_1, u_2) = u_1 + u_2$ pour tout $u_1 \in E_1$ et tout $u_2 \in E_2$. Les propriétés de la somme directe permettent de montrer facilement que f est injective et surjective. L'inverse de f est défini de la manière suivante : tout u de E s'écrit de manière unique $u = u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$ et f^{-1} est l'application $u \mapsto (u_1, u_2)$. $\square$

Proposition 4 : propriété universelle de $E_1 \times E_2$ comme produit. Soient E_1 et E_2 deux espaces vectoriels et soient $p_1 : E_1 \times E_2 \rightarrow E_1$ et $p_2 : E_1 \times E_2 \rightarrow E_2$ définis comme ci-dessus. Soient F un espace vectoriel et $f_1 : F \rightarrow E_1$ et $f_2 : F \rightarrow E_2$ deux applications linéaires. Il existe une application linéaire unique $f : F \rightarrow E_1 \times E_2$ telle que $p_1 \circ f = f_1$ et $p_2 \circ f = f_2$.

Démonstration. Si f existe, on a nécessairement, pour tout $u \in F$: $p_1(f(u)) = f_1(u)$ et $p_2(f(u)) = f_2(u)$ donc $f(u) = (f_1(u), f_2(u))$. Cette condition définit entièrement f . Il reste à vérifier que le f ainsi défini est bien linéaire et que $p_1 \circ f = f_1$ et $p_2 \circ f = f_2$. Cela se fait sans problème. $\square$

Cette proposition permet de construire des applications linéaires de but $E_1 \times E_2$ à partir d'applications linéaires de but E_1 et E_2 et de même source. Dans ce contexte, l'espace $E_1 \times E_2$ muni des applications p_1 et p_2 est appelé produit des espaces E_1 et E_2 . Le diagramme suivant, qui s'obtient juste en renversant le sens des flèches des

diagrammes précédents, permet de mieux mémoriser la propriété universelle du produit :



Généralisation. On peut généraliser ces constructions de somme directe et de produit au cas de n espaces $E_1, E_2, \dots, E_n$ et même au cas d'une famille infinie d'indices. Si $(E_i)_{i \in I}$ est une famille indexée par I d'espaces vectoriels, leur produit est l'espace des suites $(u_i)_{i \in I}$ avec $u_i \in E_i$ pour tout $i \in I$. Leur somme directe est l'espace des suites $(u_i)_{i \in I}$ avec $u_i \in E_i$ pour tout $i \in I$ tous nuls, sauf un nombre fini d'entre eux. Si I est infini, la somme directe est un sous-espace strict du produit. Nous ne détaillons pas les propriétés universelles de ces deux constructions, inutiles pour la suite de ce livre, les laissant aux lectrices et lecteurs.

9.5 PROJECTEURS

Définition : projecteur. Soit E un espace vectoriel. On appelle projecteur (ou projection) de E tout endomorphisme p de E tel que $p \circ p = p$.

Exemples.

1) Avec les notations de 9.4, les composés $j_1 \circ p_1, j_2 \circ p_2 : E_1 \times E_2 \rightarrow E_1 \times E_2$ sont des projecteurs de $E_1 \times E_2$ définis par $(u_1, u_2) \mapsto (u_1, 0)$ et $(u_1, u_2) \mapsto (0, u_2)$ respectivement.

2) L'exemple précédent et la construction de l'inverse de f dans la proposition 3 de 9.4 permettent de voir que si E admet une décomposition en somme directe $E_1 \oplus E_2$, tout u de E s'écrit de manière unique $u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$ et les applications $u \mapsto u_1$ et $u \mapsto u_2$ sont des projecteurs de E .

Ce second exemple admet une réciproque importante.

Proposition : projecteur et décomposition en somme directe. Soit E un espace vectoriel et soit p un projecteur de E . On a $E = E_1 \oplus E_2$ avec $E_1 = \text{Im}(p)$ et $E_2 = \text{ker}(p)$.

Démonstration. On a $E_1 \cap E_2 = \{0\}$. Soit, en effet, $u \in E_1 \cap E_2$; comme $u \in E_2$, il existe $v \in E$ tel que $p(v) = u$, d'où $u = p(v) = p(p(v)) = p(u) = 0$.

On a $E_1 + E_2 = E$ car si $u \in E$, on a $u = p(u) + (u - p(u))$ avec $p(u) \in E_1$ et $u - p(u) \in E_2$ puisque $p(u - p(u)) = p(u) - p(p(u)) = 0$. $\square$

Le second projecteur. Si $p : E \rightarrow E$ est un projecteur, $q = \text{id}_E - p$ est également un projecteur. On a $p + q = \text{id}$, $\ker(q) = \text{Im}(p)$ et $\text{Im}(q) = \ker(p)$.

Généralisation. Si E est somme directe des F_i , $1 \leq i \leq n$, les deux paragraphes précédents montrent qu'il existe des projecteurs p_i , $1 \leq i \leq n$, tels que $p_i(u_1 + \dots + u_n) = u_i$. On a $p_1 + \dots + p_n = \text{id}$.

C'est Erhard Schmidt (1876-1959) qui, dans des articles dont nous reparlerons au chapitre 16, a montré l'importance de la notion de projecteur dans des problèmes d'analyse fonctionnelle.

9.6 ESPACES VECTORIELS QUOTIENTS

Idée de quotient. L'idée générale de quotient sera approfondie en 18.1. Elle peut être présentée assez vaguement de la façon suivante. On connaît un ensemble E muni d'une certaine structure (espace vectoriel, groupe, anneau, etc.). On cherche à identifier certains éléments de E qui possèdent la même propriété donnée par une relation d'équivalence. On construit ainsi un ensemble Q à partir de classes d'équivalence d'éléments de E , appelé quotient et sur lequel on cherche à définir la même structure que sur E . Le lien entre E et Q est assuré par l'application $E \rightarrow Q$ qui associe à chaque élément de E sa classe d'équivalence. Le premier à avoir calculé sur des classes d'éléments est Gauss. L'idée de quotient est présente dans les travaux de Jordan dès 1873.

Quotient d'un espace vectoriel par un sous-espace. Nous allons rester en dimension finie. Soient E un espace vectoriel et F un sous-espace de E . Soit G un supplémentaire de F dans E . On a vu qu'il existe un isomorphisme $\varphi : E \rightarrow F \times G$, qui associe à un vecteur $u \in E$ le couple $(v, w) \in F \times G$ tel que $u = v + w$ et une application linéaire $q : F \times G \rightarrow G$ définie par $q(v, w) = w$. Posons $\pi = q \circ \varphi : E \rightarrow G$.

L'application linéaire π est surjective. Soit $u = v + w$ et $u' = v' + w'$ deux éléments de E , avec $v, v' \in F$ et $w, w' \in G$. Si $\pi(u) = \pi(u')$, c'est que $w = w'$ donc $u - u' = v - v' \in F$. On peut dire que u et u' sont égaux modulo F et que l'application π tue les différences modulo F .

Proposition : propriété universelle du quotient. Avec les notations précédentes, pour toute application linéaire $f : E \rightarrow H$ telle que $f(F) = \{0\}$, il existe une application

linéaire unique $g : G \rightarrow H$ telle que $g \circ \pi = f$, autrement dit le diagramme suivant est commutatif :

$$\begin{array}{ccccc}
 F & \xrightarrow{\text{incl}} & E & \xrightarrow{\pi} & G \\
 & \searrow 0 & \downarrow f & \swarrow g & \\
 & & H & &
 \end{array}$$

Démonstration. Soit $B_1 = (e_1, \dots, e_k)$ une base de F et $B_2 = (e_{k+1}, \dots, e_n)$ une base de G . On sait que $B = B_1 \cup B_2$ est une base de E . On construit l'application linéaire g en posant $g(e_i) = f(e_i)$ pour $k+1 \leq i \leq n$. Comme f et $g \circ \pi$ sont égales sur tous les vecteurs de la base B , elles sont égales. L'unicité de g résulte de la surjectivité de π .

Définition. L'espace G et l'application linéaire $\pi : E \rightarrow G$ forment ce qu'on appelle le quotient de E par F . On note $G = E/F$.

Un peu d'histoire. Les propriétés universelles sont bien commodes pour construire des applications linéaires de source une somme directe ou un quotient, ou celles de but un produit ou un noyau. Les travaux de Samuel Eilenberg (1913-1998) et de Saunders Mac Lane sont à l'origine de ces propriétés ; ils ont été publiés dans les années 1940.

Saunders Mac Lane est né en 1909 dans le Connecticut. Après ses études à Chicago, il part en 1931 préparer sa thèse à Göttingen en Allemagne. L'Institut mathématique de cette ville était, depuis près de 100 ans, l'un des centres mathématiques les plus prestigieux du monde : Gauss, Bernhard Riemann (1826-1866), Gustav Lejeune-Dirichlet (1805-1859), Félix Klein (1849-1925) ou encore Hermann Minkowski (1864-1909) y avaient travaillé. Le grand David Hilbert (1862-1943) y était encore, ainsi qu'Emmy Noether (1882-1935), Hermann Weyl (1885-1955), etc. Mais Mac Lane vit les derniers moments heureux de l'Institut. En 1933, les nazis prennent le pouvoir en Allemagne et commencent à chasser tous ceux qui sont juifs, ont un ascendant juif, etc., de leurs postes. Emmy Noether et Richard Courant (1888-1972) sont démis de leurs fonctions le 25 avril 1933, d'autres suivent le 27. Hermann Weyl, dont la femme est juive, part aux États-Unis à la fin de l'été. Très peu de mathématiciens resteront à Göttingen pendant les douze années qui suivent. Saunders Mac Lane termine sa thèse le plus rapidement possible et rentre aux États-Unis.

En 1940, Mac Lane rencontre Samuel Eilenberg, d'origine polonaise et qui vient de fuir son pays. Leur collaboration sera très fructueuse : création de la théorie des

Chapitre 9 • Sommes directes, produits, quotients

catégories, développement de la topologie algébrique, etc. En 1941, Mac Lane publie avec Garrett Birkhoff (1911-1996) l'un des tout premiers traités d'algèbre moderne : *A survey of modern algebra* qui reprend les idées du livre *Moderne Algebra* de Van der Waerden et dont l'influence sera profonde.

Saunders Mac Lane est mort le 14 avril 2005.



Saunders Mac Lane (1909-2005)

Reproduit avec l'aimable autorisation de l'Université de Chicago

➤ Vers le chapitre 10

Le chapitre 9 termine ce qu'il semble raisonnable d'enseigner en première année de Licence au sujet de l'algèbre linéaire actuellement. Le chapitre 10 est un peu à part. Il présente des notions sur l'espace vectoriel formé par les hyperplans d'un espace vectoriel, appelé espace dual. Certaines et certains pourront le trouver un peu difficile, un peu abstrait ; si sa lecture peut-être différée jusqu'au chapitre 16, il semble plus à sa place ici.

Exercices

9.1 Sommes directes

a) Soient m et n des entiers. Pour quelle valeur de p a-t-on $\mathbb{R}^p$ isomorphe à :

➤ $\mathbb{R}^m \times \mathbb{R}^n$?

➤ $\mathbb{R}^m \oplus \mathbb{R}^n$?

b) Soient E un espace vectoriel de dimension n et $B = (e_1, \dots, e_n)$ une base de E . On pose $E_i = \text{Vect}(e_i)$ pour $1 \leq i \leq n$. Montrer que : $E = \bigoplus_{1 \leq i \leq n} E_i$.

c) Soit E un espace vectoriel de dimension finie, B une base de E et (B_1, B_2) une partition de B . Montrer que $E = \text{Vect}(B_1) \oplus \text{Vect}(B_2)$.

d) Soit E un espace vectoriel de dimension finie et soient F et G deux sous-espaces de E tels que $F \cap G = \{0\}$ et $\dim(F) + \dim(G) = \dim(E)$. Montrer que $E = F \oplus G$.

e) Soient $E = F_1 \oplus F_2$ et G un sous-espace de E . On pose $H_i = G \cap F_i$ pour $1 \leq i \leq 2$. A-t-on $G = H_1 \oplus H_2$?

f) Soient E un espace vectoriel et E_1, E_2 et E_3 trois sous-espaces vectoriels de E . Montrer que $E = E_1 \oplus E_2 \oplus E_3$ si et seulement si $E = E_1 + E_2 + E_3$, $E_1 \cap E_2 = \{0\}$ et $(E_1 + E_2) \cap E_3 = \{0\}$.

g) Soient E un espace vectoriel de dimension n et F et G deux sous-espaces de E de même dimension k . Montrer qu'il existe un sous-espace de E qui est en même temps supplémentaire de F et de G .

9.2 Matrices symétriques et antisymétriques

La transposée d'une matrice $M = (a_{ij})$ est la matrice ${}^tM = (a_{ji})$ (voir 10.5 pour une présentation de ces matrices). Soit n un entier. On note E l'espace vectoriel $M_n(\mathbb{R})$ des matrices carrées d'ordre n à coefficients réels, F l'ensemble des matrices symétriques, c'est-à-dire des matrices M de E telles que $M = {}^tM$, G l'ensemble des matrices antisymétriques, c'est-à-dire des matrices M de E telles que $M = -{}^tM$. Montrer que F et G sont des sous-espaces vectoriels de E et que $E = F \oplus G$.

9.3 Projecteurs

a) Le second projecteur

Soit E un espace vectoriel et soit p un projecteur de E . On pose $E_1 = \text{Im}(p)$, $E_2 = \text{ker}(p)$ et $q = \text{id} - p$. Montrer que q est un projecteur. Exprimer $\text{ker}(q)$ et $\text{Im}(q)$ en fonction de $\text{ker}(p)$ et $\text{Im}(p)$.

b) Définir les projecteurs associés à la décomposition en somme directe de l'exemple 2 du paragraphe 1.

c) Soit E un espace vectoriel et soient p et q deux projecteurs de E . À quelles conditions sur les images et noyaux de p et q équivalent les relations $p \circ q = p$ et $p \circ q = q$?

9.4 Projecteurs qui commutent

- 1) Donner un exemple de deux projecteurs p et q de $\mathbb{R}^3$ tels que $p \circ q = q \circ p$.
- 2) Donner un exemple de deux projecteurs p et q de $\mathbb{R}^2$ tels que $p \circ q \neq q \circ p$.
- 3) Soit E un espace vectoriel et soient p, q deux projecteurs de E . On pose $E_1 = \ker(p) \cap \ker(q)$, $E_2 = \text{Im}(p) \cap \ker(q)$, $E_3 = \ker(p) \cap \text{Im}(q)$, $E_4 = \text{Im}(p) \cap \text{Im}(q)$.
 - a) On suppose que $E = E_1 \oplus E_2 \oplus E_3 \oplus E_4$. Montrer que $p \circ q = q \circ p$.
 - b) Montrer la réciproque.
 - c) Montrer qu'alors $p \circ q$ et $p + q - p \circ q$ sont des projecteurs. Déterminer l'image et le noyau de ces projecteurs.

9.5 Soient E un espace vectoriel, p un projecteur de E , a un réel et w un vecteur de E . Résoudre et discuter suivant les valeurs de a et w l'équation :

$$u + ap(u) = w \quad (E)$$

9.6 Construction d'applications linéaires

Soient E_1, E_2, E'_1 et E'_2 des espaces vectoriels et soient $f_1 : E_1 \rightarrow E'_1$ et $f_2 : E_2 \rightarrow E'_2$ des applications linéaires. Construire avec ces deux applications des applications linéaires :

- a) $g : E_1 \oplus E_2 \rightarrow E'_1 \oplus E'_2$;
- b) $h : E_1 \times E_2 \rightarrow E'_1 \times E'_2$;

en précisant la valeur de ces applications en un vecteur de l'espace de départ.

Solutions

- 9.1 a)** Les résultats sur la dimension d'une somme directe et d'un produit donnent $p = m + n$.
- b, c)** Cela résulte des propriétés d'une base dans les deux cas.
- d)** $F + G$ est somme directe de F et de G . On a $\dim(F + G) = \dim(F) + \dim(G) = \dim(E)$. Donc $F + G = E$ et la conclusion.
- e)** Prendre $E = \mathbb{R}^2$, $G = \text{Vect}((1, 1))$, $F_1 = \text{Vect}((1, 0))$, $F_2 = \text{Vect}((0, 1))$ pour obtenir un contre-exemple.
- f)** Soit $u = u_1 + u_2 + u_3 = u'_1 + u'_2 + u'_3$ deux écritures d'un vecteur de E en fonction de vecteurs des trois sous-espaces : $u_1, u'_1 \in E_1$, etc. Comme $(E_1 + E_2) \cap E_3 = \{0\}$, $u_1 - u'_1 + u_2 - u'_2 = u'_3 - u_3$ implique que les deux membres sont nuls, d'où $u'_3 = u_3$ et $u_1 + u_2 = u'_1 + u'_2$; la condition $E_1 \cap E_2 = \{0\}$ donne alors $u_1 = u'_1$, $u_2 = u'_2$. Le reste est facile.

g) Raisonnons par récurrence sur $s = n - k$. Si $s = 0$, il n'y a rien à démontrer. Si $s > 0$ et si F et G sont différents, alors $F \cup G$ n'est pas un sous-espace de E . Soit u_1 un vecteur de E tel que $u_1 \notin F \cup G$. On peut appliquer l'hypothèse de récurrence à $F \oplus \mathbb{R}u_1$ et $G \oplus \mathbb{R}u_1$ pour trouver un sous-espace H supplémentaire de ces deux espaces. Alors $H \oplus \mathbb{R}u_1$ est supplémentaire de F et de G .

9.2 Comme $M = {}^tM = -{}^tM$ implique $M = 0$, la seule matrice symétrique et antisymétrique est la matrice nulle. Si la décomposition en somme directe est possible, toute matrice M de E s'écrit $M = S + A$ où S est une matrice symétrique et A une matrice antisymétrique. On a donc ${}^tM = {}^t(S + A) = S - A$, donc $s = \frac{1}{2}(M + {}^tM)$ et $A = \frac{1}{2}(M - {}^tM)$. Il est facile de vérifier que ces deux matrices conviennent. Par exemple :

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b' & c' \\ b' & e & f' \\ c' & f' & i \end{pmatrix} + \begin{pmatrix} 0 & b'' & c'' \\ -b'' & 0 & f'' \\ -c'' & -f'' & 0 \end{pmatrix}$$

avec $b' = \frac{b+d}{2}$, $b'' = \frac{b-d}{2}$, etc.

On remarque que les matrices antisymétriques ont des coefficients nuls sur la diagonale.

9.3 a) $\ker(q) = \text{Im}(p)$; $\text{Im}(q) = \ker(p)$.

b) À la fonction f , les projecteurs associent les fonctions $x \mapsto \frac{f(x) + f(-x)}{2}$ et $x \mapsto \frac{f(x) - f(-x)}{2}$.

c) $p \circ q = p$ équivaut à $\ker(q) \subset \ker(p)$; pour le voir, écrire $u = u_1 + u_2$ avec $u_1 \in \ker(q)$ et $u_2 \in \text{Im}(q)$;

$p \circ q = q$ équivaut à $\text{Im}(q) \subset \text{Im}(p)$ (même indication).

9.4 1) Prendre $p : (x, y, z) \mapsto (x, y, 0)$ et $q : (x, y, z) \mapsto (x, 0, z)$. On a $p \circ q = q \circ p : (x, y, z) \mapsto (x, 0, 0)$.

2) En général, tout projecteur p commute avec $\text{id} - p$. Prendre deux projections orthogonales de $\mathbb{R}^2$ sur des droites vectorielles distinctes et non orthogonales.

3) a) Simple calcul.

b) On a $E = \ker(p) \oplus \text{Im}(p)$. On va décomposer $\ker(p)$ et $\text{Im}(p)$ en somme directe.

Soit $u \in \ker(p)$; comme $E = \ker(q) \oplus \text{Im}(q)$, u s'écrit $u_1 + u_2$ avec $u_1 \in \ker(q)$ et $u_2 \in \text{Im}(q)$. On a $p(u_2) = p(q(u)) = q(p(u)) = 0$ et $p(u_1) = p(u - u_2) = 0$ donc $u_1, u_2 \in \ker(p)$ et $\ker(p) = E_1 \oplus E_3$.

Soit maintenant $u = p(v) = u_1 + u_2 \in \text{Im}(p)$ avec $u_1 \in \ker(q)$, $u_2 \in \text{Im}(q)$ et $v \in E$.
On a $u_2 = q(u) = q(p(v)) = p(q(v))$ et $u_1 = u - u_2 = p(v - q(v))$, donc $u_1, u_2 \in \text{Im}(p)$ et $\text{Im}(p) = E_2 \oplus E_4$.

c) On vérifie que $p \circ q$ et $p + q - p \circ q$ sont des projecteurs par des calculs simples.

$$\text{Im}(p \circ q) = E_4, \ker(p \circ q) = E_1 \oplus E_2 \oplus E_3.$$

$$\ker(p + q - p \circ q) = E_1, \text{Im}(p + q - p \circ q) = E_2 \oplus E_3 \oplus E_4.$$

On peut remarquer que $p' = \text{id} - p$, $q' = \text{id} - q$ sont des projecteurs qui commutent et que $p + q - p \circ q = \text{id} - p' \circ q'$.

9.5 Posons $f = \text{id} + ap$. Le problème revient à résoudre l'équation $f(u) = w$. Comme f est linéaire, on peut appliquer le résultat général sur la résolution des équations linéaires. On cherche d'abord $\ker(f) = \{u \mid f(u) = 0\}$. Si $f(u) = 0$, on a $p(u + ap(u)) = 0$, donc $(1 + a)p(u) = 0$.

Si $a \neq -1$, cette condition impose $p(u) = 0$ et donc $u = 0$, d'où $\ker(f) = \{0\}$. L'équation (E) a donc au plus une solution. En appliquant p à (E), on trouve $p(u) = \frac{1}{1+a}p(w)$; en reportant cette valeur dans (E), on trouve $S(E) = \{w - \frac{a}{1+a}p(w)\}$.

Si $a = -1$, on a $f = \text{id} - p$; f est un projecteur de noyau $\text{Im}(p)$ et d'image $\ker(p)$.

Si $w \notin \ker(p)$, $S(E) = \emptyset$. Si $w \in \ker(p)$, $w = f(w')$, alors $S(E) = w' + \text{Im}(p)$.

9.6 Utiliser les propriétés universelles de la somme et du produit.

10.1 INTRODUCTION

La dualité est une notion d'algèbre linéaire dont la mise au point a été longue. Elle est déjà dans l'esprit des mathématiciens du début du XIX^e siècle qui voient dans la formule donnant l'équation d'un plan : $ax + by + cz = 0$ une *dualité* entre les coefficients a, b, c et les variables x, y, z . Cette propriété permet de transformer mécaniquement des résultats de géométrie en écrivant plan à la place de point, point à la place de plan, etc.

Dans ce chapitre, nous allons identifier les hyperplans d'un espace vectoriel E aux éléments d'un autre espace vectoriel, celui des formes linéaires sur E . Ceci va nous permettre de calculer avec eux comme on ne pouvait le faire avant.

Comme dans le reste de ce livre, nous allons nous limiter à la dimension finie. Les résultats profonds sur la dualité dans des espaces fonctionnels trouvent leur origine dans les travaux de Banach des années 1920 qui ne peuvent être étudiés ici.

Rappelons (voir 7.2) qu'on note $L(E, F)$ l'ensemble des applications linéaires de E dans F . On a défini en 7.2 deux lois sur cet ensemble :

- la somme de deux applications linéaires $f, g : E \rightarrow F$ est l'application linéaire $f + g : E \rightarrow F$ définie par $(f + g)(u) = f(u) + g(u)$;
- le produit d'une application linéaire $f : E \rightarrow F$ par un scalaire a est l'application linéaire $af : E \rightarrow F$ définie par $(af)(u) = af(u)$.

L'ensemble $L(E, F)$ muni des lois d'addition et de multiplication par un scalaire précédentes est un espace vectoriel.

Définition. Soit K un corps. On appelle dual d'un K -espace vectoriel E l'espace vectoriel $L(E, K)$. On le note E^* . Les éléments de E^* sont des applications linéaires de E dans K ; on les appelle formes linéaires sur E .

Pour ne pas rendre ce chapitre trop abstrait, nous prendrons $K = \mathbb{R}$ dans tout ce chapitre, mais tout ce qui va être dit vaut pour un corps quelconque.

On appelle bidual d'un espace vectoriel E le dual de E^* . On le note E^{**} .

10.2 FORMES LINÉAIRES ET HYPERPLANS

Commençons par décrire le dual d'un espace vectoriel E de dimension n . Un élément de E^* est une forme linéaire $f: E \rightarrow \mathbb{R}$. Son noyau $f^{-1}(0) = \ker(f)$ est un sous-espace vectoriel de E . Si $f = 0$, $\text{Im}(f) = \{0\}$ et $\ker(f) = E$. Hormis ce cas particulier, on a $\text{Im}(f) = \mathbb{R}$, et le théorème du rang montre que $\dim(\ker(f)) = n - 1$; $\ker(f)$ est donc un hyperplan de E . Il est clair que les formes linéaires af proportionnelles à f ont, si $a \neq 0$, le même noyau que f ; elles définissent donc le même hyperplan.

Inversement, considérons un hyperplan H de E . Montrons que H est le noyau d'une forme linéaire f sur E . Choisissons une base $(e_1, \dots, e_{n-1})$ de H et complétons-la en une base $B = (e_1, \dots, e_n)$ de E . On a $e_n \notin H$. Pour définir f , on pose $f(e_i) = 0$ pour $1 \leq i \leq n - 1$ et $f(e_n) = 1$. On a $H \subset \ker(f) \subset E$ et $\ker(f) \neq E$ donc $\ker(f) = H$. Soit g une forme linéaire de noyau H . On a $g(e_i) = 0$ pour $1 \leq i \leq n - 1$ et $g(e_n) = a \neq 0$. On en déduit que $g = af$. Deux formes non proportionnelles définissent donc des hyperplans différents. On peut énoncer la proposition suivante.

Proposition. Soient E un espace vectoriel de dimension finie et E^* son dual.

- À toute forme linéaire non nulle f de E^* correspond un hyperplan de E : le noyau de f .
- Tout hyperplan de E est noyau d'une forme linéaire de E^* .
- Deux formes linéaires non nulles et proportionnelles définissent le même hyperplan.
- Deux formes non proportionnelles définissent des hyperplans distincts. □

Matrice d'une forme linéaire. Choisissons une base $B = (e_1, \dots, e_n)$ de E et munissons $\mathbb{R}$ de la base canonique (1) . La matrice $M(f, B, \text{can})$ d'une forme linéaire f par rapport à ces bases est une matrice ligne de la forme

$$(a_1 \dots a_n)$$

où $a_i = f(e_i)$ pour $1 \leq i \leq n$.

Équation d'un hyperplan. Détaillons comment les formes linéaires permettent de manier les hyperplans. Soient E un espace vectoriel de dimension finie n et H un hyperplan de E . Si f est une forme linéaire sur E de noyau H , l'équation $f(u) = 0$ est appelée équation de H . Comme la forme linéaire f est définie à un facteur multiplicatif non nul près, certains préfèrent souligner qu'il s'agit là d'une équation de H .

Si on choisit comme ci-dessus une base $B = (e_1, \dots, e_n)$ de E telle que $(e_1, \dots, e_{n-1})$ soit une base de H , le vecteur $u = x_1 e_1 + \dots + x_n e_n$ appartient à H si et seulement si $x_n = 0$. Si la base B est une base quelconque de E et si la matrice de f est $(a_1 \dots a_n)$, l'équation $f(u) = 0$ s'écrit :

$$a_1 x_1 + \dots + a_n x_n = 0$$

Une situation fréquente est celle où l'hyperplan H est défini par $H = \text{Vect}(u_1, \dots, u_{n-1})$, où $(u_1, \dots, u_{n-1})$ sont des vecteurs linéairement indépendants définis par leurs coordonnées dans une base $B = (e_1, \dots, e_n)$ de E :

$$u_i = a_{i1} e_1 + \dots + a_{in} e_n$$

pour $1 \leq i \leq n$. On cherche à déterminer une forme linéaire non nulle f de noyau H , ce qui revient à déterminer les coefficients de la matrice $M(f, B, \text{can}) = (\alpha_1, \dots, \alpha_n)$.

On peut résoudre le système linéaire de $n - 1$ équations $f(u_i) = 0$ en les n inconnues $\alpha_1, \dots, \alpha_n$. Comme les vecteurs $(u_1, \dots, u_{n-1})$ sont linéairement indépendants, le système est de rang $n - 1$ et l'espace des solutions est de dimension 1. Les solutions non nulles sont définies à une constante multiplicative non nulle près.

On peut aussi écrire qu'un vecteur $u = (x_1, \dots, x_n)$ appartient à H si et seulement si l'algorithme du pivot appliqué aux vecteurs $(u_1, \dots, u_{n-1}, u)$ donne une dernière colonne nulle.

Déterminons, par exemple, l'équation du plan vectoriel engendré par $u_1 = (2, 1, 1)$ et $u_2 = (5, 3, 2)$.

La première méthode conduit au système :

$$\begin{cases} 2\alpha_1 + \alpha_2 + \alpha_3 = 0 \\ 5\alpha_1 + 3\alpha_2 + 2\alpha_3 = 0 \end{cases}$$

qui donne $\alpha_1 = 1$, $\alpha_2 = -1$, $\alpha_3 = -1$ à une constante non nulle près, d'où l'équation $x - y - z = 0$.

La méthode du pivot donne les tableaux :

$$\begin{array}{ccc|ccc} 2 & 5 & x & 2 & 0 & 0 & 2 & 0 & 0 \\ 1 & 3 & y & 1 & 1/2 & y-x/2 & 1 & 1/2 & 0 \\ 1 & 2 & z & 1 & -1/2 & z-x/2 & 1 & -1/2 & z+y-x \end{array}$$

Le système est de rang 2 si et seulement si la troisième colonne du dernier tableau est nulle, ce qui redonne l'équation de H (à un coefficient près *a priori*).

10.3 BASE DUALE

Soit E un espace vectoriel de dimension finie n et soit $B = (e_1, \dots, e_n)$ une base de E . On sait que pour définir un élément de E^* , c'est-à-dire une forme linéaire sur E , il suffit de se donner les images des vecteurs de la base B . On définit les formes linéaires $e_1^* : E \rightarrow \mathbb{R}$ par $e_1^*(e_1) = 1$ et $e_1^*(e_i) = 0$ pour $i \geq 2$, $e_2^* : E \rightarrow \mathbb{R}$ par $e_2^*(e_2) = 1$ et $e_2^*(e_i) = 0$ pour $i \neq 2$, etc.

Calcul de $f(u)$. Soit $u = x_1 e_1 + \dots + x_n e_n$ un vecteur de E et soit $f = a_1 e_1^* + \dots + a_n e_n^*$ une forme linéaire de E^* . On a :

$$f(u) = a_1 x_1 + \dots + a_n x_n.$$

Proposition : base duale. La famille $B^* = (e_1^*, \dots, e_n^*)$ est une base de E^* appelée base duale de la base B et $\dim(E^*) = \dim(E) = n$.

Démonstration. Montrons d'abord que B^* est une famille libre de E^* . Supposons que la forme linéaire nulle s'écrive $0 = \lambda_1 e_1^* + \dots + \lambda_n e_n^*$; on montre que tous les coefficients sont nécessairement nuls en appliquant la forme linéaire définie par le second membre en e_i : on obtient $0 = (\lambda_1 e_1^* + \dots + \lambda_n e_n^*)(e_i) = \lambda_i$ pour $1 \leq i \leq n$.

Montrons maintenant que B^* est une famille génératrice de E^* . Soit f une forme linéaire de E^* . Posons $\lambda_i = f(e_i)$ pour $1 \leq i \leq n$ et $g = \lambda_1 e_1^* + \dots + \lambda_n e_n^*$. Les formes linéaires f et g sont égales sur les vecteurs de la base B car $g(e_i) = \lambda_i = f(e_i)$ pour $1 \leq i \leq n$. Par conséquent $f = g$, ce qui permet de conclure. $\square$

Commentaire. Les bases B et B^* permettent de construire un isomorphisme $\varphi : E \rightarrow E^*$ en posant $\varphi(e_i) = e_i^*$ pour $1 \leq i \leq n$. Cet isomorphisme dépend du choix de la base B . On dit qu'il n'est pas canonique. Cet isomorphisme ne subsiste pas en dimension infinie.

Attention : la forme linéaire e_i^* n'est pas définie uniquement par le vecteur e_i . Elle est définie par toute la base à laquelle appartient e_i . La notation e_i^* ne mentionne pas la base à laquelle appartient e_i , ce qui ne crée pas d'ambiguïté en général.

Rang dans le dual. Considérons le système linéaire L à m équations et n inconnues :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

Posons $E = \mathbb{R}^n$, notons B la base canonique de E et B^* la base duale. Les lignes du système L s'écrivent $f_i(u) = 0$ avec $f_i = a_{i1}e_1^* + \dots + a_{in}e_n^*$ avec $1 \leq i \leq m$. Les coordonnées de f_i dans la base B^* sont donc $a_{i1}, \dots, a_{in}$.

Quand on résout le système L par la méthode du pivot (voir chapitre 4), les opérations sur les lignes sont exactement les opérations de la méthode du pivot du chapitre 3 pour déterminer le rang du système de vecteurs $(f_1, \dots, f_m)$ où les f_i sont décrits par leurs coordonnées par rapport à la base B^* . Le rang du système L est donc égal au rang du système $(f_1, \dots, f_m)$ dans E^* .

On voit donc que calculer avec les lignes d'un système linéaire revient à calculer avec les formes linéaires définies par ces lignes, c'est-à-dire avec des vecteurs de l'espace vectoriel E^* donnés par leurs coordonnées dans une base.

10.4 ORTHOGONAL D'UN SOUS-ESPACE

Définition : orthogonalité de formes linéaires et de vecteurs. Soit E un espace vectoriel et E^* son dual. On dit qu'une forme linéaire f de E^* et un vecteur u de E sont orthogonaux si $f(u) = 0$.

La formule donnant $f(u)$ montre que les coordonnées de u et f définissent des vecteurs de $\mathbb{R}^n$ orthogonaux pour le produit scalaire usuel (voir chapitre 16), ce qui explique la terminologie.

Orthogonal d'un sous-espace de E . Soit F un sous-espace vectoriel de E . On note $F^\perp$ le sous-ensemble de E^* des formes linéaires orthogonales à tous les éléments de F :

$$F^\perp = \{ f \in E^* \mid \forall u \in F : f(u) = 0 \}.$$

Un élément de $F^\perp$ est donc une forme linéaire qui s'annule pour tout élément de F . Autrement dit, son noyau contient F . Comme les formes linéaires non nulles correspondent aux hyperplans de E , les éléments non nuls de $F^\perp$ correspondent aux hyperplans de E contenant F .

Exemple. Considérons l'espace $\mathbb{R}^3$ muni de sa base canonique $B = (e_1, e_2, e_3)$ et posons $F = \text{Vect}(e_1)$. Les éléments de $F^\perp$ sont les formes linéaires $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ telles que $f(F) = \{0\}$. Notons $B^* = (e_1^*, e_2^*, e_3^*)$ la base duale de B . La forme linéaire $f = a_1 e_1^* + a_2 e_2^* + a_3 e_3^*$ est dans $F^\perp$ si et seulement si $f(e_1) = 0$, c'est-à-dire si et seulement si $a_1 = 0$. On en déduit $F^\perp = \text{Vect}(e_2^*, e_3^*)$; les éléments non nuls de $F^\perp$ correspondent aux plans contenant l'axe engendré par e_1 .

Orthogonal d'un sous-espace de E^* . Soit G un sous-espace vectoriel de E^* . On note $G^\perp$ le sous-ensemble de E des vecteurs orthogonaux à tous les éléments de G :

$$G^\perp = \{ u \in E \mid \forall f \in G : f(u) = 0 \}.$$

Les éléments de $G^\perp$ sont les vecteurs contenus dans tous les hyperplans de E définis par les éléments de G , autrement dit l'intersection des hyperplans définis par les éléments de G .

On peut généraliser ces définitions à des sous-ensembles quelconques de E et de E^* en définissant leurs orthogonaux comme les orthogonaux des sous-espaces qu'ils engendrent.

Proposition 1 : dimension de l'orthogonal d'un sous-espace de E . Soit F un sous-espace vectoriel de E ; alors $F^\perp$ est un sous-espace vectoriel de E^* et

$$\dim(F) + \dim(F^\perp) = \dim(E).$$

Démonstration. Posons $\dim(F) = r$ et soit $(e_1, \dots, e_r)$ une base de F . Complétons cette famille en une base $B = (e_1, \dots, e_n)$ de E et notons $B^* = (e_1^*, \dots, e_n^*)$ la base duale de B . Une forme linéaire $f = x_1 e_1^* + \dots + x_n e_n^*$ est dans $F^\perp$ si et seulement si $f(e_i) = 0$ pour $1 \leq i \leq r$, c'est-à-dire si et seulement si $x_1 = \dots = x_r = 0$. On a donc $F^\perp = \text{Vect}(e_{r+1}^*, \dots, e_n^*)$, donc $\dim(F^\perp) = n - r$ et l'égalité indiquée. $\square$

Proposition 2: dimension de l'orthogonal d'un sous-espace de E^* . Soit G un sous-espace vectoriel de E^* ; alors $G^\perp$ est un sous-espace vectoriel de E et

$$\dim(G) + \dim(G^\perp) = \dim(E).$$

Démonstration. Posons $\dim(E) = n$. Notons $B = (e_1, \dots, e_n)$ une base de E et $B^* = (e_1^*, \dots, e_n^*)$ la base duale de B .

Il est facile de vérifier que $G^\perp$ est un sous-espace vectoriel de E . Posons $\dim(G) = r$. Soient $f_1, \dots, f_r$ une base de G et posons :

$$f_i = a_{i1} e_1^* + \dots + a_{in} e_n^*$$

pour $1 \leq i \leq r$. L'algorithme du pivot permet de trouver une base de G qui soit une famille échelonnée par rapport à la base B^* . On peut donc se ramener au cas où $f_1, \dots, f_r$ est une base de G qui soit une famille triangulaire par rapport à une base B^* . Un vecteur $u = x_1 e_1 + \dots + x_n e_n$ est dans $G^\perp$ si et seulement si $f_i(u) = 0$ pour $1 \leq i \leq r$, c'est-à-dire si ses coordonnées vérifient le système linéaire L :

$$\left\{ \begin{array}{cccccc} a_{11}x_1 + & a_{12}x_2 + & \dots & \dots & \dots & +a_{1n}x_n & = 0 \\ & a_{22}x_2 + & \dots & \dots & \dots & +a_{2n}x_n & = 0 \\ & & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & a_{rr}x_r + & \dots & +a_{rn}x_n & = 0 \end{array} \right.$$

Le rang de L est r et la dimension de l'espace des solutions $S(L)$ est $n - \dim(G)$. On en déduit que $\dim(G^\perp) = n - r$, ce qui permet de conclure. $\square$

Proposition 3 : orthogonal de l'orthogonal.

$$(F^\perp)^\perp = F \qquad (G^\perp)^\perp = G$$

Démonstration. $(F^\perp)^\perp$ est l'espace des vecteurs de E orthogonaux à toutes les formes linéaires de $F^\perp$. Il contient donc F . Comme $\dim(F) + \dim(F^\perp) = n$ et $\dim(F^\perp) + \dim((F^\perp)^\perp) = n$, d'après les deux propositions précédentes, on voit que $\dim((F^\perp)^\perp) = \dim(F)$, donc $F = (F^\perp)^\perp$. On démontre de même la seconde égalité. $\square$

Une base de l'orthogonal $F^\perp$ d'un sous-espace est une famille de formes linéaires s'annulant sur F . Elle donne donc une famille d'équations qui décrivent F et les hyperplans définis par ces équations ont pour intersection F .

Exemple. Considérons les formes linéaires $f_1, f_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ définies pour $u = (x, y, z)$ par $f_1(u) = 2x + y + z$ et $f_2(u) = 5x + 3y + 2z$. Posons $G = \text{Vect}(f_1, f_2)$. Dans la base duale de la base canonique, on a $f_1 = 2e_1^* + e_2^* + e_3^*$ et $f_2 = 5e_1^* + 3e_2^* + 2e_3^*$, ce qui permet de voir que $\dim(G) = 2$. On en déduit $\dim(G^\perp) = 1$. On détermine $G^\perp$ comme l'ensemble des solutions du système :

$$\begin{cases} 2x + y + z = 0 \\ 5x + 3y + 2z = 0 \end{cases}$$

ce qui donne $G^\perp = \text{Vect}((1, -1, -1))$. Une forme linéaire de G s'écrit $af_1 + bf_2$. Il lui correspond un plan vectoriel de $\mathbb{R}^3$ qui contient la droite $\text{Vect}((1, -1, -1))$. De plus, tout plan contenant cette droite est défini par une forme linéaire de la forme $af_1 + bf_2$.

Proposition 4. Soit E un espace de dimension n et soit $B = (f_1, \dots, f_n)$ une base de E^* . Il existe une base $C = (e_1, \dots, e_n)$ de E telle que $B = C^*$.

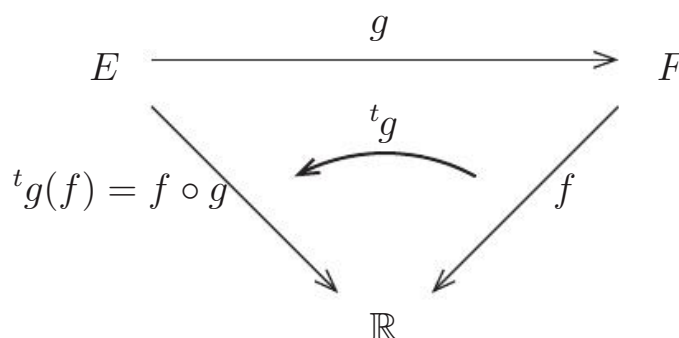
Démonstration. Pour $i = 1, \dots, n$, posons $F_i = \text{Vect}(B - \{f_i\})$. Comme $\dim(F_i) = n - 1$, on a $\dim(F_i^\perp) = 1$. Notons e_i un élément non nul de $F_i^\perp$. Il est facile de vérifier que la famille $C = (e_1, \dots, e_n)$ convient.

10.5 TRANSPOSÉE D'UNE APPLICATION LINÉAIRE

Soient E et F deux espaces vectoriels et $g : E \rightarrow F$ une application linéaire.

Définition 1 : transposée d'une application linéaire. On appelle transposée de l'application linéaire g , l'application ${}^t g : F^* \rightarrow E^*$ définie par ${}^t g(f) = f \circ g$.

La définition de ${}^t g$ se visualise bien dans le diagramme suivant.



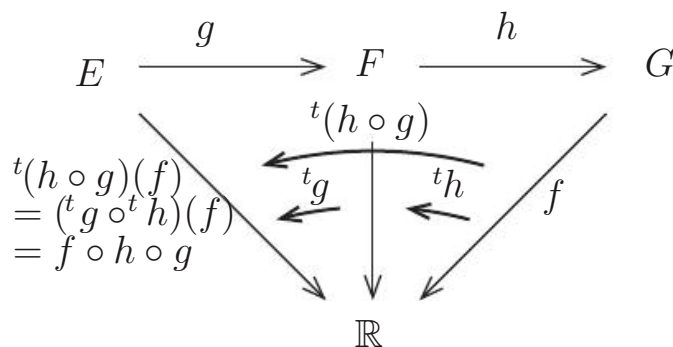
Propriétés.

1) La transposée d'une application linéaire est linéaire. Cela résulte des formules sur la compatibilité de la composition avec somme et produit par un scalaire (voir exercice 7.6) :

$${}^t g(f_1 + f_2) = (f_1 + f_2) \circ g = f_1 \circ g + f_2 \circ g = {}^t g(f_1) + {}^t g(f_2) ;$$

$${}^t g(af) = (af) \circ g = a(f \circ g) = a {}^t g(f)$$

2) Soient E, F et G trois espaces vectoriels et $g : E \rightarrow F, h : F \rightarrow G$ des applications linéaires. On a ${}^t(h \circ g) = {}^t g \circ {}^t h$, comme on le voit dans le diagramme :



Définition 2 : transposée d'une matrice. On appelle transposée d'une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ à n lignes et p colonnes la matrice à p lignes et n colonnes ${}^tA = (a_{ji})_{1 \leq j \leq p, 1 \leq i \leq n}$.

Par exemple, si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, alors ${}^tA = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$. En 1858, Cayley notait $\text{tr}.A$

la transposée de A .

Proposition 1 : matrice de la transposée d'une application linéaire. Soient E et F deux espaces vectoriels munis des bases $B = (e_1, \dots, e_p)$ et $C = (\varepsilon_1, \dots, \varepsilon_n)$ respectivement et $g : E \rightarrow F$ une application linéaire de matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ par rapport aux bases B et C .

Alors $M({}^t g, C^*, B^*) = {}^t A$.

Démonstration. Pour déterminer la matrice de ${}^t g$ par rapport aux bases indiquées, il suffit d'exprimer dans la base B^* les images par ${}^t g$ des vecteurs de la base C^* . Comme ces images sont des formes linéaires sur E , il faut calculer. On a :

$${}^t g(\varepsilon_j^*)(e_i) = (\varepsilon_j^* \circ g)(e_i) = \varepsilon_j^*(a_{i1}\varepsilon_1 + \dots + a_{in}\varepsilon_n) = a_{ji} = (a_{ji}\varepsilon_i^*)(e_i). \quad \square$$

Proposition 2 : propriétés de la transposition de matrices. Lorsque les calculs sont possibles, on a :

- 1) ${}^t({}^t A) = A$;
- 2) ${}^t(A + B) = {}^t A + {}^t B$;
- 3) ${}^t(\lambda A) = \lambda {}^t A$;
- 4) ${}^t(AB) = {}^t B {}^t A$.

Démonstration. Les trois premières propriétés sont faciles. Le 4) résulte de la propriété correspondante sur les applications linéaires : ${}^t(h \circ g) = {}^t g \circ {}^t h$. $\square$

► Vers le chapitre 11

Avant de continuer de développer les notions d'algèbre linéaire, nous allons présenter des notions d'algèbre de base (groupes, anneaux, polynômes), intéressantes en elles-mêmes et nécessaires pour étudier les chapitres suivants de ce livre.

Exercices

10.1 Petits exercices théoriques

Soit E un espace vectoriel de dimension finie.

- a) Déterminer $F^\perp$ pour $F = \{0\}$ et pour $F = E$.
- b) Déterminer $G^\perp$ pour $G = E^*$ et pour $G = \{0\}$, où 0 désigne la forme linéaire nulle.
- c) Soient F_1 et F_2 deux sous-espaces vectoriels de E , G_1 et G_2 deux sous-espaces vectoriels de E^* .

Montrer que $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$ et que $(G_1 + G_2)^\perp = G_1^\perp \cap G_2^\perp$. En déduire $F_1^\perp + F_2^\perp = (F_1 \cap F_2)^\perp$ et $G_1^\perp + G_2^\perp = (G_1 \cap G_2)^\perp$.

d) Soient $g_1, g_2 : E \rightarrow F$ deux applications linéaires. Montrer que ${}^t(g_1 + g_2) = {}^t g_1 + {}^t g_2$.

10.2 Exemples de formes linéaires

1) Soit a un nombre réel. Montrer que les « évaluations en a » suivantes sont des formes linéaires :

- a) l'application $\text{ev}_a : \mathbb{R}[X] \rightarrow \mathbb{R}$ définie par $\text{ev}_a(P) = P(a)$;
- b) l'application $\text{ev}_a : E \rightarrow \mathbb{R}$ définie par $\text{ev}_a(f) = f(a)$, où $E = \text{Fonct}(\mathbb{R}, \mathbb{R})$ note l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$.

2) Soit $E = \text{Cont}(I, \mathbb{R})$ l'espace des applications continues d'un intervalle $I = [a, b]$ dans $\mathbb{R}$.

- a) Montrer que l'application $\varphi : E \rightarrow \mathbb{R}$ définie par $\varphi(f) = \int_a^b f(t) dt$ est une forme linéaire sur E .
- b) Plus généralement, soit $p \in E$. Montrer que l'application $\psi : E \rightarrow \mathbb{R}$ définie par $\psi(f) = \int_a^b f(t)p(t) dt$ est une forme linéaire sur E .

3) Vérifier que la trace est une forme linéaire sur l'espace $M_n(K)$ des matrices carrées d'ordre n à coefficients dans un corps K .

10.3 Orthogonalité

On suppose $\mathbb{R}^3$ muni de la base canonique (e_1, e_2, e_3) et son dual muni de la base duale (e_1^*, e_2^*, e_3^*) de la base canonique.

1) Donner la dimension et une base des sous-espaces suivants de $\mathbb{R}^3$ et de leur orthogonal dans $(\mathbb{R}^3)^*$:

- a) $F_1 = \{(x, y, z) \mid y = 0\}$;
- b) $F_2 = \{(x, y, z) \mid x - 3y + 2z = 0 \text{ et } 2x - y + z = 0\}$;
- c) $F_3 = \text{Vect}((1, 2, 3), (4, -3, 1))$

2) Donner la dimension et une base des sous-espaces suivants de $(\mathbb{R}^3)^*$ et de leur orthogonal dans $\mathbb{R}^3$:

- a) $G_1 = \text{Vect}(e_1^* - e_2^* + 3e_3^*)$;
- b) $G_2 = \text{Vect}(e_1^* - e_2^* + 3e_3^*, 2e_1^* - e_2^* + 5e_3^*, 3e_1^* + 7e_3^*)$;

3) Déterminer la base duale de la base $(e_1^* - e_2^* + 3e_3^*, 2e_1^* - e_2^* + 5e_3^*, 3e_1^* + 7e_3^*)$.

10.4 Polynômes

1) a) Déterminer la base duale de la base des polynômes de Lagrange $(P_1, \dots, P_{n+1})$ de l'exercice 6.8.

b) Formule de calculs des intégrales de polynômes

Soient $[a, b]$ un intervalle de $\mathbb{R}$. Soit n un entier et $a_1, \dots, a_{n+1}$ des points de $[a, b]$. Montrer qu'il existe des réels $A_1, \dots, A_{n+1}$ tels que, pour tout polynôme de degré $\leq n$, on ait :

$$\int_a^b P(t) dt = A_1 P(a_1) + \dots + A_n P(a_n).$$

2) Donner un système d'équations linéaires, en nombre minimum, définissant, dans $\mathbb{R}_4[X]$ muni de la base $(1, X, X^2, X^3, X^4)$:

a) le sous-espace des polynômes pairs ;

b) le sous-espace des polynômes P tels que $P(0) = P''(0) = 0$.

10.5 Formule des trois niveaux

Cette formule est un cas particulier de la formule générale ci-dessus. On pose $E = \mathbb{R}_2[X]$.

a) Exprimer la forme linéaire j sur E définie par : $P \mapsto \int_{-1}^1 P(t) dt$ en fonction des formes ev_{-1} , ev_0 et ev_1 .

b) En déduire une formule donnant la valeur exacte de $\int_{-1}^1 P(t) dt$ pour tout P de E en fonction des valeurs de P en $-1, 0, 1$.

c) Soient a et b des réels. Donner une formule analogue pour le calcul de $\int_a^b P(t) dt$ pour tout P de E , en utilisant les évaluations en $a, b, \frac{a+b}{2}$.

10.6 Changement de base et espace dual

Soient K un corps, E un K -espace vectoriel de dimension finie n , $B = (e_1, \dots, e_n)$, $C = (f_1, \dots, f_n)$ deux bases de E . On note $P = (\alpha_{ij})$ la matrice de passage $M(\text{id}, C, B)$. Quelle est la matrice de passage $M(\text{id}, B^*, C^*)$? Comment calculer les vecteurs de la base C^* ?

10.7 Bidual

Soient K un corps et E un K -espace vectoriel. On note E^* le dual de E et E^{**} le dual de E^* , c'est-à-dire l'espace vectoriel $L(E^*, K)$ des formes linéaires sur E^* , qu'on a appelé bidual de E et noté E^{**} .

On définit une application $j_E : E \rightarrow E^{**}$ en définissant, pour $u \in E$, l'application $j_E(u) : E^* \rightarrow E$ par $j_E(u)(f) = f(u)$ pour $f \in E^*$.

- 1) Montrer que $j_E(u)$ est une forme linéaire sur E^* .
- 2) Montrer que j_E est une application linéaire.
- 3) On suppose que E est de dimension finie. Montrer que j_E est un isomorphisme d'espaces vectoriels (on remarquera que cet isomorphisme ne dépend pas du choix d'une base de E ; on dit qu'il est canonique).
- 4) Si $h : E \rightarrow F$ est une application linéaire (avec E et F de dimensions finies), définir l'application h^{**} rendant le diagramme suivant commutatif :

$$\begin{array}{ccc} E & \xrightarrow{j_E} & E^{**} \\ h \downarrow & & \downarrow h^{**} \\ F & \xrightarrow{j_F} & F^{**} \end{array}$$

Solutions

10.1 a) $\{0\}^\perp = E^*$; $E^\perp = \{0\}$.

b) $(E^*)^\perp = \{0\}$; $\{0\}^\perp = E$.

c) Si f s'annule sur $F_1 + F_2$, elle s'annule sur F_1 et sur F_2 , donc

$$(F_1 + F_2)^\perp \subset F_1^\perp \cap F_2^\perp.$$

Réciproquement, si f s'annule sur F_1 et sur F_2 , alors f s'annule sur $F_1 + F_2$, car si $v \in F_1 + F_2$, v s'écrit $v_1 + v_2$ avec $v_1 \in F_1$ et $v_2 \in F_2$ et $f(v) = f(v_1) + f(v_2) = 0$.

Donc f est élément de $(F_1 + F_2)^\perp$. Finalement, $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$.

En appliquant cette formule à $G_1^\perp$ et $G_2^\perp$ on obtient :

$$(G_1^\perp + G_2^\perp)^\perp = (G_1^\perp)^\perp \cap (G_2^\perp)^\perp = G_1 \cap G_2, \text{ d'où } G_1^\perp + G_2^\perp = (G_1 \cap G_2)^\perp.$$

On en déduit les deux autres formules ; on peut aussi raisonner avec des éléments.

d) Soit f une forme linéaire quelconque sur F . On a

$$\begin{aligned} {}^t(g_1 + g_2)(f) &= f \circ (g_1 + g_2) = f \circ g_1 + f \circ g_2 \\ &= {}^tg_1(f) + {}^tg_2(f) = ({}^tg_1 + {}^tg_2)(f). \end{aligned}$$

10.2 Les exemples sont importants et les vérifications faciles.

10.3 1) a) Base de F_1 : (e_1, e_3) ; base de $F_1^\perp$: (e_2^*) .

b) Base de F_2 : $(-e_1 + 3e_2 + 5e_3)$; base de $F_2^\perp$: $(e_1^* - 3e_2^* + 2e_3^*, 2e_1^* - e_2^* + e_3^*)$.

c) Base de F_3 : $((1, 2, 3), (4, -3, 1))$; base de $F_3^\perp$: $(e_1^* + e_2^* - e_3^*)$.

2) a) Base de $G_1^\perp$: $((1, 1, 0), (-3, 0, 1))$.

b) $G_2^\perp = \{0\}$, base vide.

3) Le premier vecteur $u_1 = (x, y, z)$ de la base cherchée vérifie :

$$(e_1^* - e_2^* + 3e_3^*)(u_1) = 1, (2e_1^* - e_2^* + 5e_3^*)(u_1) = 0, (3e_1^* + 7e_3^*)(u_1) = 0$$

ce qui conduit à la résolution du système linéaire :

$$\begin{cases} x & -y & +3z & = 1 \\ 2x & -y & +5z & = 0 \\ 3x & & +7z & = 0 \end{cases}$$

On trouve $u_1 = (-7, 1, 3)$. On trouve de même $u_2 = (7, -2, -3)$ et $u_3 = (-2, 1, 1)$.

10.4 1) a) La forme linéaire ev_{a_1} qui évalue les polynômes en a_1 s'annule pour tous les P_i , $i \geq 2$ et $\text{ev}_{a_1}(P_1) = 1$. Plus généralement : $\text{ev}_{a_i}(P_j) = 0$ si $i \neq j$ et $\text{ev}_{a_i}(P_i) = 1$. La base duale de la base des polynômes de Lagrange est la base formée par les évaluations aux points a_i .

b) Il suffit d'écrire la forme linéaire $P \mapsto \int_a^b P(t) dt$ définie sur $\mathbb{R}_n[X]$ dans la base précédente.

2) Écrivons $P = a_0 \cdot 1 + a_1 X + a_2 X^2 + a_3 X^3 + a_4 X^4$. On trouve les équations :

a) $a_1 = a_3 = 0$.

b) $a_0 = a_2 = 0$.

10.5 a) On a $J = A_{-1} \text{ev}_{-1} + A_0 \text{ev}_0 + A_1 \text{ev}_1$. Pour trouver les trois coefficients, on écrit les valeurs des deux membres en les polynômes d'une base de $\mathbb{R}_2[X]$ comme $(1, X, X^2)$. On trouve :

$$\begin{cases} 2 = J(1) = A_{-1} + A_0 + A_1 \\ 0 = J(X) = -A_{-1} + A_1 \\ \frac{2}{3} = J(X^2) = A_{-1} + A_1 \end{cases}$$

d'où $A_1 = A_{-1} = \frac{1}{3}$ et $A_0 = \frac{4}{3}$.

$$\text{b)} \int_{-1}^1 P(t) dt = \frac{1}{3}(P(-1) + 4P(0) + P(1)).$$

c) Le changement de variable $t \mapsto \frac{2t-a-b}{b-a}$ donne :

$$\int_a^b P(t) dt = \frac{b-a}{6}(P(a) + 4P\left(\frac{a+b}{2}\right) + P(b))$$

pour tout P de E . On montre que cette formule appliquée à une fonction f de classe C^4 sur $[a, b]$ donne une valeur qui diffère de l'intégrale exacte de moins de $\left(\frac{b-a}{2}\right)^5 \frac{M}{90}$ où M est un majorant de $|f^{(4)}(t)|$ sur l'intervalle $[a, b]$ (Voir J.-P.

Escofier, *Toute l'analyse de la licence*, Dunod, 2014, exercice 11.11).

10.6 Posons $M(\text{id}, B^*, C^*) = (\beta_{kl})$, d'où, pour $1 \leq l \leq n$, $e_l^* = \sum_{1 \leq k \leq n} \beta_{kl} f_k^*$. On a $\beta_{kl} = e_l^*(f_k) = e_l^*\left(\sum_{1 \leq i \leq n} \alpha_{ik} e_i\right) = \alpha_{lk}$.

Par conséquent, la matrice de passage $Q = M(\text{id}, B^*, C^*)$ est la transposée de P , un résultat qui permet de calculer la base C^* en inversant Q . Les vecteurs de C^* sont les colonnes de la matrice ${}^t P^{-1}$ (ou les lignes de la matrice P^{-1}).

10.7 1,2) On vérifie successivement que $j_E(u)$ puis j_E sont des applications linéaires.

3) Comme les dimensions de E , E^* et E^{**} sont égales, il suffit de montrer que j_E est injective, ce qui revient à montrer que son noyau est réduit à 0. Si $u \in \ker(j_E)$, $j_E(u) : E^* \rightarrow K$ est l'application nulle, c'est-à-dire que, pour tout $f \in E^*$, on a $j_E(u)(f) = f(u) = 0$. Mais, si $u \neq 0$, on peut construire une forme linéaire ne s'annulant pas en u (compléter u en une base de E et définir f sur cette base avec $f(u) = 1$). On a donc $u = 0$ et j_E est un isomorphisme.

Cet isomorphisme ne dépend pas du choix d'une base de E . C'est l'exemple sur lequel s'appuient Eilenberg et Mac Lane, dans leur article fondateur de la théorie des catégories en 1945, où ils qualifient de *naturel* cet isomorphisme.

4) Un élément de E^{**} est de la forme $j_E(u)$. On définit son image par h^{**} comme étant $j_F(u)$, ce qui revient à poser $h^{**} = j_F \circ h \circ j_E^{-1}$. Si on identifie E et F à leurs biduaux, on peut écrire $h^{**} = h$.

Deuxième année

Cette deuxième partie du cours d'algèbre correspond à ce qu'on peut étudier en deuxième année. Nous allons aborder de nouveaux objets essentiels pour faire de l'algèbre et, plus généralement, des mathématiques : groupes, anneaux, corps, polynômes. Nous reviendrons ensuite à l'algèbre linéaire pour présenter les déterminants, donner quelques résultats sur la réduction des matrices et, enfin, parler des problèmes d'orthogonalité dans les espaces euclidiens.

Le style de rédaction va être un peu plus rapide, demandant un peu plus de travail aux lectrices et lecteurs. Bon courage...

11.1 INTRODUCTION

Les premiers chapitres ont présenté les bases de l'algèbre linéaire. Nous avons cherché à montrer comment différents problèmes d'algèbre linéaire peuvent être placés dans un cadre unificateur, celui des espaces vectoriels et des applications linéaires. Nous espérons avoir convaincu de l'intérêt de ce mouvement vers l'*abstraction* et avoir montré qu'il permet d'unifier la résolution de problèmes d'origines diverses et d'en comprendre la nature profonde.

L'histoire de cette marche vers l'abstraction doit beaucoup aux idées de Galois sur l'étude de la résolubilité par radicaux des équations algébriques. Les permutations des racines d'un polynôme donné forment un groupe et Galois montre que le problème de la résolubilité dépend des sous-groupes de ce groupe et de leurs propriétés. La beauté et la profondeur des idées de Galois allaient influencer toutes les mathématiques.

Depuis Galois, la théorie des groupes s'est considérablement développée. Les travaux de Jordan (le grand *Traité des substitutions*, de plus de 650 pages, date de 1870) sont particulièrement importants, avec la définition des notions d'homomorphisme, d'isomorphisme, l'étude des groupes de permutations, des groupes de l'algèbre linéaire, etc.

Nous ne donnerons que les premières notions sur cette théorie. Cela nous permettra déjà d'obtenir des résultats d'arithmétique et d'étudier le groupe des permutations, utile pour développer au chapitre 14 la théorie des déterminants, ainsi que de décrire des groupes de l'algèbre linéaire.

11.2 GÉNÉRALITÉS

Nous avons rencontré des groupes dans le début de ce livre : groupe abélien des éléments d'un espace vectoriel avec pour loi la somme des vecteurs, et groupe linéaire (voir 7.10), non commutatif en général. Voici la définition de ces notions, mise au point il y a une centaine d'années après de nombreuses propositions de définitions plus compliquées.

Définition 1 : groupes. On dit qu'un couple $(G, \cdot)$ où le point désigne une loi $G \times G \rightarrow G$ est un groupe si la loi notée par un point, qu'on appellera produit et qu'on notera souvent par une simple juxtaposition, comme la multiplication usuelle, possède les propriétés suivantes.

➤ **Associativité :**

pour tous x, y, z de G , on a $x(yz) = (xy)z$

➤ **Existence d'un élément neutre :**

il existe un élément noté e dans G tel que : pour tout x de G , on a :

$$ex = xe = x$$

➤ **Existence d'un élément inverse :**

pour tout x de E , il existe un élément y de E , qu'on notera x^{-1} , tel que :

$$xy = yx = e$$

On note le plus souvent un groupe $(G, \cdot)$ par G en sous-entendant la loi de groupe.

Conséquences immédiates. L'associativité permet de ne pas mettre de parenthèses dans des produits de plusieurs termes : par exemple $xyztu$ désigne aussi bien le résultat du calcul $(xy)(z(tu))$ que celui du calcul $x((yz)(tu))$.

L'élément neutre est unique : si e et e' vérifient la propriété d'être élément neutre, on a $e = ee' = e'$.

On peut faire des simplifications comme :

$$\text{si } xy = xz \text{ alors } y = z,$$

$$\text{si } yx = zx \text{ alors } y = z,$$

il suffit de multiplier par x^{-1} à gauche ou à droite les deux membres. On en déduit que l'inverse d'un élément est unique.

Enfin, l'inverse de x^{-1} est x et $(xy)^{-1} = y^{-1}x^{-1}$.

Définition 2 : groupes commutatifs ou abéliens. Le groupe est dit commutatif (on dit aussi abélien) si la loi de groupe est commutative, c'est-à-dire si la propriété supplémentaire suivante est vérifiée.

➤ **Commutativité :**

pour tous x, y de G , on a $xy = yx$

Le plus souvent, mais pas toujours, une loi commutative est notée avec un $+$.

Voici les axiomes précédents et leurs conséquences réénoncés en utilisant cette notation.

Dire que le couple $(G, +)$ est un groupe abélien signifie que la somme possède les propriétés suivantes :

➤ **Associativité**

pour tous x, y, z de G , on a $x + (y + z) = (x + y) + z$

➤ **Commutativité :**

pour tous x, y de G , on a $x + y = y + x$

➤ **Existence d'un élément neutre pour l'addition :**

il existe un élément noté 0 dans G tel que, pour tout x de G :

$$x + 0 = x$$

➤ **Existence d'un élément opposé :**

pour tout x de G , il existe un élément y de G tel que :

$$x + y = 0$$

Conséquences immédiates. Comme ci-dessus, l'associativité permet de ne pas mettre de parenthèses dans des sommes de plusieurs termes, la commutativité permet de changer l'ordre des termes dans une somme.

Compte tenu de la commutativité de la somme, on a :

$$\text{pour tout } x \text{ de } G : x + 0 = x \text{ et } 0 + x = x$$

L'élément opposé de x est noté $-x$ et la somme $x + (-y)$ est notée $x - y$. On peut donc écrire $x - x = 0$.

Enfin, on peut faire des simplifications comme :

$$\text{si } x + y = x + z \text{ alors } y = z,$$

il suffit d'ajouter $-x$ aux deux membres. Par conséquent, $x + y = y$ implique $x = 0$.

Puissances d'un élément. Soient $(G,.)$ un groupe et x un élément de G . On pose $x^2 = xx$, $x^3 = xxx$, $x^{-2} = x^{-1}x^{-1}$. Plus généralement, on définit x^k et x^{-k} comme étant les produits de x ou de x^{-1} k fois par eux-mêmes (la définition précise est une construction par récurrence). On vérifie que x^{-k} est l'inverse de x^k et que $x^{k+l} = x^k x^l$ pour tout k et tout l de $\mathbb{Z}$. En notation additive, on pose $2x = x + x$, $-2x = -x - x$,... et la dernière formule s'écrit $(k + l)x = kx + lx$.

Définition 3 : éléments conjugués. Soient G un groupe, et a et b deux éléments de G . Les éléments a et b sont dits conjugués dans G s'il existe un élément x de G tel que $a = x^{-1}bx$. La relation de conjugaison est une relation d'équivalence. Si le groupe est abélien, le seul conjugué de a est a lui-même.

11.3 EXEMPLES

Exemples de groupes abéliens.

- Le groupe $(\mathbb{Z}, +)$ des entiers avec l'addition, les groupes $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$; par contre, $(\mathbb{N}, +)$ n'est pas un groupe, puisque l'opposé d'un entier strictement positif pour l'addition n'existe pas dans $\mathbb{N}$;
- les groupes $(\mathbb{Q}^*, .)$, $(\mathbb{R}^*, .)$, $(\mathbb{C}^*, .)$ des éléments non nuls de $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ avec la loi de multiplication ; plus généralement, le groupe multiplicatif K^* des éléments non nuls d'un corps K (la définition précise d'un corps est donnée au chapitre 12) ; par contre, $(\mathbb{Z}^*, .)$ n'est pas un groupe ;
- le groupe $\{-1, 1\}$ formé par les deux éléments -1 et 1 de $\mathbb{Z}$ avec la multiplication ;
- les groupes $(\frac{\mathbb{Z}}{n\mathbb{Z}}, +)$ que nous allons étudier au chapitre suivant ;
- le groupe abélien défini par l'ensemble des vecteurs d'un espace vectoriel et la loi d'addition ; nous ne l'avons pas étudié en tant que tel dans les chapitres précédents, c'est la structure d'espace vectoriel qui nous a intéressé.

Exemples de groupes non abéliens (en général).

- Le groupe (appelé groupe linéaire) des automorphismes d'un espace vectoriel et la loi de composition des applications linéaires (voir 11.8) ;
- une autre façon de voir le groupe précédent est de le considérer comme le groupe des matrices carrées inversibles d'ordre n avec la loi de multiplication des matrices ;
- le groupe S_n des bijections (on dit permutations) de l'ensemble des entiers $\{1, \dots, n\}$ (voir 11.6).

11.4 SOUS-GROUPES

11.4.1 Sous-groupes

Définition 1. Soit $(G, \cdot)$ un groupe. On appelle sous-groupe de G , un sous-ensemble H de G tel que :

- H est stable par la multiplication : pour tout x et tout y de H , on a $xy \in H$;
- l'élément neutre de G est dans H ;
- H est stable par inversion : pour tout x de H , l'inverse de x pour la loi de G est dans H ,

l'ensemble H étant muni de la loi de groupe induite par la loi de G .

Les sous-ensembles $\{e\}$ et G d'un groupe G définissent des sous-groupes de G .

Proposition 1 : intersection de sous-groupes. Soit G un groupe.

- 1) Soient H et K des sous-groupes de G . Alors $H \cap K$ est un sous-groupe de G .
- 2) Plus généralement, soit $(H_i)_{i \in I}$ une famille de sous-groupes de G . Alors $\bigcap_{i \in I} H_i$ est un sous-groupe de G .

Démonstration. Il suffit de vérifier les trois propriétés de la définition. □

11.4.2 Sous-groupe engendré

On a vu l'intérêt de la notion de famille génératrice dans les espaces vectoriels. Dans les groupes, il est utile de connaître des ensembles d'éléments engendrant le groupe.

Proposition 2 : sous-groupe engendré par un élément. Soit $(G, \cdot)$ un groupe et x un élément de G . L'ensemble $\{x^k \mid k \in \mathbb{Z}\}$ des puissances de x est un sous-groupe de G appelé sous-groupe engendré par x et noté $\text{Gr}(x)$.

Démonstration. Il est facile de vérifier les trois propriétés de la définition. □

Tout sous-groupe de G contenant x contient $\text{Gr}(x)$. Par conséquent, $\text{Gr}(x)$ est égal à l'intersection des sous-groupes de G contenant x .

Définition 2 : sous-groupe engendré. Généralisons. Soit A une partie de G . Pour tout entier $k \geq 0$, considérons l'ensemble E_k des produits finis $x_1 \dots x_k$ où les x_i ou leurs inverses sont des éléments de A . Pour $k = 0$, cet ensemble se réduit à $\{e\}$.

Une fois habitué aux notations, on voit, en vérifiant les trois propriétés de la définition, que $\bigcup_{k \geq 0} E_k$ est un sous-groupe de G : c'est l'ensemble des éléments de G qu'on peut *fabriquer* avec les éléments de A . On note $\text{Gr}(A)$ ce sous-groupe et on l'appelle sous-groupe engendré par A .

Proposition 3 : sous-groupe engendré. $\text{Gr}(A)$ est l'intersection des sous-groupes de G contenant A ; c'est aussi le plus petit sous-groupe de G contenant A .

Démonstration. Laissée aux lecteurs et lectrices. □

Le groupe engendré par l'ensemble vide est réduit à l'élément neutre.

Définition 3 : ordre d'un groupe. Soit $(G, \cdot)$ un groupe. Si le cardinal de G est fini, autrement dit si G a un nombre fini d'éléments, on dit que G est un groupe d'ordre fini. Si $\text{card}(G) = n$, on dit que G est un groupe d'ordre n . On note $|G|$ l'ordre d'un groupe fini.

Définition 3 : ordre d'un élément. Soit G un groupe et x un élément de G . On appelle ordre de x dans G , l'ordre du groupe $\text{Gr}(x)$.

L'ordre de l'élément neutre est 1. L'ordre d'un élément peut-être infini, comme l'ordre de 1 dans $(\mathbb{Z}, +)$. Si x est d'ordre t , les éléments de $\text{Gr}(x)$ sont $e, x, \dots, x^{t-1}$ et t est le plus petit des entiers $k > 0$ tels que $x^k = e$.

Théorème de Lagrange. Soit G un groupe fini.

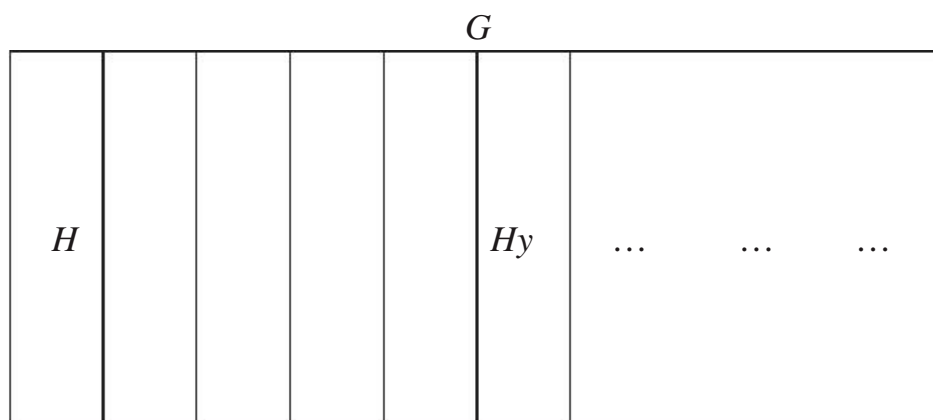
Soit H un sous-groupe de G . Alors $|H|$ divise $|G|$.

En particulier, si $x \in G$, l'ordre de x divise $|G|$.

Démonstration. L'idée de la démonstration remonte à Joseph Louis Lagrange (1736-1813) qui, en 1770, n'avait pas encore la notion de groupe à sa disposition, mais le problème sur les équations qu'il se posait était fondamentalement un problème de groupes.

On définit une relation $\mathcal{R}$ sur G en posant $x \mathcal{R} y$ si $xy^{-1} \in H$. Cette relation est une relation d'équivalence sur G , comme il est facile de le vérifier. On sait qu'une relation d'équivalence définit une partition de G en classes d'équivalence. Ces classes d'équivalence ont toutes $|H|$ éléments car l'ensemble des x équivalents à un y donné est $\{x | x \mathcal{R} y\} = \{x | xy^{-1} \in H\} = \{x | \exists h \in H (x = hy)\} = \{hy | h \in H\}$; nous noterons H_y cet ensemble. Comme $h \mapsto hy$ est une bijection de H sur H_y , H_y a $|H|$ éléments quel que soit y . La partition de G est donc formée d'ensembles ayant le même

nombre $|H|$ d'éléments, ce qui prouve que $|G|$ est un multiple de $|H|$. Un dessin peut illustrer cette partition :



La seconde affirmation se déduit de la première, puisque l'ordre d'un élément x est égal à l'ordre du sous-groupe $\text{Gr}(x)$. $\square$

11.5 HOMOMORPHISMES DE GROUPES

Définition 1 : homomorphismes de groupes. Soient deux groupes dont les lois sont notées $*$ et $'$; on appelle homomorphisme de G dans G' toute application $f : G \rightarrow G'$ telle que, pour tout x et tout y de G , on ait $f(x * y) = f(x) *' f(y)$.

Commentaire. Explicitons : cette condition s'écrit :

- $f(xy) = f(x)f(y)$ si les lois sont notées multiplicativement ;
- $f(x + y) = f(x) + f(y)$ si les deux lois sont notées additivement ;
- $f(x + y) = f(x)f(y)$ si la loi de G est notée additivement et celle de G' multiplicativement, etc.

Propriétés. Si e est l'élément neutre de G et e' celui de G' , on a $f(e) = e'$ car si x est un élément de G , on a $f(x) = f(xe) = f(x)f(e)$, d'où le résultat en simplifiant par $f(x)$.

Si x est un élément de G , on a $f(x^{-1}) = (f(x))^{-1}$ car $f(x)f(x^{-1}) = f(xx^{-1}) = f(e) = e'$.

Image. On vérifie facilement que si H est un sous-groupe de G , alors $f(H)$ est un sous-groupe de G' . En particulier, l'image de f , $\text{Im}(f) = f(G)$, est un sous-groupe de G' .

Image réciproque d'un sous-groupe. On vérifie aussi que, si H est un sous-groupe de G' , $f^{-1}(H)$ est un sous-groupe de G , puisque $f(e) = e' \in H$, $f(x^{-1}) = (f(x))^{-1} \in H$ si $x \in f^{-1}(H)$, etc.

Noyau. Le noyau de f , noté $\ker(f)$ et défini par $\ker(f) = \{x \in G \mid f(x) = e'\}$, est un sous-groupe de G , puisque c'est l'image réciproque du sous-groupe $\{e'\}$ de G' .

Proposition 1 : critère d'injectivité. *Un homomorphisme de groupes $f: G \rightarrow G'$ est injectif si et seulement si $\ker(f) = \{e\}$ où e est l'élément neutre de G .*

Démonstration. On a toujours $f(e) = e'$ donc, si f est injectif, on a nécessairement $\ker(f) = \{e\}$. Réciproquement, si $f(x) = f(y)$, alors $f(xy^{-1}) = e'$ donc $xy^{-1} = e$ donc $x = y$. $\square$

Cette proposition, analogue à celle vue en 7.5 pour les applications linéaires, donne un moyen simple de vérifier qu'un homomorphisme de groupes est injectif.

Définition 2 : isomorphisme de groupes. Un homomorphisme de groupes $f: G \rightarrow G'$ est appelé isomorphisme s'il existe un homomorphisme $g: G' \rightarrow G$ tel que $g \circ f = \text{id}(G)$ et $f \circ g = \text{id}(G')$.

Proposition 2 : critère pour être un isomorphisme. *Un homomorphisme de groupes $f: G \rightarrow G'$ tel que f soit bijective est un isomorphisme.*

Démonstration. Comme f est bijective, il existe une application $g: G' \rightarrow G$ telle que $g \circ f = \text{id}(G)$ et $f \circ g = \text{id}(G')$. Il faut vérifier que g est un homomorphisme de groupes. Si y et y' sont des éléments de G' , il existe x et x' dans G tels que $y = f(x)$, $x = g(y)$, $y' = f(x')$, $x' = g(y')$. On a $g(yy') = g(f(x)f(x')) = g(f(xx')) = xx' = g(y)g(y')$. $\square$

La propriété suivante est très utile pour définir un homomorphisme de source le groupe $(\mathbb{Z}, +)$.

Proposition 3 : propriété universelle de $(\mathbb{Z}, +)$. *Soient $G = (G, \cdot)$ un groupe et x un élément de G . Il existe un unique homomorphisme $f: (\mathbb{Z}, +) \rightarrow G$ tel que $f(1) = x$. C'est l'homomorphisme défini par $f(k) = x^k$ pour tout entier k . Si la loi de G est notée additivement, f est défini par $f(k) = kx$ pour tout entier k .*

Démonstration. Existence : on vérifie facilement qu'en posant $f(k) = x^k$, on définit bien un homomorphisme de groupe : pour tout couple (k, l) d'entiers de $\mathbb{Z}$, on a : $f(k + l) = x^{k+l} = x^k x^l = f(k) f(l)$, d'après la formule sur les puissances.

Unicité : si $g: \mathbb{Z} \rightarrow G$ est un homomorphisme tel que $g(1) = x$, on voit par récurrence que $g(k) = f(k)$ pour tout entier $k \geq 1$, puis on en déduit que $g(k) = f(k)$ pour tout entier k de $\mathbb{Z}$. $\square$

On remarquera que l'image de l'homomorphisme f de la proposition précédente est le sous-groupe $\text{Gr}(x)$ engendré par x .

11.6 ÉTUDE DES GROUPES DE PERMUTATIONS

Les résultats de ce paragraphe seront utiles dans le chapitre sur les déterminants. Beaucoup ont d'ailleurs été obtenus pour clarifier leur étude.

Définition 1 : permutations, groupe de permutations et groupe symétrique.

Soit E un ensemble. L'ensemble S_E des bijections de E muni de la loi de composition des applications est un groupe (si $E = \emptyset$, S_E est réduit à un élément). En particulier, pour tout entier $n > 0$, l'ensemble des bijections de l'ensemble $\{1, \dots, n\}$ dans lui-même muni de la loi de composition des applications est un groupe qu'on notera S_n et qu'on appellera groupe symétrique. On appellera permutations les éléments de S_n (Jordan parlait de substitutions).

Notation. Pour simplifier la notation, on notera le composé par juxtaposition : $\sigma'\sigma$ signifie $\sigma' \circ \sigma$, en n'oubliant pas que ce produit n'est pas commutatif en général.

Proposition 1. Soit E un ensemble ayant n éléments. Le groupe S_E est isomorphe au groupe S_n .

Démonstration. Puisque E a n éléments, il existe une bijection $\varphi : \{1, \dots, n\} \rightarrow E$. On définit une application $\Phi : S_E \rightarrow S_n$ en considérant le diagramme :

$$\begin{array}{ccc} E & \xrightarrow{f} & E \\ \varphi \uparrow & & \downarrow \varphi^{-1} \\ \{1, \dots, n\} & \xrightarrow{\Phi(f)} & \{1, \dots, n\} \end{array}$$

On pose $\Phi(f) = \varphi^{-1} \circ f \circ \varphi$ et on vérifie que Φ est bien un isomorphisme. $\square$

Pour pouvoir faire des calculs dans le groupe S_n , la première chose à faire est de choisir des notations agréables pour les permutations. Cauchy, dans ses mémoires de 1812, ne donne pas de notation. Galois, qui a lu les mémoires de Cauchy, écrit l'un au-dessous de l'autre les éléments de E et leurs images, ce qui conduit à l'écriture :

$$\sigma = \begin{pmatrix} x_1 & \dots & x_n \\ \sigma(x_1) & \dots & \sigma(x_n) \end{pmatrix}$$

pour une permutation σ de S_n . Par exemple, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix}$ est la permutation de $\{1, 2, 3, 4, 5\}$ telle que $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 4$, $\sigma(4) = 3$, $\sigma(5) = 5$.

Définition 2 : transposition. On appelle transposition de S_n une permutation de S_n qui échange deux éléments de $\{1, \dots, n\}$ et laisse les autres fixes. La notation précédente serait lourde pour une telle permutation et on note souvent $\tau = (i\ j)$ la transposition définie par $\tau(i) = j$, $\tau(j) = i$ et, pour tout entier $k \neq i, j$, $\tau(k) = k$. L'entier n n'apparaît pas dans cette notation, mais le contexte ne laisse, en général, aucune ambiguïté. Comme $\tau\tau = \text{id}$, τ est un élément d'ordre 2 de S_n .

Proposition 2 : ordre de S_n . On a : $|S_n| = n!$.

Démonstration. Montrons la formule par récurrence sur n . Le résultat est vrai pour $n = 0$ et $n = 1$; supposons qu'il soit vrai pour un entier $n \geq 1$. Pour compter les éléments de S_{n+1} , on définit une partition de S_{n+1} en $n + 1$ sous-ensembles disjoints E_k , $1 \leq k \leq n + 1$: E_k est l'ensemble des permutations σ de S_{n+1} telles que $\sigma(n+1) = k$. L'application $\sigma \mapsto (k\ n+1)\sigma$ (ici $(k\ n+1)$ désigne la transposition échangeant k et $n+1$) définit une bijection de E_k sur E_{n+1} , ce qui prouve que les $n + 1$ ensembles E_k ont tous le même nombre d'éléments. Puisque les éléments de E_{n+1} laissent $n + 1$ fixe, on peut les identifier à des permutations de $\{1, \dots, n\}$. D'après l'hypothèse de récurrence, le cardinal de E_{n+1} est $n!$, donc l'ordre de S_{n+1} est $(n + 1)n! = (n + 1)!$.

Définition 3 : cycles. On appelle cycle de longueur $r > 1$ de S_n , ou r -cycle, une permutation $c \in S_n$ telle qu'il existe des éléments $x_1, \dots, x_r$ de $\{1, \dots, n\}$ vérifiant $c(x_1) = x_2, \dots, c(x_{r-1}) = x_r$ et $c(x_r) = x_1$ et telle que c laisse fixes les autres éléments de $\{1, \dots, n\}$. La notation usuelle pour un cycle est $c = (x_1 \dots x_r)$, l'entier n étant sous-entendu. Par exemple, le cycle $(1\ 3\ 5)$ de S_5 est la permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 4 & 1 \end{pmatrix}$$

En particulier, une transposition est un cycle de longueur 2.

On appelle cycles disjoints des cycles $(x_1 \dots x_r)$ et $(y_1 \dots y_s)$ tels que $\{x_1, \dots, x_r\} \cap \{y_1, \dots, y_s\} = \emptyset$. Des cycles disjoints commutent entre eux.

Un cycle de longueur r est d'ordre r et le produit de deux cycles disjoints de longueurs r et s est d'ordre $\text{ppcm}(r, s)$ (le ppcm est défini en 12.12).

Proposition 3 : décomposition en produit de cycles disjoints. Toute permutation se décompose en un produit de cycles disjoints (de manière unique à l'ordre des cycles près).

Exemple. Plutôt que de donner une démonstration (on peut raisonner par récurrence ou voir la proposition 4 de 18.7), montrons sur un exemple comment cette décomposition s'obtient facilement. Considérons la permutation σ de S_{11} définie par :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 3 & 9 & 5 & 2 & 1 & 6 & 8 & 7 & 10 & 4 & 11 \end{pmatrix}$$

On regarde l'image de 1, l'image de cette image, etc., jusqu'à ce qu'on revienne à 1 ; on obtient le cycle $c = (1 \ 3 \ 5)$. Comme il existe des éléments en dehors de $\{1, 3, 5\}$ que σ ne laisse pas invariants, on en prend un, par exemple 2, on regarde son image, l'image de cette image, etc., jusqu'à ce qu'on revienne à 2, ce qui donne le cycle $c' = (2 \ 9 \ 10 \ 4)$. On recommence une nouvelle fois et on trouve la transposition $c'' = (7 \ 8)$. Enfin, 6 et 11 sont fixes. On écrit $\sigma = (1 \ 3 \ 5)(2 \ 9 \ 10 \ 4)(7 \ 8)$.

Finalement, on a $\sigma = c \circ c' \circ c''$, où les trois cycles, qui ne font pas intervenir les mêmes chiffres, commutent entre eux.

Cette décomposition permet de calculer l'ordre de σ qui est le ppcm des ordres des cycles intervenant dans la décomposition ; pour la permutation σ ci-dessus, c'est $\text{ppcm}(3, 4, 2) = 12$ (l'ordre d'un élément de S_n peut donc être supérieur à n).

Théorème. *Le groupe S_n est engendré par l'ensemble des transpositions.*

Démonstration. Raisonnons par récurrence sur n . Pour $n = 1$ ou $n = 2$, le résultat est vrai.

Soit σ une permutation de S_{n+1} . Si $\sigma(n+1) = n+1$, la restriction de σ à l'ensemble $\{1, \dots, n\}$ est une permutation σ' de cet ensemble. Par hypothèse de récurrence, σ' est un produit de transpositions de l'ensemble $\{1, \dots, n\}$. Ces transpositions définissent des transpositions de l'ensemble $\{1, \dots, n+1\}$ dont le produit est σ .

Si $\sigma(n+1) \neq n+1$, notons τ la transposition $((n+1) \ \sigma(n+1))$. La permutation $\tau \circ \sigma$ laisse $n+1$ invariant ; on peut donc lui appliquer le résultat précédent ; $\tau \circ \sigma$ est donc égal à un produit de transpositions et, en multipliant les deux membres par τ , on voit qu'il en est de même de σ . $\square$

Décomposition en produit de transpositions. Nous venons de montrer que toute permutation est un produit de transpositions et la démonstration précédente donne une méthode pour obtenir une telle décomposition.

Voici une seconde méthode souvent très commode. On commence par écrire la permutation comme produit de cycles disjoints. Il reste alors à décomposer chaque cycle en produit de transpositions.

On part de la formule :

$$(x_1 \dots x_r)(x_r \ x_1) = (x_2 \dots x_r)$$

qui se vérifie facilement en remarquant que le membre de gauche est un composé dans lequel $x_1 \mapsto x_r \mapsto x_1, x_2 \mapsto x_3, \dots, x_r \mapsto x_1 \mapsto x_2$, ce qui montre qu'il est égal au cycle $(x_2 \dots x_r)$. On montre alors, par récurrence sur r , que :

$$(x_1 \dots x_r) = (x_r x_{r-1}) \dots (x_r x_2)(x_r x_1).$$

11.7 SIGNATURE D'UNE PERMUTATION

L'écriture d'une permutation comme produit de transpositions a une propriété importante et difficile à démontrer.

Théorème. Si σ est une permutation de S_n qui s'écrit comme produit de transpositions de deux façons différentes : $\sigma = \tau_1 \circ \dots \circ \tau_r = \tau'_1 \circ \dots \circ \tau'_{r'}$, alors r et r' sont de même parité.

Pour montrer ce théorème, nous allons d'abord étudier la fonction suivante.

Définition 1 : signature d'une permutation. On définit sur S_n une fonction, notée ε et appelée signature, par la formule :

$$\varepsilon(\sigma) = \prod_{(i,j) \in E} \frac{\sigma(j) - \sigma(i)}{j - i}$$

où $E = \{(i,j) \mid 1 \leq i < j \leq n\}$.

Par exemple, pour le cycle $c = (1 \ 3 \ 2)$ de S_3 , on a :

$$\varepsilon(c) = \frac{(\sigma(2) - \sigma(1))(\sigma(3) - \sigma(1))(\sigma(3) - \sigma(2))}{(2-1)(3-1)(3-2)} = 1$$

Signature d'une transposition. Si τ est une transposition de S_n , on a $\varepsilon(\tau) = -1$.

En effet, posons $\tau = (k \ l)$, $1 \leq k < l \leq n$. Si $\{i, j\} \cap \{k, l\} = \emptyset$, on a $\frac{\tau(j) - \tau(i)}{j - i} = 1$. Si $i < k$, on a $\frac{\tau(k) - \tau(i)}{k - i} \frac{\tau(l) - \tau(i)}{l - i} = 1$; de même si $k < i < l$ ou si $l < i$. Enfin $\frac{\tau(l) - \tau(k)}{l - k} = -1$. On en déduit $\varepsilon(\tau) = -1$.

Proposition : propriétés de la signature. L'application ε définie par la formule précédente est un homomorphisme $\varepsilon : S_n \rightarrow \{-1, 1\}$ du groupe S_n dans le groupe multiplicatif $\{-1, 1\}$.

Démonstration. La démonstration est délicate et peut être passée. Soit $\sigma \in S_n$.

Montrons d'abord que $\varepsilon(\sigma) \in \{-1, 1\}$. Ce n'est pas évident sur la formule ; on va voir que les facteurs du numérateur et ceux du dénominateur sont les mêmes au signe près et à une permutation près.

Pour cela, définissons la fonction $f: E \rightarrow E$ par :

- $f(i, j) = (\sigma(i), \sigma(j))$ si $\sigma(i) < \sigma(j)$;
- $f(i, j) = (\sigma(j), \sigma(i))$ si $\sigma(j) < \sigma(i)$.

On vérifie que f est injective ; comme E est fini, f est donc bijective. Par conséquent, en réordonnant le numérateur de $\varepsilon(\sigma)$, on voit que :

$$\left| \prod_{(i,j) \in E} (\sigma(j) - \sigma(i)) \right| = \left| \prod_{(i,j) \in E} (j - i) \right|$$

d'où $|\varepsilon(\sigma)| = 1$.

Montrons maintenant que ε est un homomorphisme de groupes. Soient σ et σ' dans S_n . On a :

$$\varepsilon(\sigma'\sigma) = \prod_{(i,j) \in E} \frac{(\sigma'\sigma)(j) - (\sigma'\sigma)(i)}{j - i}$$

Cette expression peut s'écrire :

$$\prod_{(i,j) \in E} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{(i,j) \in E} \frac{\sigma(j) - \sigma(i)}{j - i}$$

Le second facteur est $\varepsilon(\sigma)$; le premier est $\varepsilon(\sigma')$ puisque, quand les couples (i, j) décrivent E , il en est de même des couples $f(\sigma(j), \sigma(i))$ et que :

$$\begin{aligned} & \prod_{(i,j) \in E} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{(i,j) \in E_1} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{(i,j) \in E_2} \frac{\sigma'(\sigma(i)) - \sigma'(\sigma(j))}{\sigma(i) - \sigma(j)} \end{aligned}$$

avec $E_1 = \{(i, j) \in E \mid \sigma(i) < \sigma(j)\}$ et $E_2 = \{(i, j) \in E \mid \sigma(j) < \sigma(i)\}$. □

Fin de la démonstration du théorème. Si σ est un produit de r transpositions, on a $\varepsilon(\sigma) = (-1)^r$ puisque la signature d'une transposition est -1 . Si σ est aussi un produit de r' transpositions, on a donc $(-1)^r = (-1)^{r'}$, ce qui prouve que les parités de r et de r' sont les mêmes. □

Définition 2 : permutations paires et impaires. On dira qu'une permutation produit d'un nombre pair de transpositions est une permutation paire et qu'une permutation produit d'un nombre impair de transpositions est une permutation impaire.

Si σ est un r -cycle, on a vu ci-dessus qu'il s'écrit comme produit de $r - 1$ transpositions. On a donc $\varepsilon(c) = (-1)^{r-1}$.

Un cycle de longueur paire est donc une permutation impaire et un cycle de longueur impaire est donc une permutation paire.

Définition 3 : groupe alterné. Le noyau de l'homomorphisme de signature est un sous-groupe de S_n appelé groupe alterné et noté A_n . Ses éléments sont donc les permutations paires de $\{1, \dots, n\}$.

Ordre du groupe alterné. On suppose $n \geq 2$. Soit τ une transposition de S_n . Considérons l'application $\sigma \mapsto \tau\sigma$ de S_n dans lui-même. Cette application est bijective, transforme une permutation paire en permutation impaire et *vice versa*. Par conséquent l'ordre de A_n est la moitié de l'ordre de S_n , soit $n!/2$. Par exemple, $|A_3| = 3$, $|A_4| = 12$, $|A_5| = 60$.

11.8 GROUPE LINÉAIRE

Groupe $GL(E)$. Soit E un espace vectoriel. Le composé de deux automorphismes de E est un automorphisme de E , l'application identique est l'élément neutre pour la composition des automorphismes, l'inverse d'un automorphisme de E est un automorphisme de E . Par conséquent, les automorphismes de E forment un groupe noté $GL(E)$ pour la composition des applications. Ce groupe est appelé groupe linéaire de E . Si $E = \mathbb{R}^n$, on notera $GL(\mathbb{R}^n)$ ou $GL(n, \mathbb{R})$ ce groupe.

On a vu que tout espace vectoriel de dimension n est isomorphe à l'espace vectoriel $\mathbb{R}^n$ (voir 7.10). On va montrer que le groupe linéaire d'un espace de dimension n est isomorphe à $GL(\mathbb{R}^n)$. Les propriétés du groupe linéaire d'un espace vectoriel de dimension n seront donc celles de $GL(\mathbb{R}^n)$ et ne nécessitent pas d'étude supplémentaire. La proposition suivante énonce ces résultats.

Proposition 1. *Soit E un espace vectoriel de dimension n . L'ensemble des isomorphismes de E , muni de la loi de composition des isomorphismes, est un groupe. Ce groupe est isomorphe au groupe $GL(\mathbb{R}^n)$ des isomorphismes de l'espace vectoriel $\mathbb{R}^n$.*

Démonstration. Nous savons qu'il existe un isomorphisme $\varphi : \mathbb{R}^n \rightarrow E$. On définit alors un isomorphisme $\Phi : GL(E) \rightarrow GL(\mathbb{R}^n)$ en posant :

$$\Phi(f) = \varphi^{-1} \circ f \circ \varphi.$$

La définition de Φ sera plus claire en considérant le diagramme :

$$\begin{array}{ccc} E & \xrightarrow{f} & E \\ \varphi \uparrow & & \downarrow \varphi^{-1} \\ \mathbb{R}^n & \xrightarrow{\Phi(f)} & \mathbb{R}^n \end{array}$$

Quand on demande que ce diagramme soit commutatif, cela signifie que les deux chemins allant de $\mathbb{R}^n$ à $\mathbb{R}^n$, celui par la ligne du bas et celui qui passe par la ligne du haut, donnent le même résultat, ce qui définit $\Phi(f)$.

Pour vérifier que Φ est un homomorphisme de groupes, il suffit de vérifier que $\Phi(g \circ f) = \Phi(g) \circ \Phi(f)$, ce qui résulte de :

$$\varphi^{-1} \circ (g \circ f) \circ \varphi = (\varphi^{-1} \circ g \circ \varphi) \circ (\varphi^{-1} \circ f \circ \varphi). \quad \square$$

Nous allons ramener l'étude du groupe linéaire d'un $\mathbb{R}$ -espace vectoriel de dimension n à celle du groupe des matrices carrées inversibles d'ordre n à coefficients dans $\mathbb{R}$.

Proposition 2. *Les matrices carrées inversibles d'ordre n à coefficients dans $\mathbb{R}$ forment un groupe pour le produit des matrices. Ce groupe est isomorphe à $GL(\mathbb{R}^n)$.*

Démonstration. Munissons $\mathbb{R}^n$ de la base canonique. À toute application linéaire de $GL(\mathbb{R}^n)$, associons sa matrice dans la base canonique. Cette application est bijective et c'est un homomorphisme de groupes. C'est donc un isomorphisme de groupes. $\square$

11.9 CENTRE DU GROUPE LINÉAIRE

Ce paragraphe et le suivant ne sont pas utiles pour lire la suite du livre. Notons $Z(G)$ l'ensemble des éléments d'un groupe G qui commutent avec tous les éléments de G : $Z(G) = \{x \in G \mid \forall y \in G : xy = yx\}$. On appelle $Z(G)$ le centre de G (l'explication du Z vient de l'allemand *Zentrum* = centre).

Proposition. Le centre d'un groupe G est un sous-groupe de G .

Démonstration. Elle est facile et laissée aux lecteurs et lectrices. $\square$

Proposition : centre du groupe linéaire.

- *Le centre du groupe des matrices carrées inversibles d'ordre n est $\{aI \mid a \in \mathbb{R}^*\}$;*
- *les éléments du centre de $GL(E)$ sont les homothéties de rapport non nul.*

Démonstration. Il est clair d'abord que $\{aI \mid a \in \mathbb{R}^*\}$ est contenu dans le centre du groupe des matrices carrées inversibles d'ordre n .

Montrons qu'aucune autre matrice ne peut être dans ce groupe.

Pour cela, nous utiliserons les matrices de transvections $T_{ij}(\alpha) = I + \alpha E_{ij}$ définies en 8.4. Rappelons que la multiplication à gauche d'une matrice carrée A par la matrice de transvection $T_{ij}(\alpha)$ a pour effet de remplacer la i -ième ligne de A par la somme de la i -ième ligne et de la j -ième multipliée par α et que la multiplication à droite de A par $T_{ij}(\alpha)$ a pour effet de remplacer la j -ième colonne de A par la somme de la j -ième colonne et de la i -ième multipliée par α .

Soit $A = (a_{ij})$ une matrice carrée d'ordre n . Supposons qu'il existe i, j , $1 \leq i, j \leq n$ et $i \neq j$, telle que $a_{ij} \neq 0$. Le coefficient d'indice $(1, j)$ de $T_{1i}(1)A$ est $a_{1j} + a_{ij}$; le coefficient d'indice $(1, j)$ de $AT_{1i}(1)$ est a_{1j} . Donc A n'est pas dans le centre.

Supposons maintenant que A soit une matrice diagonale et qu'il existe i, j , $1 \leq i, j \leq n$ et $i \neq j$, telle que $a_{ii} \neq a_{jj}$. Le coefficient d'indice (i, j) de $T_{ij}(1)A$ est a_{jj} ; le coefficient d'indice $(1, j)$ de $AT_{ij}(1)$ est a_{ii} . Donc A n'est pas dans le centre. $\square$

11.10 GÉNÉRATEURS DU GROUPE LINÉAIRE

Fixons un entier $n > 0$ et notons $T = \{T_{ij}(\alpha) \mid \alpha \in \mathbb{R}^*\}$ et D l'ensemble des matrices diagonales inversibles (c'est-à-dire sans terme nul sur la diagonale) de $GL(\mathbb{R}^n)$.

Proposition. L'ensemble $T \cup D$ engendre le groupe $GL(\mathbb{R}^n)$.

Démonstration. Remarquons d'abord qu'une matrice inversible $A = (a_{ij})$ de $GL(\mathbb{R}^n)$ n'a aucune colonne ne contenant que des 0.

La première colonne de A n'est pas nulle. Si $a_{11} \neq 0$, la multiplication de A à gauche par des matrices de transvection de la forme $T_{i1}(\alpha)$ a l'effet d'ajouter à la i -ième ligne α fois la première, et on peut choisir α de façon à ce que le terme d'indice $(i, 1)$ du produit soit nul. On répète cette opération pour tous les i , $2 \leq i \leq n$.

Si $a_{11} = 0$, il existe un i tel que $a_{i1} \neq 0$ et la matrice $T_{1i}(1)A$ a son terme d'indice $(1, 1)$ non nul, ce qui ramène au cas précédent.

On procède de même, en multipliant à gauche par des matrices de T , pour obtenir une matrice dont les termes de la diagonale sont non nuls et tous les autres termes nuls. On obtient donc une égalité de la forme $UA = V$ où U est un produit de matrices de T et $V \in D$. Comme l'inverse d'une matrice de T est une matrice de T , on voit que $A = U^{-1}V$ est un produit d'éléments de $T \cup D$, d'où le résultat. $\square$

On remarquera que la méthode de démonstration de cette proposition est calquée sur la méthode du pivot. On remarquera aussi que $A^{-1} = V^{-1}U$.

➤ Vers le chapitre 12

Après avoir étudié la notion de groupe, nous allons aborder la notion d'anneau, ce qui va nous permettre d'aborder l'étude de l'anneau $\mathbb{Z}$ et de l'arithmétique puis, au chapitre 13, l'étude des polynômes.

Exercices

11.1 Ordre d'un élément

- a) Soit G un groupe d'ordre premier p . Quel est l'ordre d'un élément x de G distinct de l'élément neutre ? Que peut-on en déduire pour $\text{Gr}(x)$?
- b) Soit G un groupe et a un élément d'ordre t de G . On suppose qu'il existe un entier n tel que $a^n = e$. Montrer que t divise n (utiliser la division euclidienne de n par t).
- c) (On pourra utiliser les résultats de 12.12) Soit G un groupe abélien. Soient a et b deux éléments de G d'ordres respectifs s et t premiers entre eux. Quel est l'ordre de ab ?
- d) On pose

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

Montrer que A , B et AB sont dans $\text{GL}(\mathbb{R}^2)$ et déterminer leurs ordres. Le résultat vous surprend-il ?

- e) On considère la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 10 & 5 & 7 & 11 & 9 & 2 & 1 & 12 & 6 & 13 & 3 & 14 & 4 & 8 \end{pmatrix}$$

de S_{14} .

- Décomposer σ en produit de cycles disjoints.
- Calculer σ^{2005} (on modifiera l'exposant suivant l'année où l'on fait ce calcul).

- f) Soit G un groupe fini et a et b dans G . Montrer les propriétés suivantes.

- a et a^{-1} ont même ordre.
- a et bab^{-1} ont même ordre.
- ab et ba ont même ordre.

- g) Déterminer les ordres possibles des éléments de S_4 et, pour chacun, le nombre d'éléments ayant cet ordre. Reprendre la même question pour les groupes S_5 et S_6 (utiliser la décomposition d'une permutation en produit de cycles disjoints).

11.2 Petits exercices théoriques

- a) Soient a et b deux éléments d'un groupe G . Montrer que a et b commutent si et seulement si $aba^{-1} = b$.
- b) Soit G un groupe. Montrer que H est un sous-groupe de G au sens de la définition 11.4.1 si et seulement si H est non vide et, pour tout x et tout y de H , $xy^{-1} \in H$.
- c) Soit G un groupe. Montrer que si A engendre G et si $A \subset B \subset G$, alors B engendre G .
- d) Dans un groupe fini d'ordre pair, montrer qu'il existe un élément, distinct de l'élément neutre, qui est son propre inverse.
- e) Montrer qu'un sous-ensemble non vide H de S_n est un sous-groupe de S_n si et seulement si il est stable par composition.

11.3 Conjugué d'un cycle

- a) Soit σ une permutation de S_n et $c = (x_1 \dots x_r)$ un cycle de longueur r de S_n . Montrer que le composé $\sigma' = \sigma \circ c \circ \sigma^{-1}$ est le cycle $(\sigma(x_1) \dots \sigma(x_r))$.
- b) Application : calculer le produit des trois 3-cycles
 $c_1 \circ c_2 \circ c_3 = (1 \ 2 \ 3)(2 \ 3 \ 4)(1 \ 3 \ 2)$ dans S_4 .

11.4 Transpositions

Dans cet exercice, on pourra utiliser la formule de l'exercice précédent.

- a) Montrer que l'ensemble E_1 des n transpositions $\{(1 \ 2), (1 \ 3), \dots, (1 \ n)\}$ engendre S_n .
- b) Montrer que l'ensemble E_2 des n transpositions $\{(1 \ 2), (2 \ 3), \dots, ((n-1) \ n)\}$ engendre S_n .
- c) Montrer que l'ensemble E_3 formé par la transposition $(1 \ 2)$ et le cycle $c = (2 \ 3 \dots n)$ engendre S_n .
- d) Déterminer le centre de S_n en utilisant des transpositions.

11.5 Sous-groupes de A_4

- a) Montrer que les éléments de A_4 autres que l'élément neutre sont huit 3-cycles et trois produits de deux transpositions définies sur des sous-ensembles disjoints (on appellera ces permutations *doubles transpositions*).
- b) Montrer que les trois doubles transpositions sont conjuguées entre elles dans A_4 .
- c) Montrer qu'un sous-groupe de A_4 contenant les cycles $(1 \ 2 \ 3)$ et $(1 \ 2 \ 4)$ est égal à A_4 . En déduire qu'un sous-groupe de A_4 contenant deux 3-cycles définis sur des sous-ensembles distincts de $\{1, 2, 3, 4\}$ est égal à A_4 .
- d) En utilisant le c), montrer que A_4 ne possède pas de sous-groupe d'ordre 6.

11.6 Le jeu de taquin

Ce jeu se joue seul avec un carré de 16 cases dont une vide et la seule façon de changer la disposition des cases est de faire glisser une case dans la case vide. On suppose que la position de départ est la suivante :

13	14	15	vide
9	10	11	12
5	6	7	8
1	2	3	4

Peut-on trouver une suite de mouvements amenant à la position suivante ?

13	15	14	vide
9	10	11	12
5	6	7	8
1	2	3	4

Solutions

11.1 a) Dans un groupe fini, l'ordre d'un élément divise l'ordre du groupe, donc l'ordre de x divise p . Comme p est premier, c'est 1 ou p . Comme $x \neq e$, l'ordre est p : x engendre G et $\text{Gr}(x) = G$.

b) On a $n = tq + r$ avec $0 \leq r < t$. Donc $e = a^n = a^{tq+r} = a^r$; si $r \neq 0$, t n'est pas le plus petit des entiers $k > 0$ tels que $a^k = e$, donc $r = 0$ et t divise n .

c) D'une part, $(ab)^{st} = a^{st} b^{st} = (a^s)^t (b^t)^s = e$, donc l'ordre de ab est un diviseur u de st . D'autre part, $(ab)^u = e$ implique $(ab)^{us} = b^{us} = e$ donc t divise us ; comme t est premier avec s , il divise u ; de même s divise u ; donc st divise u ; finalement, $st = u$; l'ordre de ab est st .

d) On trouve que A, A^2, A^3 sont différents de I et que $A^4 = I$, ce qui prouve que A est d'ordre 4 ; de même B est d'ordre 3. On trouve $(AB)^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, donc AB est

d'ordre infini. Si A et B commuteraient, les puissances de AB auraient une périodicité de 12, mais ce n'est pas le cas ici.

e) On trouve $\sigma = (1\ 10\ 13\ 4\ 11\ 3\ 7)(2\ 5\ 9\ 6)(8\ 12\ 14)$; σ est produit de trois cycles disjoints c, c', c'' qui commutent entre eux et sont d'ordre 3, 4, 7. Ces nombres étant premiers entre eux, l'ordre de σ est $3 \times 4 \times 7 = 84$. On a $\sigma^{2005} = c^{2005}c'^{2005}c''^{2005}$. Comme $2005 = 7 \times 286 + 3 = 4 \times 501 + 1 = 3 \times 668 + 1$, on a $\sigma^{2005} = c^3c'c'' = (1\ 4\ 7\ 13\ 3\ 10\ 11)(2\ 5\ 9\ 6)(8\ 12\ 14)$.

f) Prendre la définition de l'ordre comme le plus petit des entiers $k > 0$ tels que $x^k = e$ pour les deux premières questions. Pour la troisième, appliquer la seconde propriété à ab et $b(ab)b^{-1} = ba$.

g) La décomposition d'une permutation en produit de cycles disjoints montre, dans S_6 par exemple, que les éléments d'ordre 3 sont des 3-cycles ou des produits de deux 3-cycles disjoints, que les éléments d'ordre 4 sont des 4-cycles ou des produits d'un 4-cycle par une transposition des deux éléments restants, que les éléments d'ordre 6 sont des 6-cycles ou des produits d'un 3-cycle et d'une des trois transpositions entre deux des trois éléments restants. Dans le tableau suivant, les sommes détaillent les nombres d'éléments correspondants aux différentes décompositions possibles. Les sommes des lignes sont respectivement $4! = 24$, $5! = 120$, $6! = 720$.

	1	2	3	4	5	6
S_4	1	6+3	8	6		
S_5	1	10+15	20	30	24	20
S_6	1	15+45+15	40+40	90+90	144	120+120

11.2 a) L'égalité $ab = ba$ équivaut à $aba^{-1} = b$ (en multipliant à droite par a^{-1}).

b) La définition 11.4.1 implique la propriété. Réciproquement, si H est non vide, soit a dans H . On a d'abord $e \in H$ car $e = aa^{-1}$, puis, pour tout x dans H , on a $x^{-1} \in H$ car $x^{-1} = ex^{-1}$. Enfin, pour tous x et y de H , on a $y^{-1} \in H$ donc $xy = x(y^{-1})^{-1}$ est dans H .

c) $\text{Gr}(B)$ contient $\text{Gr}(A) = G$, donc est nécessairement égal à G .

d) Considérons les ensembles $\{x, x^{-1}\}$ quand x décrit G . Ils forment une partition de G . S'ils avaient tous deux éléments distincts pour $x \neq e$, le nombre d'éléments du groupe serait impair.

e) Soit σ un élément de H ; comme σ est d'ordre fini, l'élément neutre est une puissance de σ , donc appartient à H . Si σ est d'ordre t , son inverse est σ^{t-1} . Les mathématiciens du XIX^e siècle qui travaillaient sur des groupes de permutations n'avaient donc pas besoin de la définition générale de groupe donnée maintenant. On a une propriété analogue pour les sous-groupes d'un groupe fini quelconque.

11.3 a) Si $x \notin \{\sigma(x_1), \dots, \sigma(x_r)\}$, alors $\sigma^{-1}(x) \notin \{x_1, \dots, x_r\}$, donc $\sigma'(x) = \sigma(\sigma^{-1}(x)) = x$. Si $x = \sigma(x_i)$ avec $1 \leq i < r$, alors $\sigma'(x) = \sigma(x_{i+1})$ et si $x = \sigma(x_r)$, alors $\sigma'(x) = \sigma(x_1)$.

On trouve donc :

$$\sigma \circ (x_1 \dots x_r) \circ \sigma^{-1} = (\sigma(x_1) \dots \sigma(x_r))$$

Cette formule est extrêmement utile.

b) Comme $c_3 = c_1^{-1}$, la formule précédente donne :

$$c_1 \circ c_2 \circ c_3 = c_1 \circ c_2 \circ c_1^{-1} = (c_1(2)c_1(3)c_1(4)) = (314).$$

11.4 On sait que l'ensemble des transpositions engendre S_n . Si on réussit à fabriquer toutes les transpositions à partir de l'un des deux ensembles proposés, on pourra donc fabriquer n'importe quelle permutation. Nous utilisons la formule de l'exercice précédent en remarquant que l'inverse d'une transposition est cette transposition elle-même.

a) $(ij) = (1i)(1j)(1i)$; toute transposition est donc produit de transpositions de E_1 , donc E_1 engendre S_n .

b) $(13) = (23)(12)(23)$, $(14) = (34)(13)(34) = (34)(23)(12)(23)(34)$, etc. Comme tout élément de E_1 est produit d'éléments de E_2 et que E_1 engendre S_n , E_2 engendre S_n .

c) $c^i(12)c^{-i} = (1i+2)$ pour $1 \leq i \leq n-2$; comme tout élément de E_1 est produit d'éléments de E_3 et que E_1 engendre S_n , E_3 engendre S_n .

d) Pour $n = 1$ et $n = 2$ le groupe S_n est commutatif et égal à son centre. Supposons $n > 2$ et soit $\sigma \in S_n$. La commutation de σ avec la transposition (ij) s'écrit $\sigma(ij)\sigma^{-1} = (ij)$; on doit donc avoir $(\sigma(i) \sigma(j)) = (ij)$ pour tout i et tout j , ce qui impose $\sigma = \text{id}$.

11.5 a) Les éléments de A_4 autres que e sont :

- les 3-cycles : (123) , (132) , (124) , (142) , (134) , (143) , (234) , (243) ;
- les doubles transpositions $(12)(34)$, $(13)(24)$, $(14)(23)$.

b) On a par exemple $(123)(12)(34)(132) = (23)(14) = (14)(23)$ (faire un calcul de composé de permutations ou appliquer l'exercice 11.3 avec $\sigma = (123)$).

c) Avec (123) , (124) , on fabrique facilement 7 éléments distincts. Comme l'ordre du sous-groupe engendré par ces 3-cycles divise 12, ce ne peut-être que 12. Le cas général peut se ramener à celui-là en renumérotant les éléments.

d) Supposons que H soit un sous-groupe d'ordre 6. Il doit contenir un 3-cycle mais ne peut en contenir deux définis sur des ensembles distincts d'après la question précédente. Il contient donc une double transposition. On peut toujours se ramener au

cas où le 3-cycle est $c = (1\ 2\ 3)$ et où la double transposition est $t = (1\ 2)(3\ 4)$. Le sous-groupe contient alors le 3-cycle $tct^{-1} = (2\ 1\ 4)$ et le c donne une contradiction.

11.6 Notons 16 la case vide. Chaque coup est une transposition de S_{16} . Comme la position recherchée est donnée par la transposition $(14\ 15)$, qui est une permutation impaire, on doit donc faire un nombre impair de coups. Mais la case 16 ne peut retourner à sa position initiale en un nombre impair de coups car elle participe à tous les mouvements et la somme de ses deux coordonnées varie de 1 à chaque coup.

ARITHMÉTIQUE, ANNEAUX

12

12.1 INTRODUCTION

L'étude des nombres remonte aux premiers temps des mathématiques. Les propriétés des entiers sont explorées depuis les Babyloniens vers – 1800. Les *Éléments* d'Euclide en – 300 contiennent plusieurs chapitres d'arithmétique, avec en particulier le célèbre algorithme de recherche du pgcd.

12.2 DIVISION EUCLIDIENNE DANS $\mathbb{Z}$

Proposition. Soient a et b dans $\mathbb{Z}$, $b > 0$. Il existe un couple unique d'entiers q et r tels que : $a = bq + r$ et $0 \leq r < b$; q est appelé le quotient et r le reste de la division euclidienne de a par b .

Démonstration. Notons q la partie entière de $\frac{a}{b}$. On $q \leq \frac{a}{b} < q+1$ donc $bq \leq a < bq + b$; $r = a - bq$ convient. Si deux couples (q, r) , (q', r') satisfont les mêmes conditions, on a $b(q - q') = r' - r$; comme $|r' - r| < b$, on a nécessairement $q = q'$ et $r = r'$. $\square$

Par exemple, en divisant par 5, on a : $17 = 5 \times 3 + 2$, $-17 = 5 \times (-4) + 3$.

Si $b < 0$, la proposition montre, en divisant par $-b$, qu'il existe q et r uniques tels que $a = bq + r$ et $0 \leq r < |b|$. Le nom donné à cette division vient d'Euclide qui utilise cette division au livre 7 de ses *Éléments*.

12.3 CONGRUENCE MODULO n , DÉFINITION DE $\frac{\mathbb{Z}}{n\mathbb{Z}}$

C'est Gauss qui a donné, à la première page de son livre *Disquisitiones Arithmeticae* (en 1801, il a alors 24 ans), les propriétés des nombres congrus modulo n . Une traduction française (*Recherches arithmétiques*) de A.-C.-M. Pouillet-Delisle paraît dès 1807 chez Courcier.

RECHERCHES ARITHMÉTIQUES.

SECTION PREMIÈRE.

Des Nombres congrus en général.

1. Si un nombre a divise la différence des nombres b et c , b et c sont dits *congrus* suivant a , sinon *incongrus*. a s'appellera le module ; chacun des nombres b et c , *résidus* de l'autre dans le premier cas, et *non résidus* dans le second.

Les nombres peuvent être positifs ou négatifs, mais entiers. Quant au module il doit évidemment être pris absolument, c'est-à-dire, sans aucun signe.

Ainsi -9 et $+16$ sont *congrus* par rapport au module 5 ; -7 est *résidu* de 15 par rapport au module 11, et *non résidu* par rapport au module 3.

Au reste 0 étant divisible par tous les nombres, il s'ensuit qu'on peut regarder tout nombre comme congru avec lui-même par rapport à un module quelconque.

2. Tous les résidus d'un nombre donné a suivant le module m ; sont compris dans la formule $a + km$, k étant un entier indéterminé. Les plus faciles des propositions que nous allons exposer

A

Carl Friedrich GAUSS,
Recherches arithmétiques.

Reprint par les Édition Jacques Gabay, Paris, 1989

Définition : congruence modulo n . Soit n un entier naturel.

Si x et y sont dans $\mathbb{Z}$, on dit que x est congru à y modulo n si $x - y$ est divisible par n , c'est-à-dire si x est de la forme $y + kn$ avec k entier. On écrit $x \equiv y \pmod{n}$ (notation de Gauss) ou $x = y \pmod{n}$.

Il est facile de vérifier que cette relation est une relation réflexive, symétrique et transitive ; c'est donc une relation d'équivalence.

Rappel sur les relations d'équivalence. On sait qu'une relation d'équivalence $\mathcal{R}$ sur un ensemble E définit une partition de cet ensemble : les éléments de cette partition sont les sous-ensembles de E définis par les classes d'équivalence modulo $\mathcal{R}$. On appelle quotient de l'ensemble E par la relation $\mathcal{R}$ l'ensemble de ces classes. Chaque classe doit être pensée comme un élément de l'ensemble quotient ; ce point de vue moderne est souligné par Gauss.

Si on note $E/\mathcal{R}$ l'ensemble quotient, l'application $\pi : E \rightarrow E/\mathcal{R}$, appelée application ou projection canonique, qui associe à chaque x de E sa classe modulo $\mathcal{R}$ est essentielle pour faire le lien entre l'ensemble E et son quotient. Elle est surjective. Pour parler commodément des classes d'équivalence, éléments de $E/\mathcal{R}$, il est commode de choisir un élément de la classe pour la représenter, un *représentant* de la classe d'équivalence : si $\alpha \in E/\mathcal{R}$, on choisit un élément x de E tel que $\pi(x) = \alpha$ pour pouvoir parler de α comme de la classe de l'élément x de E ; tout élément équivalent à x définit aussi α , mais on s'efforce, si possible, de bien choisir les représentants.

L'ensemble quotient de $\mathbb{Z}$ par la relation de congruence modulo n sera noté $\frac{\mathbb{Z}}{n\mathbb{Z}}$ et ses éléments seront appelés entiers modulo n .

Par exemple, il y a deux classes d'entiers modulo 2 : les nombres pairs et les nombres impairs ; par conséquent, $\frac{\mathbb{Z}}{2\mathbb{Z}}$ a deux éléments : la classe de 0 et la classe de 1.

Pour $n > 0$, la division euclidienne de x par n s'écrit $x = qn + r$ avec $0 \leq r \leq n - 1$; tout élément x de $\mathbb{Z}$ est donc congru à un entier de l'ensemble $\{k, 0 \leq k \leq n - 1\}$. De plus, deux éléments distincts de $\{k, 0 \leq k \leq n - 1\}$ ne sont pas congrus entre eux, donc $\frac{\mathbb{Z}}{n\mathbb{Z}}$ a n éléments. Les éléments congrus à 0 sont les multiples de n . On conçoit qu'il soit plus agréable, en général, de désigner un élément de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ par son représentant entre 0 et $n - 1$; par exemple, les éléments de $\frac{\mathbb{Z}}{17\mathbb{Z}}$ sont mieux représentés par 0, 1, ..., 16 que par 51, 52, ..., 67.

Pour $n = 0$, la relation d'équivalence est la relation d'égalité et le quotient est $\mathbb{Z}$.

Problèmes de notation. Soit x un entier de $\mathbb{Z}$. On notera $\bar{x}$ ou $x \pmod{n}$ ou simplement x la classe de x dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Dans $\frac{\mathbb{Z}}{5\mathbb{Z}}$, on écrira 2 mod 5 ou même 2 pour noter la classe de 2 modulo 5 ; par exemple, on écrira $2 = 7$ dans $\frac{\mathbb{Z}}{5\mathbb{Z}}$ mais en pensant bien

qu'il s'agit d'éléments de $\frac{\mathbb{Z}}{5\mathbb{Z}}$ et que cet abus de langage ne signifie nullement que $\frac{\mathbb{Z}}{5\mathbb{Z}}$ soit inclus dans $\mathbb{Z}$. Le contexte doit toujours permettre de lever l'ambiguïté. Comme ces abus de langage risquent d'être gênants en première lecture, nous utiliserons les barres dans les paragraphes suivants.

Le lien entre $\mathbb{Z}$ et $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est, nous venons de le dire dans le cas général, mais il nous semble utile d'insister, l'application surjective $\pi : \mathbb{Z} \rightarrow \frac{\mathbb{Z}}{n\mathbb{Z}}$ définie par $\pi(x) = \bar{x}$, la classe de x modulo n . Elle n'est bien sûr pas injective, sauf si $n = 0$.

12.4 ADDITION ET MULTIPLICATION DANS $\frac{\mathbb{Z}}{n\mathbb{Z}}$

L'ensemble des entiers $\mathbb{Z}$ est muni de deux opérations : l'addition et la multiplication. Cela va nous permettre de munir de manière naturelle $\frac{\mathbb{Z}}{n\mathbb{Z}}$ d'une addition et d'une multiplication.

La définition de l'addition dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ part de la remarque suivante : si $x = x' \bmod n$ et $y = y' \bmod n$, on a $x + y = x' + y' \bmod n$. La somme modulo n de deux éléments ne dépend donc que de la valeur de chacun modulo n .

Ceci permet de définir une loi dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$, encore notée $+$ de la façon suivante. Soient a et b deux éléments de $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Comme π est surjectif, il existe x et y dans $\mathbb{Z}$ tels que $\pi(x) = \bar{x} = a$ et $\pi(y) = \bar{y} = b$. On pose $a + b = \pi(x + y)$ et cette définition ne dépend pas des choix de x et de y . Autrement dit :

$$\overline{x + y} = \bar{x} + \bar{y}$$

La loi ainsi définie est une loi de groupe. La vérification est immédiate ; par exemple, l'associativité se démontre en remarquant que

$$\begin{aligned} & [\pi(x) + \pi(y)] + \pi(z) \\ &= \pi(x + y) + \pi(z) = \pi[(x + y) + z] = \pi[x + (y + z)] = \pi(x) + \pi(y + z) \\ &= \pi(x) + [\pi(y) + \pi(z)]. \end{aligned}$$

L'élément neutre est $\pi(0) = \bar{0}$, l'élément opposé à $\pi(x)$ est $\pi(-x) = \pi(n - x)$, autrement dit : $-\bar{x} = \overline{-x}$. Par exemple, dans $\frac{\mathbb{Z}}{5\mathbb{Z}}$, l'opposé de $\bar{3}$ est $\bar{2}$ car $\bar{2} + \bar{3} = \overline{2 + 3} = \bar{0}$.

Par conséquent, π est un homomorphisme de groupes.

De même, on définit une loi de multiplication, notée par juxtaposition, dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ en partant de la remarque suivante : si $x = x' \bmod n$ et $y = y' \bmod n$, on a $x = x' + kn$, $y = y' + k'n$ d'où $xy = x'y' \bmod n$. Alors, si $a = \pi(x)$ et $b = \pi(y)$ sont deux éléments de $\frac{\mathbb{Z}}{n\mathbb{Z}}$, on pose $ab = \pi(xy)$, autrement dit : $\overline{xy} = \overline{x} \overline{y}$.

Par exemple, dans $\frac{\mathbb{Z}}{34\mathbb{Z}}$, on a : $\overline{5 \times 7} = \overline{1}$, $\overline{3 \times 13} = \overline{5}$; on a aussi $\overline{6 \times 17} = \overline{0}$, ce qui montre que le produit de deux éléments non nuls de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ peut être nul.

12.5 STRUCTURES D'ANNEAU COMMUTATIF UNITAIRE ET DE CORPS

Les ensembles $\mathbb{Z}$ et $\frac{\mathbb{Z}}{n\mathbb{Z}}$ sont munis de deux lois. Avec la seule loi d'addition, ils ont une structure de groupe. Avec les deux lois, ils ont une structure plus riche qu'on appelle structure d'anneau commutatif unitaire. Les premiers exemples d'anneaux que nous allons rencontrer dans ce livre sont l'anneau $\mathbb{Z}$ des entiers, l'anneau $\frac{\mathbb{Z}}{n\mathbb{Z}}$ des entiers modulo n , et dans le chapitre suivant, l'anneau $K[X]$ des polynômes à coefficients dans un corps K . La notion sera développée dans le chapitre 19.

Les notions d'anneau et d'idéal ont été introduites en 1871 par Dedekind pour clarifier la théorie des nombres en même temps que la notion de corps (voir 5.5). Le nom d'anneau (*Ring* en allemand) a été donné en 1897 par Hilbert.

Définition 1 : anneau commutatif unitaire. Un anneau commutatif unitaire (nous dirons anneau le plus souvent) est un triplet $(A, +, \cdot)$, qu'on notera souvent A , en sous-entendant les opérations, formé d'un ensemble et de deux lois de composition : l'une appelée addition : $+: A \times A \rightarrow A$, l'autre appelée multiplication $\cdot: A \times A \rightarrow A$. La somme de deux éléments x, y de A est notée $x + y$ et leur produit $x \cdot y$ ou le plus souvent xy . Ces opérations vérifient les propriétés suivantes.

Propriétés de l'addition

➤ Associativité :

$$\text{pour tous } x, y, z \text{ de } A, \text{ on a : } x + (y + z) = (x + y) + z$$

➤ Commutativité :

$$\text{pour tous } x, y \text{ de } A, \text{ on a : } x + y = y + x$$

➤ Existence d'un élément neutre, noté 0 :

$$\text{pour tout } x \text{ de } A : x + 0 = x$$

Chapitre 12 • Arithmétique, anneaux

- Existence d'un élément opposé :

pour tout x de A , il existe y tel que : $x + y = 0$

Cet élément est noté $-x$.

Autrement dit, $(A,+)$ est un groupe commutatif.

Propriétés de la multiplication

- Associativité :

pour tous x, y, z de A , on a : $x(yz) = (xy)z$

- Commutativité :

pour tous x, y de A , on a : $xy = yx$

- Existence d'un élément neutre, noté 1 :

pour tout x de A on a $x1 = x$

Propriété de distributivité :

pour tous x, y, z de A , on a : $x(y + z) = xy + xz$

Un anneau nul est un anneau ayant un seul élément, jouant à la fois le rôle du 1 et du 0 .

Il est facile (mais un peu laborieux) de vérifier que $\mathbb{Z}$ est muni d'une structure d'anneau avec l'addition et la multiplication usuelle, et que $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est muni d'une structure d'anneau avec l'addition et la multiplication que nous avons défini en 12.4.

Les calculs dans un anneau commutatif unitaire s'effectuent comme on en a l'habitude dans les exemples donnés au début de cette section ; par exemple, si a, b, c, d sont quatre éléments d'un anneau A , on a $0.a = 0$, $(-a)b = a(-b) = -ab$, $(a + b)(c + d) = ac + ad + bc + bd$; si on se donne un entier $n \geq 1$, on démontre par récurrence la formule du binôme : $(a + b)^n = \sum_{0 \leq k \leq n} C_n^k a^k b^{n-k}$.

Un élément a d'un anneau A est dit nilpotent s'il existe une puissance de a qui est nulle, autrement dit, s'il existe un entier $n \geq 0$ tel que $a^n = 0$. Par exemple, 4 est un élément nilpotent de l'anneau $\mathbb{Z}/32\mathbb{Z}$ puisque $4^3 = 64 = 0 \bmod 32$.

Définition 2 : corps commutatif. Un corps (on ne précisera pas *commutatif*, en général) est un triplet $(K, +, \cdot)$, qu'on notera K , sans préciser les opérations, formé d'un ensemble et de deux lois de composition $K \times K \rightarrow K$, l'une appelée addition notée $+$, l'autre appelée multiplication notée avec un. ou par juxtaposition. Ce triplet est un anneau commutatif unitaire non nul tel que tout élément non nul possède un inverse pour la multiplication, autrement dit : pour tout $x \neq 0$ de K , il existe y tel que : $xy = 1$; cet élément est noté x^{-1} .

On sait que $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont des corps. On peut noter qu'il s'en faut de peu que $\mathbb{Z}$ possède toutes ces propriétés : la seule qui fasse défaut est l'existence d'inverse de tout entier : seuls 1 et -1 en ont.

D'autres exemples de corps ont pris une place importante dans les mathématiques actuelles : les corps finis, dont les premiers exemples sont donnés par Galois. Par exemple, les corps $\frac{\mathbb{Z}}{p\mathbb{Z}}$ pour p premier. Indiquons qu'on peut construire facilement les corps finis en quotientant des anneaux de polynômes (voir chapitre 21). Les corps finis ayant un nombre d'éléments égal à une puissance de 2 sont très utilisés dans les problèmes d'informatique, de transport de l'information et de cryptographie.

Anneau non commutatif unitaire. Deux exemples que nous avons rencontré aux chapitres 7 et 8 ne rentrent pas exactement dans ce cadre. Il s'agit de l'anneau $L(E)$ des endomorphismes d'un K -espace vectoriel E et de l'anneau $M_n(K)$ des matrices carrées d'ordre n fixé à coefficients dans K . Dans ces deux anneaux, la multiplication (la loi de composition des endomorphismes, la loi de multiplication des matrices) ne vérifient pas une des propriétés de la structure d'anneau : la commutativité.

12.6 HOMOMORPHISMES D'ANNEAUX

Comme toujours quand on définit une structure, il est intéressant de considérer les applications qui respectent cette structure. Pour les espaces vectoriels, nous avons défini les applications linéaires, pour les groupes, ce sont les homomorphismes de groupes, pour les anneaux ce sont les homomorphismes d'anneaux.

Définition : homomorphismes d'anneaux. Soient $(A, +, \cdot)$ et $(A', +', \cdot')$ deux anneaux avec des éléments unités 1 et $1'$. Les homomorphismes d'anneaux de A dans A' sont les applications $f : A \rightarrow A'$ tels que $f(x + y) = f(x) +' f(y)$, $f(xy) = f(x) \cdot' f(y)$ pour tous x, y de A et $f(1) = 1'$.

Un homomorphisme d'anneaux f est, en particulier, un homomorphisme des groupes sous-jacents et $f(0) = 0$.

Le premier exemple d'homomorphisme d'anneaux que nous ayons étudié est l'application $\pi : \mathbb{Z} \rightarrow \frac{\mathbb{Z}}{n\mathbb{Z}}$.

Un isomorphisme d'anneaux commutatifs unitaires $f : A \rightarrow A'$ est un homomorphisme d'anneaux commutatifs unitaires tel qu'il existe un homomorphisme d'anneaux commutatifs unitaires $g : A' \rightarrow A$ tel que $g \circ f = \text{id}(A)$ et $f \circ g = \text{id}(A')$. Il revient au même de demander que f soit homomorphisme d'anneaux commutatifs

unitaires bijectif ; la vérification est analogue à celle faite pour les isomorphismes de groupes en 11.5.

Homomorphismes injectifs. La proposition 1 de 11.5 donne un critère d'injectivité pour les homomorphismes d'anneaux : $f: A \rightarrow A'$ est injectif si et seulement si $\ker(f) = \{0\}$. D'autre part, un homomorphisme de corps $f: K \rightarrow K'$ est toujours injectif, car si $x \in K^*$, $f(x)f(x^{-1}) = f(1) = 1$ montre que $f(x) \neq 0$ donc que $x \notin \ker(f)$.

Proposition : propriété universelle de $\mathbb{Z}$ en tant qu'anneau commutatif unitaire. Soit A un anneau. Il existe un unique homomorphisme d'anneaux commutatifs unitaires $f: \mathbb{Z} \rightarrow A$.

Démonstration. Si on ne considère que les structures de groupes, on a vu qu'il existe un unique homomorphisme de groupes $f: \mathbb{Z} \rightarrow A$ défini par $f(n) = n.1$ pour $n \in \mathbb{Z}$. Il suffit maintenant de vérifier que f est un homomorphisme d'anneaux, c'est-à-dire que $f(mn) = f(m)f(n)$ pour tout m et tout n de $\mathbb{Z}$, ce que nous laissons en exercice. $\square$

12.7 UTILISATIONS DES CONGRUENCES

Les utilisations des congruences sont extrêmement diverses. Donnons un exemple d'applications aux équations diophantiennes, équations polynomiales pour lesquelles on recherche des solutions entières, du nom d'un grand mathématicien de l'école d'Alexandrie, au III^e siècle : Diophante. Existe-t-il des entiers x et y tels que $x^2 - 2y^6 = 17$?

Si de tels entiers existent, leurs images $\bar{x}$ et $\bar{y}$, par l'homomorphisme $\pi: \mathbb{Z} \rightarrow \frac{\mathbb{Z}}{7\mathbb{Z}}$, vérifient $\bar{x}^2 - 2\bar{y}^6 = \bar{3}$ dans $\frac{\mathbb{Z}}{7\mathbb{Z}}$. Dressons les tables des valeurs prises par $\bar{x}^2$ et $\bar{y}^6$ quand $\bar{x}$ et $\bar{y}$ parcourent $\frac{\mathbb{Z}}{7\mathbb{Z}}$:

$\bar{x}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{y}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$
$\bar{x}^2$	$\bar{0}$	$\bar{1}$	$\bar{4}$	$\bar{2}$	$\bar{2}$	$\bar{4}$	$\bar{1}$	$\bar{y}^6$	$\bar{0}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$

On voit que $\bar{x}^2 - 2\bar{y}^6$ peut prendre les valeurs de $\bar{x}^2$ diminuées de $\bar{0}$ ou $\bar{2}$, soit $\bar{0}, \bar{1}, \bar{2}, \bar{4}$ et $\bar{5}, \bar{6}, \bar{0}, \bar{2}$. Comme $\bar{3}$ n'est jamais obtenu, l'équation initiale n'a pas de solutions entières. Dans ce type d'applications, la première difficulté est de trouver le nombre modulo lequel on va raisonner ; il n'est même pas sûr, *a priori*, qu'il en existe.

12.8 ÉLÉMENTS INVERSIBLES

Un élément u d'un anneau A est dit inversible s'il existe un élément v de A tel que $uv = 1$. L'élément v est dit inverse de u et noté u^{-1} .

Par exemple, $\bar{2}$ est inversible dans $\mathbb{Z}/15\mathbb{Z}$ et d'inverse $\bar{8}$ car $\bar{2}.\bar{8} = \bar{16} = \bar{1}$; dans ce même anneau, $\bar{4}$ est son propre inverse. Par contre, $\bar{3}$ n'est pas inversible, sinon on aurait $\bar{5} = \bar{3}.x.\bar{5} = \bar{15}x = \bar{0}$. Plus généralement, un diviseur de 0 dans un anneau, c'est-à-dire un élément $x \neq 0$ tel qu'il existe $y \neq 0$ avec $xy = 0$, n'est pas inversible.

Un élément nilpotent est un diviseur de 0 et n'est pas inversible.

Définition : anneau intègre. Un anneau non nul est dit intègre si la relation $ab = 0$ implique $a = 0$ ou $b = 0$.

Dans un anneau intègre, le produit de deux éléments non nuls est non nul. Par conséquent, dans un anneau intègre, une égalité du type $au = av$ avec $a \neq 0$, implique $a(u - v) = 0$ donc $u = v$, autrement dit, on peut simplifier par a .

Par exemple, tout corps est un anneau intègre puisque $ab = 0$ et $a \neq 0$ impliquent $b = 0$; $\mathbb{Z}$ est également un anneau intègre, mais $\frac{\mathbb{Z}}{n\mathbb{Z}}$ n'est pas intègre si n n'est pas premier : si $n = ab$ avec $1 < a, b < n$, alors $\bar{a}\bar{b} = \bar{0}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$, mais $\bar{a}$ et $\bar{b}$ sont tous deux non nuls dans cet anneau.

12.9 IDÉAL

Cette section présente des notions de grande importance pour la théorie des anneaux. Elles sont rapidement présentées ici, dans le but de fournir un langage commode et efficace pour les développements arithmétiques des sections suivantes. Elles seront approfondies au chapitre 19.

Définition 1 : idéal. Un idéal d'un anneau A est un sous-groupe I de A tel que : pour tout x de I et tout a de A , on a $ax \in I$.

L'intersection d'une famille d'idéaux de A est un idéal de A .

Proposition : noyau d'un homomorphisme. Si $f : A \rightarrow A'$ est un homomorphisme d'anneaux commutatifs unitaires, $\ker(f) = f^{-1}(0)$ est un idéal de A . $\square$

L'exemple type de ce chapitre est l'idéal $n\mathbb{Z}$ des multiples de n dans $\mathbb{Z}$ qui est le noyau de l'homomorphisme $\pi : \mathbb{Z} \rightarrow \frac{\mathbb{Z}}{n\mathbb{Z}}$. On notera que A et $\{0\}$ sont des idéaux de A .

Définition 2 : idéal engendré. L'idéal engendré par une partie B d'un anneau A est le plus petit idéal de A contenant B .

Cet idéal existe : c'est l'intersection de la famille, non vide, des idéaux de A contenant B . On le note (B) . On peut décrire l'idéal engendré par B comme l'ensemble des *combinaisons linéaires* à coefficients dans A de familles d'éléments de B , c'est-à-dire comme l'ensemble des sommes de la forme $a_1b_1 + \dots + a_nb_n$ où $n \in \mathbb{N}$, $a_1, \dots, a_n \in A$ et $b_1, \dots, b_n \in B$: en effet, on vérifie facilement que ces sommes forment un idéal et qu'elles appartiennent à tout idéal contenant B .

Définition 3 : idéal principal. Un idéal est dit principal lorsqu'il est engendré par un seul élément ; il est de la forme $\{ax | x \in A\}$ et noté aA ou (a) .

Définition 4 : anneau principal. Un anneau est dit principal lorsqu'il est intègre et que tous ses idéaux sont principaux.

La condition d'intégrité permet d'affirmer que, si a et b engendrent un même idéal principal $I \neq (0)$, il existe un élément inversible u de A tel que $a = ub$. En effet, si $I = (a) = (b)$, les relations $a \in (b)$ et $b \in (a)$ impliquent l'existence de u et v dans A tels que $a = ub$ et $b = va$; on a donc $a = uva$, $(1 - uv)a = 0$, d'où $uv = 1$ car $a \neq 0$.

Comme exemples importants d'anneaux principaux, nous présentons l'anneau $\mathbb{Z}$ dans la section suivante et les anneaux de polynômes $K[X]$ au chapitre 13 ; pour d'autres exemples, voir le chapitre 19.

12.10 SOUS-GROUPES, IDÉAUX DE $\mathbb{Z}$

Déterminons les sous-groupes et les idéaux de $\mathbb{Z}$.

Proposition.

- Tout sous-groupe de $(\mathbb{Z}, +)$ est de la forme $n\mathbb{Z}$, avec $n \geq 0$ unique.
- Tout idéal de $(\mathbb{Z}, +, \cdot)$ est de la forme $n\mathbb{Z}$ avec $n \geq 0$ unique.
- L'anneau $(\mathbb{Z}, +, \cdot)$ est un anneau principal.

Démonstration. L'ensemble $n\mathbb{Z}$ des multiples de n est clairement un sous-groupe et même un idéal de $\mathbb{Z}$.

Réciproquement, soit H un sous-groupe de $\mathbb{Z}$. Soit $H = \{0\} = 0\mathbb{Z}$, soit $H \neq \{0\}$. Dans ce second cas, il possède un élément k non nul. Si $k < 0$, alors $-k$ est un élément > 0 de H , donc H possède toujours un élément > 0 . Notons n le plus petit des éléments > 0 de H . Les multiples de n sont dans H , ce qui montre que $n\mathbb{Z} \subset H$. Soit a un élément de H . Par division euclidienne de a par n , il existe des entiers q et r tels que $a = nq + r$ et $0 \leq r < n$. On a $r = a - nq \in H$ et comme $0 \leq r < n$, la seule possibilité est $r = 0$, ce qui prouve que $a = nq$ donc que $H = n\mathbb{Z}$.

Un idéal de $\mathbb{Z}$ étant un sous-groupe de $\mathbb{Z}$ est donc aussi de la forme $n\mathbb{Z}$. C'est donc un idéal principal. Par conséquent, $\mathbb{Z}$ est un anneau principal. $\square$

12.11 DIVISIBILITÉ, NOMBRES PREMIERS

Définition 1 : divisibilité. Soient a et b dans $\mathbb{Z}$. On dit que b divise a (ou que b est un diviseur de a , ou encore que a est un multiple de b) s'il existe c dans $\mathbb{Z}$ tel que $a = bc$. On note $b|a$.

Comme 1 et -1 sont inversibles dans $\mathbb{Z}$, ils divisent tout entier.

Tout entier divise 0 mais 0 ne divise aucun entier non nul.

Si $a|b$ et si $b|a$, a et b étant non nuls, alors il existe c et c' dans $\mathbb{Z}$ tels que $b = ac$ et $a = bc'$; on a $a(1 - cc') = 0$, donc $cc' = 1$; par conséquent, c est inversible, $c = \pm 1$ et $a = \pm b$.

La relation de divisibilité n'est donc pas une relation d'ordre : elle est réflexive et transitive mais elle n'est pas antisymétrique ; on dit que c'est une relation de pré-ordre.

On peut exprimer la relation de divisibilité en termes d'idéaux. La relation $b|a$ équivaut à l'inclusion $(a) \subset (b)$ puisque :

- si $b|a$, il existe c tel que $a = bc$; l'idéal engendré par b contient a donc $(a) \subset (b)$;
- réciproquement, si $(a) \subset (b)$, $a \in (b)$ donc a est un multiple de b .

Cette reformulation de la relation de divisibilité est importante pour la suite : elle permet de démontrer rapidement les propriétés de la section 12.

Définition 2 : nombres premiers. Les éléments $p \neq 0$ non inversibles de $\mathbb{Z}$ qui admettent pour seuls diviseurs $\pm p$ et ± 1 sont appelés nombres premiers.

En général, on parlera de nombres premiers > 0 . Par exemple, 23, 239, 1 093 sont des nombres premiers, 243 et 343 n'en sont pas. On peut remarquer que 1 n'est pas premier, étant inversible.

Définition 3 : nombres composés. Un nombre n qui n'est ni inversible ni premier est appelé composé.

Un nombre composé possède d'autres diviseurs que ± 1 et $\pm n$; par exemple, $243 = 3 \times 81$, $343 = 7 \times 49$ sont composés.

Proposition 1. *Tout entier est produit de facteurs premiers.*

Démonstration. Par récurrence. $\square$

Proposition 2 : infinité des nombres premiers. *L'ensemble des nombres premiers est infini.*

Démonstration. Raisonnons par l'absurde et supposons que l'ensemble des nombres premiers a un nombre fini d'éléments $p_1, \dots, p_n$. Notons P le produit de tous ces nombres et posons $N = P + 1$. Soit p un diviseur premier de N . D'après l'hypothèse, il existe k , $1 \leq k \leq n$ tel que $p = p_k$. Comme p_k divise P et N , il divise $N - P = 1$, ce qui est absurde. $\square$

Commentaire. L'énoncé que donne Euclide de ce résultat est différent du point de la philosophie des mathématiques. À son époque, on refusait d'envisager des ensembles infinis et Euclide énonce : « Les nombres premiers sont en plus grande quantité que toute quantité proposée de nombres premiers ».

12.12 PGCD, PPCM, NOMBRES PREMIERS ENTRE EUX

12.12.1 Pgcd

Définition 1 : pgcd. Soient a et b dans $\mathbb{Z}$. On dit que d est un plus grand commun diviseur (on abrège en pgcd) de a et b si $d|a$, $d|b$ et si tout diviseur commun de a et b divise d .

Les entiers d et $-d$ sont alors pgcd de a et b ; quand on parle du pgcd de a et b , on pense souvent à celui qui est positif ; on écrit $d = \text{pgcd}(a, b)$ ou parfois $d = (a, b)$. Les Anglais utilisent l'abréviation gcd (pour *greatest common divisor*).

Plus généralement, on définit le pgcd de n entiers $a_1, a_2, \dots, a_n$

Existence du pgcd. Soient a et b deux entiers. L'idéal I engendré par a et b étant un idéal de $\mathbb{Z}$ est un idéal principal. Un générateur d de I est un pgcd de a et b . Pour le vérifier, il suffit d'exprimer la relation de divisibilité en termes d'idéaux.

De même, $\text{pgcd}(a_1, \dots, a_n)$ est un générateur de l'idéal engendré par $a_1, \dots, a_n$. Le calcul du pgcd de n entiers se ramène à des calculs de pgcd de deux entiers grâce à la formule : $\text{pgcd}(a, b, c) = \text{pgcd}(\text{pgcd}(a, b), c)$.

Calcul du pgcd dans $\mathbb{Z}$: algorithme d'Euclide. Soient a et b dans $\mathbb{Z}$. On a $\text{pgcd}(a, b) = \text{pgcd}(b, a)$, $\text{pgcd}(a, 0) = a$, $\text{pgcd}(-a, b) = \text{pgcd}(a, b)$, $\text{pgcd}(a, a) = a$. Il suffit donc de montrer comment calculer le pgcd de deux nombres a et b dans le cas $a > b > 0$.

L'algorithme de calcul du pgcd est donné par Euclide (proposition 2 du livre 7 des Éléments ; Euclide fait des soustractions successives et non des divisions). Il est basé sur la remarque suivante. Quand on écrit la division euclidienne de a par b : $a = bq + r$, un diviseur commun à a et b divise $r = a - bq$; d'autre part, un diviseur commun à b et r divise a . On a donc $\text{pgcd}(a, b) = \text{pgcd}(b, r)$.

Soient donc deux entiers a et b , $a > b > 0$. On définit par récurrence une suite d'entiers r_k en posant : $r_0 = a$, $r_1 = b$, et si $k \geq 1$ et si $r_k \neq 0$, r_{k+1} est le reste de la division euclidienne de r_{k-1} par r_k : $r_{k-1} = r_k q_k + r_{k+1}$, avec $0 \leq r_{k+1} < r_k$; si $r_k = 0$, on arrête ; on note N ce dernier indice. Le pgcd est donc *le dernier reste non nul* : $\text{pgcd}(a, b) = r_{N-1}$. Montrons-le sur un exemple.

Exemple. Posons $a = 2\,652$ et $b = 2\,310$. Les divisions successives donnent :

$$\begin{array}{rcl} 2652 & = & 2310 + 342 \\ 2310 & = & 342 \times 6 + 258 \\ 342 & = & 258 + 84 \\ 258 & = & 84 \times 3 + 6 \\ 84 & = & 6 \times 14 + 0 \end{array}$$

On en déduit : $\text{pgcd}(2\,652, 2\,310) = 6$.

En effet, on a

$$\begin{aligned} \text{pgcd}(2\,652, 2\,310) &= \text{pgcd}(2\,310, 342) = \text{pgcd}(342, 258) = \text{pgcd} \\ (258, 84) &= \text{pgcd}(84, 6) = \text{pgcd}(6, 0) = 6. \end{aligned}$$

Il est facile de généraliser cette démonstration.

Remarque : On pourrait penser à une autre méthode : décomposer a et b en produits de facteurs premiers. Si a et b admettent les décompositions en produit de facteurs premiers $\prod_{1 \leq i \leq m} p_i^{\alpha_i}$ et $\prod_{1 \leq i \leq m} p_i^{\beta_i}$ (certains des exposants peuvent être nuls dans l'une ou l'autre des deux décompositions), on obtient le pgcd de a et de b en prenant le produit des p_i à des puissances $\inf(\alpha_i, \beta_i)$. Par exemple, si $a = 2^4 \times 3^7 \times 11$ et si $b = 2^5 \times 3^2 \times 13$, on a $\text{pgcd}(a, b) = 2^4 \times 3^2$. Cependant, ce procédé est impraticable dès que a et b deviennent un peu grands, car on est incapable, en général, de trouver les décompositions de a et b en facteurs premiers (voir 22.12).

12.12.2 Identité de Bézout

Proposition 1 : identité de Bézout.

- 1) Soient a et b dans $\mathbb{Z}$ et $d = \text{pgcd}(a, b)$. Il existe des entiers u et v tels que $ua + vb = d$.
- 2) Soient $a_1, \dots, a_n$ dans $\mathbb{Z}$ et $d = \text{pgcd}(a_1, \dots, a_n)$. Il existe des entiers $u_1, \dots, u_n$ tels que $u_1 a_1 + \dots + u_n a_n = d$.

Démonstration. 1) L'idéal (d) est l'idéal engendré par a et b donc contient nécessairement l'ensemble E des éléments de la forme $ma + nb$ avec m et n dans $\mathbb{Z}$, d'après 12.9 ; comme E est visiblement un idéal de $\mathbb{Z}$, on a $E = (d)$ donc il existe des entiers u et v tels que $ua + vb = d$.

2) On raisonne comme ci-dessus : l'idéal (d) est l'idéal engendré par $a_1, \dots, a_n$; il contient nécessairement l'ensemble des éléments de la forme $m_1 a_1 + \dots + m_n a_n$, etc. $\square$

Cette identité apparaît dans le livre de Claude Gaspard Bachet de Méziriac (1581-1638), *Problèmes plaisants et délectables qui se font par les nombres*, 1624, 150 ans avant les travaux de Bézout.

La première idée pour trouver les nombres u et v est de *remonter* les calculs faits pour calculer le pgcd. Pour trouver u et v tels que $6 = 2652u + 2310v$, on peut reprendre les égalités ci-dessus en partant de la dernière :

$$\begin{aligned} 6 &= 258 - 84 \times 3 \\ &= 258 - (342 - 258) \times 3 \\ &= 258 \times 4 - 342 \times 3 \\ &= (2310 - 342 \times 6) \times 4 - 342 \times 3 \\ &= 2310 \times 4 - 342 \times 27 \\ &= 2310 \times 4 - (2652 - 2310) \times 27 \\ &= 2652 \times (-27) + 2310 \times 31 \end{aligned}$$

Ainsi, on pourra prendre $u = -27$ et $v = 31$.

Une meilleure méthode est proposée en exercice.

Définition 2 : nombres premiers entre eux

- 1) Deux entiers sont dits premiers entre eux si leur pgcd est ± 1 , c'est-à-dire si l'idéal qu'ils engendrent est $\mathbb{Z}$.
- 2) On dit de même que n entiers $a_1, \dots, a_n$ sont premiers entre eux si leur pgcd est ± 1 .

Proposition 2 : critère pour que des entiers soient premiers entre eux.

- 1) Deux entiers a et b sont premiers entre eux si et seulement s'il existe des entiers u et v tels que $ua + vb = 1$.
- 2) Des entiers $a_1, \dots, a_n$ sont premiers entre eux si et seulement s'il existe des entiers $u_1, \dots, u_n$ tels que $u_1a_1 + \dots + u_na_n = 1$.

Démonstration. 1) Dans un sens, c'est l'identité de Bézout. Réciproquement, si de tels entiers existent, tout diviseur commun à a et b divise 1, donc $\text{pgcd}(a, b) = 1$. $\square$

On ne confondra pas les énoncés « $a_1, \dots, a_n$ sont premiers entre eux » et « $a_1, \dots, a_n$ sont premiers entre eux deux à deux » (considérer, par exemple, les entiers 2, 3 et 6).

12.12.3 Applications de l'identité de Bézout

Inverse d'un entier modulo n . Soit a un entier non premier avec n et non nul modulo n . Posons $d = \text{pgcd}(a, n)$. Il existe donc k et l , $1 < l < n$, tels que $a = kd$ et $n = ld$. On a $la = lkd = kn$, donc $\overline{la} = 0$, ce qui prouve que a n'est pas inversible modulo n .

Soit a un entier premier avec n . L'identité de Bézout montre qu'il exist des entiers u et v tels que $ua + vn = 1$. On a donc : $ua = 1 \pmod n$, ce qui montre que $\bar{u}$ est l'inverse de $\bar{a}$ dans l'anneau $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Pour calculer l'inverse de a , on peut donc calculer le u de l'identité de Bézout.

19. Si les nombres a, b, c , etc. sont premiers avec k , leur produit l'est aussi.

En effet, puisqu'aucun des nombres a, b, c , etc. n'a de facteurs premiers communs avec k , et que le produit de ces nombres ne peut avoir de facteurs premiers qui n'appartiennent à quelqu'un d'entr'eux, ce produit n'aura non plus aucun facteur premier commun avec k .

Si les nombres a, b, c , etc. sont premiers entr'eux, et que k soit divisible par chacun d'eux, il le sera aussi par leur produit.

C'est une suite des n^{os} 17 et 18. Soit en effet p un diviseur premier quelconque du produit abc etc. et qu'il ait l'exposant π , quelqu'un des nombres a, b, c , etc. sera divisible par p^π , par conséquent k , qui est divisible par ce nombre, le sera aussi par p^π : il en sera de même des autres diviseurs du produit.

Donc, si deux nombres m, n sont congrus suivant plusieurs modules a, b, c , etc. premiers entr'eux, ils le seront aussi suivant leur produit. En effet, puisque $m - n$ est divisible par chacun des nombres a, b, c , etc., il le sera aussi par leur produit.

Enfin, si a est premier avec b , et que ak soit divisible par b , k sera aussi divisible par b . En effet, puisque ak est divisible par a et par b , il le sera par leur produit ; donc $\frac{ak}{a} = \frac{k}{b}$ sera un entier.

Carl Friedrich GAUSS,

Recherches arithmétiques.

Reprint par les Éditions Jacques Gabay, Paris, 1989, page 9.

Proposition 3 : lemme de Gauss. Soient a, b, c trois entiers non nuls. Si $a \mid bc$ et si $(a, b) = 1$ alors $a \mid c$.

Démonstration. En effet, $(a, b) = 1$ entraîne l'existence d'entiers u et v tels que $ua + vb = 1$; on a donc $uac + vbc = c$; comme a divise bc , il divise le membre de gauche ; par conséquent, il divise c . $\square$

En particulier, si un nombre premier divise un produit d'entiers, il divise l'un des facteurs.

Voir le texte de Gauss (ci-dessus), page 9 des *Recherches arithmétiques*.

Corollaire. Si deux entiers a et b premiers entre eux divisent un entier c , leur produit divise c .

Démonstration. Si a divise c , on peut écrire $c = ac_1$; comme b divise $c = ac_1$ et est premier avec a , il divise c_1 . On a donc $c_1 = bc_2$ d'où $c = abc_2$. $\square$

Proposition 4 : unicité de la décomposition en produit de facteurs premiers. Soit n un entier non nul. Il existe des nombres premiers $p_1, \dots, p_r$ et des entiers strictement positifs $k_1, \dots, k_r$ uniques tels que :

$$n = u \prod_{1 \leq i \leq r} p_i^{k_i}$$

avec $u = \pm 1$.

Démonstration. L'existence de la décomposition est donnée par la proposition 1 de 12.11. Supposons que $n = p_1^{k_1} \dots p_r^{k_r} = q_1^{l_1} \dots q_s^{l_s}$ où les p_i et les q_j sont des nombres premiers et où les exposants sont des entiers ≥ 0 . Si les deux écritures ne sont pas identiques, on peut simplifier par les facteurs communs. Soit p un facteur premier restant dans le premier membre. Il divise le produit de nombres premiers restant dans le second membre sans diviser aucun d'eux, ce qui contredit le lemme de Gauss. $\square$

Un peu d'histoire. À la page 7 de ses *Recherches arithmétiques*, Gauss explique : *Il est évident par les éléments, que l'on peut toujours décomposer un nombre quelconque en facteurs premiers ; mais on suppose à tort tacitement que cette décomposition ne soit possible que d'une seule manière.* Et Gauss donne une démonstration de cette unicité que personne, avant lui, n'avait perçue comme nécessaire.

12.12.4 ppcm dans $\mathbb{Z}$

On dit que m est un plus petit commun multiple (on abrège en ppcm) de a et b si m est multiple de a et b et si tout multiple commun de a et b est multiple de m ; m et $-m$ sont alors ppcm de a et b ; quand on parle du ppcm de a et b , on pense souvent à celui qui est > 0 ; on écrit $m = \text{ppcm}(a, b)$.

Montrons l'existence de ppcm (a, b) . L'idéal $(a) \cap (b)$ de $\mathbb{Z}$ est principal, engendré par un élément m unique au signe près ; m est un ppcm de a et b comme on le voit en exprimant la relation de divisibilité en termes d'idéaux.

On généralise facilement à un ensemble fini quelconque d'éléments de $\mathbb{Z}$.

Notons que le ppcm et le pgcd sont liés par la formule :

$$ab = \text{ppcm}(a, b) \text{ pgcd}(a, b).$$

Cette formule permet de calculer le ppcm à partir du pgcd, donc à l'aide de divisions euclidiennes. Comme pour le pgcd, on ne peut calculer, en général, le ppcm en cherchant à décomposer les nombres a et b en produits de facteurs premiers ; si on dispose cependant de telles décompositions, le ppcm de a et de b est, avec les notations utilisées dans la remarque finale de 12.12.1, le produit des p_i à des puissances $\sup(\alpha_i, \beta_i)$. Par exemple, si $a = 2^4 \times 3^7 \times 11$ et si $b = 2^5 \times 3^2 \times 13$, on a $\text{ppcm}(a, b) = 2^5 \times 3^7 \times 11 \times 13$.

12.13 LES CORPS $\frac{\mathbb{Z}}{p\mathbb{Z}}$

Proposition 1 : critère pour que $\frac{\mathbb{Z}}{p\mathbb{Z}}$ soit un corps. Soit p un entier. L'anneau $\frac{\mathbb{Z}}{p\mathbb{Z}}$ est un corps si et seulement si p est premier.

Démonstration. Supposons p non premier. Il existe des entiers $a < p$ et $b < p$ tels que $ab = p$. Comme a et b ne sont pas inversibles modulo p , $\frac{\mathbb{Z}}{p\mathbb{Z}}$ n'est pas un corps.

Supposons maintenant p premier. Pour montrer que $\frac{\mathbb{Z}}{p\mathbb{Z}}$ est un corps, il suffit de montrer que tout élément non nul est inversible. Mais si un entier a est premier avec p , il est inversible dans $\frac{\mathbb{Z}}{p\mathbb{Z}}$ d'après 12.12.3. $\square$

Commentaire. Ainsi, voilà de nouveaux corps, finis, sur lesquels construire des espaces vectoriels pour faire de l'algèbre linéaire. Dans les chapitres 5 à 10, nous avons développé la théorie des espaces vectoriels sur le corps des réels. Vous pouvez vérifier que nous n'avons jamais utilisé de propriété particulière du corps des nombres réels. Les résultats que nous avons démontrés alors sont donc valables pour les espaces vectoriels sur un corps quelconque, par exemple sur le corps $\frac{\mathbb{Z}}{p\mathbb{Z}}$ pour p premier.

Proposition 2 : petit théorème de Fermat. Soit p un nombre premier. Dans le corps $K = \frac{\mathbb{Z}}{p\mathbb{Z}}$, on a :

- 1) $x^{p-1} = 1$ pour tout x non nul ;
- 2) $x^p = x$ pour tout x .

Démonstration. La seconde égalité résulte de la première. La première n'est évidemment pas vraie pour $x = 0$. Tout $x \neq 0$ est un élément du groupe multiplicatif $K^* = \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^*$. Ce groupe étant fini d'ordre $p - 1$, le théorème de Lagrange donne le résultat. $\square$

Commentaire. Pierre de Fermat (1601?-1665) (un article de Klaus Barner publié en 2001 donne des arguments convaincants pour placer la naissance de Fermat en 1607) a rarement donné les démonstrations des résultats qu'il énonçait. Ce qu'on appelle son petit théorème est énoncé dans des lettres de 1640. Leibniz et Euler en donneront des démonstrations, 40 et 96 ans plus tard.

Nous présenterons au chapitre 22 les utilisations actuelles et inattendues de l'arithmétique et de la théorie des corps finis.

➤ Vers le chapitre 13

Le chapitre suivant étudie les polynômes. Il s'agit encore d'anneaux et certains résultats sont très semblables à ceux obtenus en arithmétique dans ce chapitre.

Exercices

Les entiers ne sont pas distingués de leur classe modulo un nombre dans les exercices qui suivent ; le contexte permet de lever cette ambiguïté.

12.1 Petits calculs

- 1) Calculer $37 + 55 \bmod 63$, $37 \times 55 \bmod 63$.
- 2) Calculer $\text{pgcd}(433014481, 18000)$ en décomposant 18 000 en produit de facteurs premiers.
- 3) Pour cet exercice, on pourra s'aider d'une petite calculatrice. La seconde question n'est pas faite pour rebuter les lecteurs et lectrices, mais pour comparer les deux méthodes de calcul du pgcd.
 - a) Calculer le pgcd de $a = 42098$ et de $b = 36146$ avec des divisions euclidiennes.
 - b) Retrouver ce résultat en décomposant a et b en produit de facteurs premiers.
 - c) Déterminer des entiers u et v tels que $\text{pgcd}(a, b) = au + bv$.
 - d) En déduire tous les couples (s, t) d'entiers tels que $\text{pgcd}(a, b) = as + bt$.
- 4) Quel est l'inverse de 583 dans $\frac{\mathbb{Z}}{679\mathbb{Z}}$?
- 5) Montrer que, pour tout entier $a \geq 0$, 3^{a+1} divise $2^{3^a} + 1$.

12.2 Critères de divisibilité

- 1) Soit n un entier et $S(n)$ la somme des chiffres $a_0, \dots, a_N$ de l'écriture décimale de $n = a_N \dots a_0 = \sum_{0 \leq k \leq N} a_k 10^k$.
 - a) Comparer $n \bmod 3$ et $S(n) \bmod 3$; en déduire un critère de divisibilité par 3.
 - b) Comparer $n \bmod 9$ et $S(n) \bmod 9$; en déduire un critère de divisibilité par 9.
 - c) Donner un critère de divisibilité par 11 en étudiant les puissances de 10 modulo 11.

2) Soit n un entier. On écrit $n = 10a + b$ en séparant le nombre de dizaines a et le nombre d'unités b de n .

- a) Montrer que n est divisible par 13 équivaut à : $a + 4b$ divisible par 13.
- b) Montrer que n est divisible par 19 équivaut à : $a + 2b$ divisible par 19.
- c) Montrer à l'aide de ces critères que 107 341 est divisible par 13 et que 156 883 est divisible par 119.

12.3 Calcul pratique de u et de v

On reprend les résultats présentés dans le paragraphe 12.12. Soient a et b des entiers, $r_0, r_1, \dots, r_N$ la suite pour déterminer leur pgcd et d leur pgcd. L'idée de remonter les calculs pour déterminer les u et v tels que $d = ua + vb$ suppose que tous les calculs aient été mémorisés. Pour un programme informatique, on peut faire mieux en déterminant à chaque étape des entiers u_k et v_k tels que $r_k = au_k + bv_k$.

- a) Montrer qu'on peut choisir comme suites (u_k) et (v_k) les suites définies par $u_0 = 1, u_1 = 0$ et $u_{k+2} = u_k - u_{k+1}q_{k+1}$, $v_0 = 0, v_1 = 1$ et $v_{k+2} = v_k - v_{k+1}q_{k+1}$.
- b) Quelle est la valeur de l'indice tel que $u_k = u, v_k = v$?
- c) Reprendre les calculs de l'exercice 12.1 avec cet algorithme.

12.4 Théorème de Lamé

Les nombres de Fibonacci F_n ont été présentés en 2.5. On reprend les notations des résultats du paragraphe 12.12. Soient a et b des entiers, $r_0, r_1, \dots, r_N$ la suite pour déterminer leur pgcd d ; on a $d = r_{N-1}$ et $r_N = 0$. On note m le nombre de divisions successives pour obtenir r_N et p le nombre de chiffres de b en écriture décimale. Le théorème de Lamé dit que $m \leq 5p$.

- a) Montrer que $r_{N-k} \geq F_k$ pour $k = 1, \dots, N$.
- b) Montrer que $N \geq \frac{p}{\log(\rho_1)} + \frac{\log(\sqrt{5})}{\log(\rho_1)}$ (la notation $\log$ désigne le logarithme décimal) ; en déduire le théorème de Lamé.
- c) Que pensez-vous de cette majoration si (a, b) est l'un des couples suivants : (F_6, F_5) , (F_{11}, F_{10}) , (F_{16}, F_{15}) ? La majoration est-elle optimale si $p = 4$?

Lamé est connu pour avoir démontré le cas $n = 7$ du théorème de Fermat et pour avoir, le premier sans doute, noté un polynôme par une seule lettre.

12.5 Utilisations des congruences

Dans les exercices suivants, on pourra raisonner modulo un entier convenable.

- 1) Pour quelles valeurs de l'entier n a-t-on $4n^2 + 17n - 31$ divisible par 5 ?
- 2) Montrer que, pour tout $n \geq 3$, l'un des nombres $2^n - 1$ et $2^n + 1$ est composé.
- 3) Montrer qu'il n'existe qu'un seul nombre premier n tel que $8n^2 + 1$ soit premier.

- 4) Montrer que la somme des cubes de trois entiers consécutifs est divisible par 9.
- 5) a) Montrer que, si $x^2 + y^2 = 0 \pmod{3}$, on a $x, y = 0 \pmod{3}$.
 b) Montrer que $x^2 + y^2 = 7\,500\,000$ n'a pas de solution dans $\mathbb{Z}$.
- 6) Montrer que $7x^2 - 3y^3 = 5$ n'a pas de solutions entières.
- 7) a) Montrer qu'un nombre de la forme $8n + 7$ n'est pas la somme de trois carrés d'entiers.
 b) Montrer que l'ensemble des entiers qui sont sommes de trois carrés d'entiers n'est pas stable par multiplication.

12.6 Petit théorème de Fermat

- 1) Soient p un nombre premier et k un entier tel que $1 \leq k < p$. Montrer que C_p^k est divisible par p .
- 2) Montrer que $a^p = a \pmod{p}$ pour tout entier a par récurrence sur a .

Commentaire. Le résultat du 1) est la base des premières démonstrations du théorème de Fermat par Leibniz (vers 1680) et Euler (1736).

- 3) Vérifier que le calcul de 2^{n-1} permet de montrer que n n'est pas premier pour $n = 15, 21, 91, 143$ (ce n'est sûrement pas la meilleure façon de montrer que ces entiers ne sont pas premiers).
- 4) On pose $n = 341$.
 a) Calculer $2^{n-1} \pmod{n}$; peut-on en déduire que n n'est pas premier ?
 b) Calculer $2^{(n-1)/4} \pmod{n}$ et $2^{(n-1)/2} \pmod{n}$. Montrer qu'on peut en déduire que n n'est pas premier.

Commentaire. Le 4) présente, sur un exemple, l'idée du test de Rabin-Miller (voir 22.11.1).

12.7 Trinôme du second degré avec racine double

Soient $a, b, c \in \mathbb{Z}$ tels que le trinôme $ax^2 + bx + c$ ait une racine double dans $\mathbb{Q}$. Montrer qu'il existe des entiers r, s, d tels que $ax^2 + bx + c = d(rx + s)^2$.

12.8 Formule de Wilson

- 1) a) Soit p un nombre premier impair. Montrer que $(p-1)! = -1 \pmod{p}$ (on pourra considérer le polynôme $X^{p-1} - 1$ de $(\mathbb{Z}/p\mathbb{Z})[X]$ ou regrouper les éléments du groupe $(\mathbb{Z}/p\mathbb{Z})^*$ avec leurs inverses).

Cette formule, nommée d'après John Wilson (1741-1793), était déjà connue de Ibn al Haytham (965-1039, souvent appelé Alhazen) et de Leibniz.

- b) Montrer la réciproque (on supposera n composé et on étudiera différents cas). On notera que ce résultat ne donne pas de moyen de vérifier si un nombre est ou

non premier, le calcul de la factorielle supposant d'effectuer un nombre de produits quasiment égal au nombre étudié.

2) On définit une fonction f sur l'ensemble des entiers supérieurs ou égaux à 3 par

$$f(n) = 2 + n \left(\left\lfloor \frac{(n+1)! + 1}{n+2} \right\rfloor - \left\lfloor \frac{(n+1)!}{n+2} \right\rfloor \right).$$

Calculer $f(n)$ pour tout $n \geq 3$ (on distinguera les cas $n+2$ premier et $n+2$ composé).

12.9 Infinité des nombres premiers, une idée d'Euler

Soit $\mathcal{P}$ l'ensemble de tous les nombres premiers. Pour tout réel $x \geq 2$, posons

$$u_x = \prod_{p \in \mathcal{P}, p \leq x} \frac{1}{(1 - 1/p)}.$$

1) Montrer que

$$u_x > \prod_{p \in \mathcal{P}, p \leq x} (1 + 1/p + \dots + 1/p^m)$$

pour tout entier m .

2) Montrer que $u_x > \sum_{n \leq x} 1/n$.

3) En déduire que $\mathcal{P}$ est infini.

4) Montrer que, pour $0 \leq u < 1$, on a

$$-\ln(1-u) - u \leq (u^2/2) \left(\frac{1}{1-u} \right).$$

5) En déduire que

$$\log u_x - \sum_{p \in \mathcal{P}, p \leq x} 1/p < 1/2,$$

puis que la série $\sum_{p \in \mathcal{P}} \frac{1}{p}$ est divergente, autrement dit que

$$\lim_{x \rightarrow \infty} \sum_{p \in \mathcal{P}, p \leq x} 1/p = +\infty.$$

12.10 Plimpton 322

On appelle triplet pythagoricien un triplet (a, b, c) d'entiers strictement positifs tels que $a^2 + b^2 = c^2$; autrement dit, a, b, c sont les longueurs des côtés d'un triangle rectangle, c étant la longueur de l'hypoténuse du triangle.

a) Soient m et n des entiers premiers entre eux et k un entier quelconque. Montrer que :

$$a = k(m^2 - n^2), \quad b = 2kmn, \quad c = k(m^2 + n^2)$$

est un triplet pythagoricien.

Dans la suite de ce problème, on propose de montrer la réciproque. Soient donc a, b, c un triplet pythagoricien tel que a, b, c n'aient aucun diviseur premier commun.

b) Montrer que (a, b) , (b, c) , (c, a) sont des couples d'entiers premiers entre eux.

c) Montrer que a ou b est pair (on pourra raisonner modulo 4).

On suppose désormais b pair et on pose $b = 2b'$.

d) Montrer que $u = \frac{c+a}{2}$ et $v = \frac{c-a}{2}$ sont des entiers premiers entre eux.

e) Montrer que u et v sont des carrés d'entiers.

f) Conclure.

g) Pour quelles valeurs de m et n obtient-on $a = 10441$ et $c = 20809$?

Commentaire. La tablette babylonienne Plimpton 322 (du nom du collectionneur George Plimpton) date des années – 1800 environ. C'est l'un des plus anciens textes mathématiques connus. Cette tablette merveilleuse donne une liste de 15 triplets pythagoriciens, comme celui du g). Une recherche au hasard de tels triplets est exclue et les premiers commentateurs de la tablette, Otto Neugebauer (1899-1990) et Abraham Sachs (1914-1983), pensaient que les Babyloniens connaissaient déjà le résultat du a), 1500 ans avant Euclide¹.

12.11 Méthode de descente infinie

Pour montrer qu'un problème arithmétique n'a pas de solution, Fermat a proposé une méthode qu'il a appelée descente infinie (voir la citation de Fermat en 21.4). Fermat l'utilise d'abord en supposant que le problème à une solution et construit une solution plus petite pour un certain critère ; en réitérant ce procédé, il aboutit à une contradiction.

Nous allons appliquer cette méthode pour montrer la conjecture de Fermat pour $n = 4$: il n'existe pas d'entiers naturels x, y, z non nuls tels que $x^4 + y^4 = z^4$ (E).

La démonstration commence en supposant qu'il existe des entiers a, b, c non nuls tels que

$$a^4 + b^4 = c^2.$$

1) Montrer qu'on peut supposer a, b, c premiers entre eux deux à deux.

2) Montrer qu'il existe, à permutation éventuelle de a et b , des entiers u, v premiers entre eux non nuls tels que :

$$a^2 = u^2 - v^2 \quad b^2 = 2uv \quad c = u^2 + v^2.$$

3) Montrer que u ou v est pair, puis que u est impair et que v est pair.

1. Pour une présentation détaillée des résultats de cette tablette et une discussion de ses résultats, voir *Faire des mathématiques à partir de leur histoire*, tome 6, IREM de Rennes, octobre 2004, 114 p.

4) Montrer qu'il existe des entiers m et n premiers entre eux tels que $a = m^2 - n^2$, $v = 2mn$ et $u = m^2 + n^2$.

5) Montrer que m, n, u sont des carrés : r^2, s^2, p^2 respectivement et que $p^2 = r^4 + s^4$. En déduire une solution de E et conclure.

Solutions

12.1 1) $37 + 55 = 29 \bmod 63$; $37 \times 55 = 19 \bmod 63$.

2) Les facteurs premiers de 18000 sont 2, 3, 5 et aucun d'entre eux ne divise 433014 481. Le pgcd est donc 1. On notera la dissymétrie de la difficulté : l'autre nombre est le carré du nombre premier 20809 et sa décomposition n'est pas évidente.

3) a) Posons $r_0 = a$, $r_1 = b$ et notons q_{k+1} et r_{k+2} les quotient et reste de la division euclidienne de r_k par r_{k+1} obtenus tant que r_{k+1} est non nul. On peut faire un tableau des résultats.

k	r_k	q_k
0	42 098	
1	36 146	1
2	5 952	6
3	434	13
4	310	1
5	124	2
6	62	2
7	0	

Le pgcd est 62.

b) La décomposition des deux nombres en produit de facteurs premiers n'est pas évidente à la main : $42\,098 = 2 \times 7 \times 31 \times 97$, $36\,146 = 2 \times 11 \times 31 \times 53$; le pgcd est $2 \times 31 = 62$.

c) On remonte les calculs :

$$\begin{aligned}
 62 &= 310 - 2 \times 124 \\
 &= 310 - 2(434 - 310) \\
 &= 3 \times 310 - 2 \times 434 \\
 &= 3 \times (5952 - 13 \times 434) - 2 \times 434 \\
 &= \dots\dots\dots \\
 &= 249 \times 42\,098 - 290 \times 36\,146
 \end{aligned}$$

ce qui donne $u = 249$, $v = -290$.

d) Si on a $62 = as + bt = au + bv$, on a $a(s - u) + b(t - v) = 0$. En simplifiant par 62, on obtient : $7 \times 97(s - u) = 11 \times 53(v - t)$. Le lemme de Gauss montre que 11×53 divise $s - u$. En posant $s - u = 11 \times 53 \times k$, avec $k \in \mathbb{Z}$, on obtient : $s = 249 + 11 \times 53k$, $t = -290 - 7 \times 97k$.

4) On applique l'algorithme de recherche du pgcd :

$$\begin{array}{rcl} 679 & = & 583 + 96 \\ 583 & = & 96 \times 6 + 7 \\ 96 & = & 7 \times 13 + 5 \\ 7 & = & 5 + 2 \\ 5 & = & 2 \times 2 + 1 \end{array}$$

et on remonte les calculs :

$$\begin{array}{rcl} 1 & = & 5 - 2 \times 2 \\ & = & 5 \times 3 - 7 \times 2 \\ & = & 96 \times 3 - 7 \times 41 \\ & = & 96 \times 249 - 583 \times 41 \\ & = & 679 \times 249 - 583 \times 290 \end{array}$$

L'inverse est $-290 = 389$.

5) Raisonnons par récurrence sur a . Pour $a = 0, 1$, le résultat est clair. Supposons qu'il soit vrai pour un entier $a > 0$. Comme $x^3 + 1 = (x + 1)(x^2 - x + 1) = (x + 1)((x + 1)^2 - 3x)$, on a $2^{3^{a+1}} + 1 = (2^{3^a} + 1)((2^{3^a} + 1)^2 - 3 \times 2^{3^a})$; le premier facteur est divisible par 3^{a+1} et le second par 3, ce qui permet de terminer la récurrence.

12.2 1) a) Comme $10 \equiv 1 \pmod{3}$, on a $S(n) = \sum_{0 \leq k \leq N} a_k 10^k \equiv \sum_{0 \leq k \leq N} a_k \pmod{3}$, d'où le critère : n est divisible par 3 si et seulement si la somme de ses chiffres est divisible par 3 (et cette somme est divisible par 3 si et seulement si la somme de ses chiffres est divisible par 3, etc.)

b) Raisonnement identique : n est divisible par 9 si et seulement si la somme de ses chiffres est divisible par 9.

c) Comme $10 \equiv -1 \pmod{11}$, on a $S(n) = \sum_{0 \leq k \leq N} a_k 10^k \equiv \sum_{0 \leq k \leq N} (-1)^k a_k \pmod{11}$, d'où le critère : n est divisible par 11 si et seulement si la somme alternée de ses chiffres est divisible par 11. Par exemple, 435 281 est divisible par 11 car $4 - 3 + 5 - 2 + 8 - 1 = 11$.

2) a) Que 13 divise n équivaut à ce que 13 divise $4n$; comme $4n = 39a + a + 4b$, le critère est démontré.

b) Que 19 divise n équivaut à ce que 19 divise $2n$; comme $2n = 19a + a + 2b$, le critère est démontré.

c) On applique les critères précédents en cascade. Pour tester si 13 divise 107 341, on calcule successivement 10738, 1 105, 130 et 13 et on conclut positivement. Pour tester si 19 divise 156 883, on calcule successivement 15 694, 1 577, 171 et 19 et on conclut aussi positivement.

12.3 a) La vérification du pas de récurrence est immédiate : si $r_k = au_k + bv_k$, $r_{k+1} = au_{k+1} + bv_{k+1}$, $r_{k+2} = r_k - r_{k+1}q_{k+1} = au_k + bv_k - (au_{k+1} + bv_{k+1})q_{k+1} = au_{k+2} + bv_{k+2}$.

b) Il faut prendre les valeurs d'indice $N - 1$.

c) Les résultats sont donnés à partir de ceux de l'exercice 12.1.c).

Pour $k = 6$, on retrouve $u = 249$ et $v = -290$.

k	r_k	q_k	u_k	v_k
0	42 098		1	0
1	36 146	1	0	1
2	5 952	6	1	-1
3	434	13	-6	7
4	310	1	79	-92
5	124	2	-85	99
6	62	2	249	-290
7	0			

12.4 a) Raisonnons par récurrence sur k . Comme $r_{N-1} \geq 1 = F_1$ et $r_{N-2} \geq 2 = F_2$, la propriété est vraie pour $k = 1$ et pour $k = 2$. Supposons la propriété vraie jusqu'à l'entier k , $2 \leq k < N$. On a $r_{N-(k+1)} = q_{N-k}r_{N-k} + r_{N-(k-1)}$ et l'hypothèse de récurrence donne $r_{N-(k+1)} \geq F_k + F_{k-1}$, donc $r_{N-(k+1)} \geq F_{k+1}$, ce qui permet de conclure.

b) Comme $10^{p-1} \leq b < 10^p$, on a $p \geq \log(b + 1)$.

Le **a)** montre que $b = r_1 \geq F_{N-1}$. On sait (voir 2.5), que $F_{N-1} = \frac{1}{\sqrt{5}}(\rho_1^N - \rho_2^N)$ avec

$$\rho_1 = \frac{1+\sqrt{5}}{2} \text{ et } \rho_2 = \frac{1-\sqrt{5}}{2}. \text{ Comme } |\rho_2^N| < 1, \text{ on a } b + 1 \geq \frac{\rho_1^N}{\sqrt{5}}, \text{ d'où}$$

$$N \leq \frac{\log(b+1)}{\log(\rho_1)} + \frac{\log(\sqrt{5})}{\log(\rho_1)}, \text{ donc } N \leq Ap + B \text{ avec } A = \frac{1}{\log(\rho_1)}, B = \frac{\log(\sqrt{5})}{\log(\rho_1)}.$$

Une calculatrice donne des valeurs approchées par excès : $A \approx 4,78$, $B \approx 1,67$. Comme $m = N - 1$, on obtient $m \leq Ap + B - 1$; comme m est entier, on a $m \leq [Ap + B - 1]$. Il reste à montrer que pour tout p strictement positif, on a $[Ap + B - 1] \leq 5p$. On vérifie directement l'inégalité pour $p = 1, 2, 3$; pour $p \geq 4$, on a $Ap + B - 1 < 4,79p + 0,68 = 5p + 0,68 - 0,21p < 5p$; d'où le théorème. On remarquera que l'estimation ne tient compte que de la taille de b et que a peut être aussi grand qu'on veut sans que cela influe sur le nombre m d'itérations.

c) Le nombre d'itérations pour déterminer le pgcd de F_{n+1} et F_n est n puisque la suite des relations $F_{k+1} = F_k + F_{k-1}$ est exactement la suite des divisions euclidiennes qu'il faut effectuer pour obtenir $\text{pgcd}(F_{n+1}, F_n) = 1$. Si $m = n = 5$, $F_5 = 8$ a un chiffre ; si $m = n = 10$, $F_{10} = 89$ a deux chiffres ; si $m = n = 15$, $F_{15} = 987$ a trois chiffres ; dans ces trois cas, la majoration du théorème de Lamé est optimale. Pour les nombres à 4 chiffres, le plus grand nombre de Fibonacci est $F_{19} = 6765$; la majoration n'est plus optimale pour le couple (F_{20}, F_{19}) puisque $m = 19$ et $4 \times 5 = 20$; elle n'est donc plus optimale pour les nombres b de quatre chiffres.

12.5 1) Modulo 5, l'expression est égale à $-(n-1)^2$. Elle est nulle modulo 5 si et seulement si $n \equiv 1 \pmod{5}$.

2) Étant donnés trois nombres successifs, comme $2^n - 1$, 2^n , $2^n + 1$, l'un des trois est divisible par 3. Comme 2^n n'est pas divisible par 3, on peut conclure. On peut aussi écrire les puissances de 2 modulo 3 : comme $2 \equiv -1 \pmod{3}$, elles sont égales à ± 1 modulo 3, par conséquent $2^n - 1$ ou $2^n + 1$ est nul modulo 3.

3) Si $n = 3$, $8n^2 + 1 = 73$ est premier. Si $n \neq 3$ est premier, on a $n \equiv \pm 1 \pmod{3}$, donc $8n^2 + 1 \equiv 0 \pmod{3}$; donc $8n^2 + 1$ n'est pas premier.

4) Écrire la somme sous la forme $(n-1)^3 + n^3 + (n+1)^3 = 3n(n^2 + 2)$ et vérifier que $n(n^2 + 2) \equiv 0 \pmod{3}$ pour $n \equiv 0, 1, 2 \pmod{3}$.

5) a) Si x ou y ne sont pas tous deux nuls modulo 3, on a $x^2 \equiv 1 \pmod{3}$ ou $y^2 \equiv 1 \pmod{3}$ et $x^2 + y^2$ n'est pas nul modulo 3.

b) Si l'équation possède une solution, comme le second membre est divisible par 3, c'est que x et y sont nuls modulo 3 ; alors $x^2 + y^2$ est divisible par 9, mais le second membre n'est pas divisible par 9.

6) On peut raisonner modulo 3 ou modulo 7 : $x^2 \equiv 2 \pmod{3}$ n'a pas de solution et $-3y^3 \equiv 5 \pmod{7}$ non plus.

7) a) Les carrés modulo 8 sont 0, 1, 4 : impossible d'obtenir 7 en faisant la somme de trois de ces nombres.

b) On voit que $3 = 0^2 + 1^2 + 1^2$ et $5 = 0^2 + 1^2 + 2^2$ sont des sommes de trois carrés. Leur produit est $15 = 8 + 7$, qui ne peut être une somme de trois carrés d'après le a).

12.6 1) Posons $C_p^k = N$. Comme $Nk!(p-k)! = p!$, on voit que p , qui divise le membre de droite, doit diviser le membre de gauche, d'après le lemme de Gauss. Comme p ne divise aucun des facteurs de $k!$ ni aucun des facteurs de $(p-k)!$, il divise N .

2) Le résultat est vrai pour $a = 1$ et s'il est vrai pour un entier $a \geq 1$, on a $(a+1)^p = a^p + C_p^1 a^{p-1} + \dots + 1 \equiv a^p + 1 \pmod{p}$ d'après le 1) ; l'hypothèse de récurrence permet de conclure.

3) $2^{14} \equiv 4 \pmod{15}$; $2^{20} \equiv 4 \pmod{21}$; $2^{90} \equiv 64 \pmod{91}$; $2^{142} \equiv 114 \pmod{143}$ montrent que les nombres 15, 21, 91, 143 ne sont pas premiers puisque le petit théorème de Fermat n'est pas vérifié. Le calcul des puissances, à la main comme sur une petite calculatrice doit être mené en réduisant suivant le module à chaque étape. Calculer 2^{142} puis réduire modulo 143 dépasse les capacités des calculatrices et n'est pas

recommandé sur des ordinateurs puissants. D'autre part, il faut essayer de faire le minimum d'opérations :

pour calculer 2^{142} , on peut écrire $142 = 2(2(2(16 + 1) + 1) + 1)$ et calculer modulo 143 successivement $2^2, 2^4, 2^8, 2^{16}, 2^{17}, 2^{34}, 2^{35}, 2^{70}, 2^{71}, 2^{142}$.

4) a) $2^{340} = 1 \pmod{341}$; la seule conclusion de ce calcul est qu'il ne montre pas que 341 n'est pas premier.

b) $2^{170} = 1 \pmod{341}$ et $2^{85} = 32 \pmod{341}$. On a donc $32^2 = 1 \pmod{341}$. Si 341 était un nombre premier, $\frac{\mathbb{Z}}{341\mathbb{Z}}$ serait un corps ; dans un corps, l'équation $x^2 - 1 = 0$, autrement dit $(x - 1)(x + 1) = 0$ n'a que les solutions ± 1 . On voit que ce n'est pas le cas ici.

12.7 Comme $4ac = b^2$, a et c sont de même signe. Posons $d = \text{pgcd}(a, c)$ en prenant le signe de d tel que $a = a'd$, $c = c'd$ et $a', c' > 0$. Comme $4ac = b^2$, on voit que d^2 divise b^2 , donc d divise b . Posons $b = b'd$. On a $ax^2 + bx + c = d(a'x^2 + b'x + c')$ et $b'^2 = 4a'c'$ avec $(a', c') = 1$, donc il existe r et s tels que $a' = r^2$ et $c' = s^2$. On obtient alors la forme demandée.

12.8 1) a) D'après le petit théorème de Fermat, les $p - 1$ éléments non nuls de $\mathbb{Z}/p\mathbb{Z} : 1, \dots, p - 1$, sont tous racines du polynôme $T = X^{p-1} - 1$. Comme T est de degré $p - 1$, ces éléments forment l'ensemble des racines de T . Le produit de ces racines est le terme constant de T , d'où la formule de Wilson.

Voici une seconde démonstration, utilisant le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$ dont les éléments sont les $p - 1$ éléments non nuls de $\mathbb{Z}/p\mathbb{Z}$. Chaque élément a de ce groupe a un inverse a^{-1} ; en général, a^{-1} est distinct de a ; le cas où $a^{-1} = a$ équivaut à $a^2 = 1$, soit $(a - 1)(a + 1) = 0$ ou $a = \pm 1$. Dans le produit $(p - 1)!$ de tous les éléments du groupe $(\mathbb{Z}/p\mathbb{Z})^*$, le produit d'un élément et de son inverse, quand ils sont distincts, vaut 1. Les seuls termes de $(p - 1)!$ qui ne se simplifient pas ainsi sont 1 et -1 , d'où la formule de Wilson.

Par exemple, si $p = 11$, les paires d'éléments distincts inverses l'un de l'autre sont : 2 et 6, 3 et 4, 5 et 9, 7 et 8 et on a

$$10! = (2 \times 6) \times (3 \times 4) \times (5 \times 9) \times (7 \times 8) \times 10 = 10 = -1 \pmod{11}.$$

b) Montrons que l'égalité $(n - 1)! = -1 \pmod{n}$ est fautive si n est composé. D'abord, si n n'est pas le carré d'un nombre premier, on peut écrire $n = ab$ avec $1 < a, b < n$ et $a \neq b$; comme $ab = n$ est un facteur de $(n - 1)!$, on a $(n - 1)! = 0 \pmod{n}$. Ensuite, si $n = p^2$ avec p premier et $p > 2$, on a $1 < p < 2p < n$, donc $2p^2$ est un facteur de $(n - 1)!$ et $(n - 1)! = 0 \pmod{n}$. Enfin, si $n = 4$, $(n - 1)! = 6 \neq -1 \pmod{4}$.

2) Si $n + 2$ est premier, la formule de Wilson montre que $\frac{(n+1)!+1}{n+2}$ est un entier N .

Comme $\left\lfloor \frac{(n+1)!}{n+2} \right\rfloor$ est la partie entière de $N - 1/(n+2)$, il est égal à $N - 1$ et on a $f(n) = n + 2$.

Si $n + 2$ est composé, on voit que $f(n) = 2$ en distinguant les mêmes cas qu'au **1)b**.

12.9 1) Remarquons que le produit définissant u_x est un produit fini, puisque le nombre de nombres premiers p tels que $p \leq x$ est fini. Pour tout $p \in \mathcal{P}$ tel que $p \leq x$ et tout entier m , on a $(1 - 1/p)(1 + 1/p + \dots + 1/p^m) = 1 - 1/p^{m+1} < 1$, d'où la majoration indiquée.

2) Posons $v_{x,m} = \prod_{p \in \mathcal{P}, p \leq x} (1 + 1/p + \dots + 1/p^m)$ et notons $p_1, \dots, p_r$ les nombres premiers inférieurs à x rangés par ordre croissant ($p_1 = 2, p_2 = 3$, etc.). On peut écrire $v_{x,m} = (1 + 1/p_1 + \dots + 1/p_1^m) \dots (1 + 1/p_r + \dots + 1/p_r^m)$. Un terme de $v_{x,m}$ est l'inverse d'un produit de la forme $p_1^{k_1} \dots p_r^{k_r}$ avec $k_i \leq m$ pour $1 \leq i \leq r$. Tous les entiers n tels que $n \leq 2^m$ ont une décomposition en produit de facteurs premiers de cette forme, donc on a, pour tout entier m tel que $2^m > x$: $u_x > v_{x,m} > \sum_{n \leq x} 1/n$.

3) Si $\mathcal{P}$ était fini, l'ensemble $\mathcal{P}$ serait majoré par un nombre premier P . Pour $x > P$, u_x serait constant. Mais on vient de montrer que u_x majore la somme partielle $\sum_{n \leq x} 1/n$ de la série harmonique ; comme cette série est divergente, on a $\lim_{x \rightarrow \infty} u_x = +\infty$, contradiction.

4) On peut étudier les variations de la fonction $(1 - u)(-\ln(1 - u) - u) - u^2/2$ pour $0 \leq u < 1$ ou utiliser, si on le connaît, le développement en série $-\ln(1 - u) = u + u^2/2 + u^3/3 + \dots$; on a $-\ln(1 - u) - u = u^2/2 + u^3/3 + \dots \leq (u^2/2)(1 + u + u^2 + \dots) = (u^2/2) \left(\frac{1}{1-u} \right)$

5) On en déduit

$$\begin{aligned} \log u_x - \sum_{p \in \mathcal{P}, p \leq x} 1/p &= \sum_{p \in \mathcal{P}, p \leq x} -\log(1 - 1/p) - 1/p \\ &< \sum_{p \in \mathcal{P}, p \leq x} (1/2 p^2) \left(\frac{1}{1 - 1/p} \right). \end{aligned}$$

Comme $(1/2 p^2) \left(\frac{1}{1 - 1/p} \right) = \frac{1}{2p(p-1)}$ et que

$$\sum_{p \in \mathcal{P}, p \leq x} \frac{1}{2p(p-1)} < \sum_{n \geq 2} \frac{1}{2n(n-1)} = \sum_{n \geq 2} (1/2)(1/(n-1) - 1/n) = 1/2,$$

(la dernière somme s'obtient par télescopage), on a

$$\sum_{p \in \mathcal{P}, p \leq x} \frac{1}{p} > \log u_x - 1/2,$$

ce qui prouve la divergence de $\sum_{p \in \mathcal{P}} \frac{1}{p}$.

12.10 b) Si un nombre premier p divise a et b , il divise $a^2 + b^2 = c^2$ donc il divise c d'après le lemme de Gauss. On raisonne de même dans les deux autres cas.

c) Si a et b sont impairs, on a $a^2 = b^2 = 1 \pmod{4}$, donc $a^2 + b^2 = 2 \pmod{4}$. Comme un carré n'est jamais égal à 2 modulo 4, c ne peut exister.

d) Si un nombre premier p divise u et v , il divise $u + v = c$ et $u - v = a$, ce qui est impossible.

e, f) On a $uv = b'^2$. La décomposition en produit de facteurs premiers des deux membres fait apparaître dans le membre de droite des nombres premiers avec des exposants pairs. Comme u et v sont premiers entre eux, un nombre premier qui apparaît dans le membre de droite ne peut être facteur que de l'un des nombres u ou v . Comme son exposant est pair, on en déduit que u et v sont des produits de puissances paires de nombres premiers et sont donc des carrés d'entiers. En posant $u = m^2$ et $v = n^2$, on retrouve les formules du a).

g) $m = 125$, $n = 72$.

12.11

1) Si un nombre premier p divisait a, b , alors p diviserait c^2 , donc c ; s'il divisait a et c , alors il diviserait b^4 , donc b ; de même, s'il divisait b et c , il diviserait a . Dans ces trois cas, a/p , b/p et c/p^2 donneraient une solution de l'équation (E).

2) Le triplet a^2, b^2, c est un triplet pythagoricien et l'exercice 12.10 assure l'existence, à permutation éventuelle de a et b près, des entiers u et v . Comme a et c sont premiers entre eux, il en est de même de u et v .

3) Comme $b^2 = 2uv$, b est pair, donc $2uv$ est divisible par 4, ce qui conduit à uv pair, puis à u ou v est pair.

Si u est pair et v impair, l'égalité $a^2 = u^2 - v^2$ donne $a^2 = -1 \pmod{4}$, ce qui est impossible. Par conséquent, u est impair et v est pair.

4) Le triplet a, v, u est pythagoricien, car $a^2 + v^2 = u^2$. Comme v est pair, on obtient l'existence de m et n .

5) Posons $v = 2v'$. On a $b^2 = 4uv'$; comme u et v' sont premiers entre eux, ce sont des carrés et on peut poser $u = p^2$, $v' = q^2$. Comme m et n sont premiers entre eux, $v = 2q^2 = 2mn$ montre que m et n sont des carrés : $m = r^2$ et $n = s^2$. On a alors $p^2 = r^4 + s^4$. On obtient une solution (r, s, p) de (E) avec $p < u^2 + v^2 = c$. Si on est parti au départ d'une relation $x^4 + y^4 = z^2$ avec z minimal, on obtient une contradiction. L'équation $x^4 + y^4 = z^2$ n'a donc pas de solutions en nombres entiers et il en est de même pour l'équation $x^4 + y^4 = z^4$.

13.1 INTRODUCTION

C'est vers 1590 que François Viète (1540-1603) écrit les premiers calculs avec des lettres. Avant lui, les résolutions d'équations algébriques étaient traitées en proposant des exemples numériques.

Ainsi, Mohammed Al Khwarizmi (vers 780-vers 850), dans son traité d'algèbre écrit vers 825, classifie les équations du second degré en différents types pour ne considérer que des coefficients positifs. Par exemple :

un carré et 21 dirhams sont égaux à dix racines

qui correspond à l'équation $x^2 + 21 = 10x$, est une équation du type $x^2 + c = bx$ avec $b, c > 0$ (pour Al Khwarizmi qui traite de problèmes d'héritage, les nombres sont exprimés en dirhams, l'unité monétaire arabe). « Dans ce cas, précise Al Khwarizmi, saches que si tu divises en 2 la racine (il faut comprendre : le coefficient b de x), que tu la multiplies par elle-même, et que le produit soit plus petit que les dirhams (il faut comprendre : $\frac{b^2}{4} < c$) alors le problème est impossible ». On comprend ce que l'auteur veut dire, mais l'usage de lettres rend les choses plus simples.

Le calcul littéral permet à Descartes, en 1637, dans sa *Géométrie*, de dégager la notion de polynôme et de montrer comment faire la division par $X - a$.

13.2 POLYNÔMES SUR UN CORPS K

13.2.1 Nécessité de distinguer polynômes et fonctions polynomiales

Exemples. Pour comprendre la nécessité de distinguer les polynômes et les fonctions polynomiales, regardons ce qui se passe pour les polynômes à coefficients dans le corps $K = \frac{\mathbb{Z}}{2\mathbb{Z}}$. Posons $P(X) = X^2 - X$. La fonction polynomiale associée à P est

la fonction $p : K \rightarrow K$ définie par $p(x) = x^2 - x$. On a $p(0) = p(1) = 0$. Donc p est la fonction nulle alors que P n'est pas nul.

Plus généralement, si P est un polynôme quelconque à coefficients dans ce corps, la division euclidienne de P par $X^2 - X$ donne $P(X) = (X^2 - X)Q(X) + R(X)$ avec $\deg(R) < 2$. Pour tout x de K , on a donc $p(x) = r(x)$, où p et r sont les fonctions polynomiales associées à P et à R . Comme $r(x)$ est de la forme $ax + b$ avec $a, b \in K$, on voit qu'il existe 4 fonctions polynomiales distinctes alors qu'il existe une infinité de polynômes à coefficients dans K .

On peut généraliser ce raisonnement au corps $\frac{\mathbb{Z}}{p\mathbb{Z}}$ avec p premier, en considérant le polynôme $X^p - X$: la fonction polynomiale associée est $x \mapsto x^p - x$ qui est nulle pour tout x de $\frac{\mathbb{Z}}{p\mathbb{Z}}$ d'après le petit théorème de Fermat.

Cas des corps infinis. Si le corps K est infini, par exemple si $K = \mathbb{R}$, nous verrons en 13.5 que la correspondance entre polynômes et fonctions polynomiales est bijective. C'est ce qui nous a permis de ne pas distinguer jusqu'ici les deux notions.

13.2.2 Construction de l'ensemble des polynômes

Dans ce paragraphe, nous nous proposons de construire l'ensemble des polynômes à coefficients dans un corps K et d'en dégager les propriétés structurelles. Répétons que cette construction est nécessaire car l'écriture $\sum_{0 \leq k \leq n} a_k X^k$ n'a pas de sens *a priori*. Nous désignerons le corps par K ; dans les exemples que nous prendrons, le corps sera souvent celui des nombres réels ou celui des nombres complexes. On va construire les polynômes en partant de la suite de leurs coefficients : $(a_0, a_1, \dots, a_n, 0, 0, \dots)$.

L'espace vectoriel S_f . On a défini au chapitre 2, l'espace vectoriel des suites de nombres réels. Ici, nous noterons S l'espace des suites d'éléments du corps K . Considérons le sous-ensemble S_f de S des suites ayant un nombre fini de termes non nuls. Il est clair que

S_f est un sous-espace vectoriel de S , car si deux suites ont un nombre fini de termes non nuls, leur somme également et le produit de l'une d'elles par un scalaire aussi.

Pour tout entier $k \geq 0$, notons e_k la suite de S_f dont tous les termes sont nuls, sauf le k -ième qui est égal à 1. Chaque élément de S_f s'écrit comme combinaison linéaire des suites e_k : en effet, si $u = (u_k)$ est une suite de S_f dont les termes de rang $\geq n$ sont nuls, on a $u = \sum_{0 \leq k \leq n} u_k e_k$. La famille $(e_k)_{k \geq 0}$ est donc une famille génératrice de S_f .

Montrons maintenant que la famille $(e_k)_{k \geq 0}$ est une famille libre de S_f . En effet, si $0 = \sum_{0 \leq k \leq n} u_k e_k$, la suite $(u_0, \dots, u_n, 0, \dots)$ est la suite nulle, donc $u_0 = \dots = u_n = 0$. On voit que la famille infinie $(e_k)_{k \geq 0}$ est une base de S_f et que S_f est un espace de dimension infinie.

Définition du produit. Pour retrouver les polynômes, définissons un produit sur S_f qui donne ce qu'on veut. Si $u = (u_n)$ et $v = (v_n)$ sont des suites de S_f , on définit $w = uv$ par $w_h = \sum_{0 \leq p \leq n} u_p v_{h-p} = \sum_{p+q=h} u_p v_q$. On obtient bien un élément de S_f puisque si u_n est nul pour $n > M$ et si v_n est nul pour $n > N$, on vérifie que w_n est nul pour $n > M + N$.

Propriétés du produit. La loi de multiplication que nous avons définie a les propriétés attendues.

Elle est associative car :

$$[u(vw)]_h = \sum_{p+m=h} u_p (vw)_m = \sum_{p+m=h} u_p \left(\sum_{q+r=m} v_q w_r \right) = \sum_{p+q+r=h} u_p v_q w_r$$

$$\text{et de même : } [(uv)w]_n = \sum_{p+q+r=n} u_p v_q w_r.$$

On vérifie de même les propriétés suivantes :

- Elle est distributive par rapport à l'addition : $u(v + w) = uv + uw$.
- Elle est commutative : $uv = vu$.
- Enfin, pour tout scalaire a de K : $(au)v = u(av) = a(uv)$.

Changement de notation. Posons alors $X = e_1$; X s'appelle l'indéterminée (les anglais disent souvent variable ; on pourrait choisir une autre lettre : T , Y , etc.). La définition précédente nous montre que $X^2 = XX = e_2$ et plus généralement $X^n = e_n$ pour tout entier $n \geq 1$. La suite e_0 est un élément neutre pour cette multiplication et nous la noterons 1. Une suite $u = (u_k)$ de S_f dont les termes de rang $> n$ sont nuls peut maintenant s'écrire $u = u_0 e_0 + \dots + u_n e_n = u_0 1 + u_1 X + \dots + u_n X^n$. Ce simple changement de notation nous donne une notation cohérente avec celle des fonctions polynomiales.

Un polynôme peut se noter par une seule lettre comme P , Q , etc. On définira, par exemple, le polynôme $P = X^3 + 1$; on écrit souvent $P(X)$ et non P seul.

Un polynôme ne comportant qu'un seul terme est appelé monôme ; par exemple, le polynôme $-7X^3$ est un monôme.

Récapitulation. On notera désormais $K[X]$ l'ensemble S_f et ses éléments seront appelés polynômes en X .

L'ensemble $K[X]$ est muni de plusieurs lois :

- 1) une loi d'addition des polynômes entre eux ;
- 2) une loi de multiplication des polynômes entre eux ;
- 3) une loi de multiplication des polynômes par les scalaires.

Les deux premières lois munissent $K[X]$ d'une structure d'anneau commutatif unitaire ; la première et la troisième loi munissent $K[X]$ d'une structure de K -espace vectoriel. Les trois lois munissent $K[X]$ de ce qu'on appelle une structure de K -algèbre commutative, nous développerons cet aspect au paragraphe 8.

Généralisation à un anneau. On peut généraliser cette construction : on aura plus tard besoin de considérer, par exemple, des polynômes à coefficients entiers. La construction de $\mathbb{Z}[X]$ est tout à fait analogue à ce que nous venons de faire, en partant de l'ensemble des suites finies d'entiers, mais comme $\mathbb{Z}$ est un anneau et n'est pas un corps, la structure obtenue avec les lois 1 et 3 n'est pas une structure d'espace vectoriel mais de module ; nous n'en dirons pas plus.

Définition : évaluation. Étant donné un polynôme $P(X) = \sum_{0 \leq k \leq n} a_k X^k$ de $K[X]$ et un élément x de K , on appelle évaluation ou valeur de P en x l'élément $\text{ev}_x(P) = P(x) = \sum_{0 \leq k \leq n} a_k x^k$. Les propriétés de l'évaluation par rapport aux trois lois de $K[X]$ sont :

$$\begin{aligned}(P_1 + P_2)(x) &= P_1(x) + P_2(x) ; \\ (aP)(x) &= aP(x) ; \\ (P_1 P_2)(x) &= P_1(x)P_2(x).\end{aligned}$$

Autrement dit, l'évaluation en x est une application linéaire $K[X] \rightarrow K$ qui est compatible avec les produits.

La fonction polynomiale associée à un polynôme est la fonction

$$x \mapsto \text{ev}_x(P) = P(x).$$

13.3 DEGRÉ, DIVISION EUCLIDIENNE

Définition : degré. On appelle degré d'un polynôme non nul $P(X) = \sum_{0 \leq k \leq n} a_k X^k$ et on note $\deg(P)$ l'entier $d = \sup\{k \mid a_k \neq 0\}$. On pose $\deg(0) = -\infty$ en considérant $-\infty$ comme un nouveau symbole satisfaisant les relations $-\infty < n$, $-\infty + n = -\infty$ pour tout entier n de $\mathbb{N}$. Le degré d'un polynôme constant non nul est 0 et seuls les polynômes constants non nuls sont de degré 0. Si P est un polynôme de degré d , on appelle a_d coefficient dominant de P et $a_d X^d$ terme dominant de P ; si $a_d = 1$, P est dit unitaire.

On peut vérifier que :

- (1) $\deg(P + Q) \leq \sup\{\deg(P), \deg(Q)\}$;
 (2) si $\deg(P) \neq \deg(Q)$, $\deg(P + Q) = \sup\{\deg(P), \deg(Q)\}$;
 (3) $\deg(PQ) = \deg(P) + \deg(Q)$.

Ces relations sont vraies même si $P = 0$ ou si $Q = 0$.

Remarque. La relation (3) est vraie pour tout corps mais est fausse pour les anneaux en général : par exemple, $\bar{2}X \times \bar{4}X = \bar{8}X^2 = 0$ dans $\frac{\mathbb{Z}}{8\mathbb{Z}}[X]$. Pour la démontrer, il suffit de remarquer que, si P et Q sont non nuls, le coefficient dominant de PQ est le produit des coefficients dominants de P et de Q . Cette remarque montre que, si K est un corps, le produit de deux polynômes non nuls de $K[X]$ est un polynôme non nul, autrement dit l'anneau $K[X]$ est un anneau intègre.

Proposition : division euclidienne dans $K[X]$. Soient A et B des polynômes de $K[X]$ avec $B \neq 0$. Il existe un couple unique de polynômes Q et R de $K[X]$ tels que $A = BQ + R$ et $\deg(R) < \deg(B)$.

Démonstration. Montrons d'abord l'existence de Q et R en raisonnant par récurrence sur le degré de A . Si $\deg(A) < \deg(B)$, il suffit de poser $Q = 0$ et $R = A$. Si $\deg(A) \geq \deg(B)$, posons $m = \deg(A)$ et $n = \deg(B)$ et notons a_m et b_n les coefficients dominants de A et de B . Posons $A_1 = A - a_m b_n^{-1} X^{m-n} B$. Le terme de degré m de A disparaît dans cette soustraction. On a donc $\deg(A_1) < \deg(A)$ et l'hypothèse de récurrence montre qu'il existe des polynômes Q et R de $K[X]$ tels qu'on ait $A_1 = BQ + R$ et $\deg(R) < \deg(B)$.

On en déduit $A = B(Q + a_m b_n^{-1} X^{m-n}) + R$ ce qui montre l'existence du couple cherché.

Si deux couples (Q, R) et (Q_1, R_1) satisfont les mêmes conditions, on a $B(Q - Q_1) = R_1 - R$; comme $\deg(R_1 - R) < \deg(B)$, ceci implique $Q - Q_1 = 0$ et $R - R_1 = 0$. $\square$

Les polynômes Q et R sont appelés le quotient et le reste de la division euclidienne de A par B .

La démonstration montre comment procéder pour obtenir Q et R : c'est la procédure habituelle. Donnons un exemple.

$$\begin{array}{rcll}
 A = & X^3 & +X & +1 \\
 -X^2 B = & -X^3 -X^2 & & \\
 A_1 = A - X^2 B = & -X^2 & +X & +1 \\
 -(-XB) = & X^2 & +X & \\
 A_2 = A_1 + XB = & & 2X & +1 \\
 -2B = & & -2X & -2 \\
 R = A_2 - 2B = & & & -1
 \end{array}
 \quad \left| \begin{array}{l}
 B = X + 1 \\
 \hline
 Q = X^2 - X + 2
 \end{array} \right.$$

La division euclidienne de A par B s'écrit :

$$X^3 + X + 1 = (X + 1)(X^2 - X + 2) - 1$$

Remarque. Le résultat s'étend aux polynômes à coefficients dans un anneau commutatif unitaire sous la condition que le coefficient dominant de B soit inversible dans l'anneau, ce qui est le cas, par exemple, si B est unitaire. On pourra donc effectuer des divisions euclidiennes dans $\mathbb{Z}[X]$ par des polynômes unitaires.

13.4 PGCD DE POLYNÔMES

L'anneau $K[X]$ des polynômes à coefficients dans un corps K a une propriété extrêmement importante en commun avec l'anneau $\mathbb{Z}$: c'est la division euclidienne. Cette propriété permet d'obtenir dans $K[X]$ des propriétés tout à fait analogues aux propriétés de l'anneau $\mathbb{Z}$. On ne sera donc pas étonné de constater le parallélisme de ce paragraphe avec ce qui a été fait au chapitre précédent.

Éléments inversibles de $K[X]$. Soit A un élément inversible de $K[X]$ et B son inverse. On a $AB = 1$ d'où $\deg(A) + \deg(B) = 0$ et $A, B \neq 0$; comme $\deg(A)$ et $\deg(B)$ sont ≥ 0 , on a $\deg(A) = 0$; donc A est un polynôme constant non nul.

Tout polynôme non nul de $K[X]$ s'écrit sous la forme uP où u est constant non nul et P est unitaire.

Définition : divisibilité. Soient A et B dans $K[X]$. On dit que B divise A s'il existe C dans $K[X]$ tel que $A = BC$.

Propriétés. Comme un polynôme constant non nul est inversible dans $K[X]$, il divise tout polynôme.

Pour A et B non nuls, si A divise B et si B divise A , alors $A = uB$ avec $u \in K^*$. En effet, il existe des polynômes C et C_1 tels que $B = AC$, $A = BC_1$, $(1 - CC_1)A = 0$, donc $CC_1 = 1$, ce qui prouve que C est inversible donc constant.

Enfin, la relation de divisibilité s'exprime en termes d'idéaux : si B divise A , c'est que $(B) \supset (A)$.

Proposition 1. *L'anneau $K[X]$ est un anneau principal.*

Démonstration. Soit I un idéal de $K[X]$. Si I n'est pas réduit à l'idéal $\{0\}$ engendré par 0, il contient des éléments de degré ≥ 0 . L'ensemble de ces degrés admet un plus petit élément m . Notons B un polynôme de degré m de I et montrons que l'idéal (B) engendré par B est égal à I . On a déjà $(B) \subset I$. Si A est un polynôme quelconque de I , la division euclidienne de A par B montre qu'il existe des polynômes Q et R tels que $A = BQ + R$, avec $\deg(R) < \deg(B)$. Comme $R = A - BQ$, R est un élément de I , donc $R = 0$ et $A \in (B)$; d'où $I = (B)$. □

On remarquera que si B et B_1 engendrent le même idéal I , B divise B_1 et B_1 divise B ; par conséquent, il existe un élément u de K^* tel que $B_1 = uB$. Il existe donc, si $I \neq 0$, un unique polynôme unitaire engendrant I .

Pgcd de deux polynômes. Soient A et B deux polynômes de $K[X]$. L'idéal I qu'ils engendrent (voir 12.9) est principal, engendré par un élément D unique à un élément inversible près. On dit que D est un (on dit assez souvent *le* mais c'est un abus de langage) pgcd de A et B ; on écrit $D = \text{pgcd}(A, B)$; pour tout $u \in K$ non nul, $uD = \text{pgcd}(A, B)$; on peut donc choisir D unitaire.

Comme A et B sont dans l'idéal (D) , on voit que D est un diviseur de A et de B . Soit D_1 un diviseur de A et de B ; on a $(D_1) \supset (A)$, $(D_1) \supset (B)$, donc $(D_1) \supset (D)$, autrement dit : D_1 divise D ; donc tout diviseur commun à A et B divise D . On retrouve bien l'idée de plus grand commun diviseur.

De même, étant donnés des polynômes $A_1, \dots, A_n$, leur pgcd est un polynôme D engendrant l'idéal principal engendré par $A_1, \dots, A_n$; il est défini à un polynôme constant non nul près.

Proposition 2 : identité de Bézout.

- 1) Soient A et B deux polynômes de $K[X]$ et D leur pgcd (à un élément inversible près). Il existe des polynômes U et V tels que $UA + VB = D$.
- 2) Soient $A_1, \dots, A_n$ des polynômes et D leur pgcd (à un élément inversible près). Il existe des polynômes $U_1, \dots, U_n$ tels que $U_1 A_1 + \dots + U_n A_n = D$.

Démonstration. Comme dans le chapitre précédent, on remarque que l'idéal (D) engendré par les deux polynômes A et B (resp : par les n polynômes $A_1, \dots, A_n$) est l'ensemble des polynômes de la forme $SA + TB$, avec S et T dans $K[X]$ (resp : $S_1 A_1 + \dots + S_n A_n$ avec $S_1, \dots, S_n$ dans $K[X]$). $\square$

Calcul du pgcd et de l'identité de Bézout. Étant donnés deux polynômes A et B de $K[X]$, le pgcd de ces deux polynômes se calcule exactement de la même façon que le pgcd de deux entiers : par divisions successives, jusqu'à obtenir un reste nul : le pgcd est le reste de l'avant dernière division (voir 12.12.1).

On notera que le résultat de l'algorithme n'est pas en général un polynôme unitaire ; par exemple, un ordinateur pourra trouver un rationnel avec un numérateur et un dénominateur de grande taille comme pgcd de deux polynômes de $\mathbb{Q}[X]$ premiers entre eux, alors qu'on s'attend à le voir trouver 1.

De même, les algorithmes du chapitre précédent donnent des méthodes pour trouver les polynômes U et V de l'identité de Bézout : soit on remonte les calculs, soit on définit par récurrence des suites U_k et V_k (voir 12.12.2 et l'exercice 12.3).

Définition : polynômes premiers entre eux. Des polynômes $A_1, \dots, A_n$ sont dits premiers entre eux si leur pgcd est 1 (à un élément inversible de K près).

Proposition 3.

1) Deux polynômes de $K[X]$ sont premiers entre eux si et seulement s'il existe des polynômes U et V de $K[X]$ tels que $UA + VB = 1$.

2) Des polynômes $A_1, \dots, A_n$ sont premiers entre eux si et seulement s'il existe des polynômes $U_1, \dots, U_n$ tels que $U_1 A_1 + \dots + U_n A_n = 1$.

Démonstration. Dans un sens, c'est l'identité de Bézout. La réciproque résulte de la définition du pgcd. $\square$

Pour la démonstration des deux résultats suivants, on suit la démonstration des résultats analogues de 12.12.3.

Proposition 4 : lemme de Gauss. Soient A, B, C trois polynômes de $K[X]$. Si A divise le produit BC et si $\text{pgcd}(A, B) = 1$ alors A divise C .

Corollaire. Soient A et B deux polynômes de $K[X]$ premiers entre eux. Si chacun d'eux divise un polynôme C de $K[X]$, alors leur produit AB divise C .

13.5 RACINES D'UN POLYNÔME

Définition 1 : racine d'un polynôme. Soit P un polynôme de $K[X]$. Un élément a de K est appelé racine de P (on dit aussi zéro de P) si $P(a) = 0$.

Proposition 1.

1) La division euclidienne de P par $X - a$ s'écrit $P = (X - a)Q + P(a)$.

2) Un polynôme P de $K[X]$ est divisible par $X - a$ si et seulement si $P(a) = 0$.

Démonstration. 1) La division euclidienne de P par $X - a$ donne $P = (X - a)Q + R$, avec $\deg(R) < 1$ donc R est un élément de K ; en évaluant l'égalité en a , on a : $R = P(a)$.

2) Si P est divisible par $X - a$, $P = (X - a)Q$ et $P(a) = 0$. La réciproque est donnée par le 1). $\square$

Définition 2 : multiplicité d'une racine. Soient P un polynôme non nul de $K[X]$ et a dans K . On appelle ordre ou multiplicité de a dans P l'exposant de la plus grande puissance de $X - a$ qui divise P . On le note $m(P, a)$.

Il est clair que $m(P, a) > 0$ équivaut à $P(a) = 0$, c'est-à-dire à : a est racine de P . En posant $m = m(P, a)$, on a donc $P = (X - a)^m P_1$ avec $P_1(a) \neq 0$, puisque sinon P_1 serait divisible par $X - a$ donc P serait divisible par $(X - a)^{m+1}$.

On dit que P admet une racine multiple dans le cas d'une racine de multiplicité ≥ 2 .

Proposition 2.

- 1) Soient a et b deux racines distinctes d'un polynôme P non nul de $K[X]$ de multiplicités respectives m et n ; alors P est de la forme $(X - a)^m (X - b)^n P_1$ avec $P_1(a) \neq 0$, $P_1(b) \neq 0$ et $m + n \leq \deg(P)$.
- 2) Généralisation : si $a_1, \dots, a_n$ sont des racines deux à deux distinctes d'un polynôme P non nul de $K[X]$ de multiplicités respectives $m_1, \dots, m_n$, alors P est de la forme $(X - a_1)^{m_1} \dots (X - a_n)^{m_n} P_1$ avec $P_1(a_1), \dots, P_1(a_n)$ non nuls et $m_1 + \dots + m_n \leq \deg(P)$.

Si $m_1 + \dots + m_n = \deg(P)$, alors $P = u \prod_{1 \leq i \leq n} (X - a_i)^{m_i}$ avec u dans K^* .

Démonstration. 1) Les hypothèses impliquent que $P = (X - a)^m P_2 = (X - b)^n P_3$ avec $P_2(a) \neq 0$ et $P_3(b) \neq 0$. Comme $P(a) = 0$ et que $a - b \neq 0$, on a $P_3(a) = 0$. Soit $r = m(P_3, a)$. On a donc : $P = (X - a)^m P_2 = (X - b)^n (X - a)^r P_1$ avec $P_1(a) \neq 0$. En simplifiant et en évaluant en a , les inégalités $m < r$ et $m > r$ conduisent à des impossibilités ; on a donc $m = r$ et le résultat.

2) Immédiat par récurrence sur n . □

Corollaire : degré et nombre de racines. Un polynôme P non nul de degré n de $K[X]$ a au plus n racines distinctes dans K . □

Cette proposition n'est plus valable pour un polynôme à coefficients dans un anneau. Par exemple, dans $\frac{\mathbb{Z}}{8\mathbb{Z}}[X]$ le polynôme $X^2 - 1$ est du second degré et s'annule 4 fois : pour $\bar{1}, \bar{3}, \bar{5}, \bar{7}$.

Polynômes et fonctions polynomiales. Nous avons expliqué en 13.2.1 qu'il fallait distinguer les polynômes et les fonctions polynomiales. Cette distinction est nécessaire pour les corps finis. Cependant, pour un corps infini, $\mathbb{R}$ par exemple, nous allons montrer que la correspondance entre un polynôme et la fonction polynomiale qu'il définit est une bijection.

Proposition 3. Soit K un corps infini.

- 1) Soient $P(X) = \sum_{0 \leq k \leq n} a_k X^k$ un polynôme de $K[X]$ de degré $n \geq 0$, $p : K \rightarrow K$ la fonction polynomiale associée et $x_1, \dots, x_{n+1}$ des éléments de K distincts. On ne peut avoir $p(x_1) = \dots = p(x_{n+1}) = 0$.
- 2) Le seul polynôme à coefficients dans K dont la fonction polynomiale associée est la fonction nulle est le polynôme nul.
- 3) Si deux polynômes P_1 et P_2 à coefficients dans K définissent la même fonction polynomiale, ils sont égaux.
- 4) La correspondance entre polynômes et fonctions polynomiales associées est bijective.

Démonstration. 1) Raisonnons par récurrence sur n . La proposition est vraie pour $n = 0$. Supposons $n > 0$. La division euclidienne de P par $X - x_{n+1}$ donne : $P(X) = (X - x_{n+1})Q(X) + r$ où Q est un polynôme de degré $n - 1$ et r est un élément de K ; notons q la fonction polynomiale associée à Q . Si $p(x_{n+1}) = 0$, on a $r = 0$, donc $P(X) = (X - x_{n+1})Q(X)$. Comme Q est de degré $n - 1$, on ne peut avoir $q(x_1) = \dots = q(x_n) = 0$, donc on ne peut avoir $p(x_1) = \dots = p(x_n) = 0$.

2) Si P n'est pas le polynôme nul, son degré est $n \geq 0$ et il s'annule au plus n fois d'après le 1).

3) Appliquer le 2) à $P_1 - P_2$. Le 4) en résulte. $\square$

Polynômes scindés. Un polynôme P de $K[X]$ est dit scindé sur K s'il se factorise (*se scinde*) en un produit de facteurs du premier degré de $K[X]$, autrement dit s'il admet des racines $a_1, \dots, a_n$ dans K de multiplicités respectives $m_1, \dots, m_n$ telles que $m_1 + \dots + m_n = \deg(P)$; on a alors $P = u \prod_{1 \leq i \leq n} (X - a_i)^{m_i}$ avec u dans K^* .

Exemples. Le polynôme $2(X - 3)^3(X + 2)^2$ est scindé sur $\mathbb{R}$ et sur $\mathbb{C}$.

Le polynôme $X^2 + X + 1 = (X - j)(X - j^2)$ est scindé sur $\mathbb{C}$ mais n'est pas scindé sur $\mathbb{R}$.

Théorème de d'Alembert (ou théorème fondamental de l'algèbre). *Tout polynôme de $\mathbb{C}[X]$ est scindé sur $\mathbb{C}$.*

Ce théorème est un résultat difficile que les Français appellent théorème de d'Alembert et les anglo-saxons théorème fondamental de l'algèbre. Albert Girard (1595-1632) avait déjà énoncé ce théorème en 1629, sans démonstration ; Jean Le Rond d'Alembert (1717-1783), Euler, Pierre Simon de Laplace (1749-1827) en donnent des démonstrations plus ou moins convaincantes, plus ou moins complètes dans la seconde moitié du XVIII^e siècle. On considère que c'est Gauss qui a donné les premières démonstrations satisfaisantes. Une démonstration fréquemment donnée aujourd'hui utilise la théorie des fonctions de variable complexe enseignée en second cycle.

13.6 DÉRIVATION

La notion de dérivation pour les fonctions de variables réelles, en particulier pour les fonctions polynomiales, peut se définir formellement sur les polynômes.

On appelle dérivation l'application K -linéaire $D : K[X] \rightarrow K[X]$ définie par $D(X^n) = nX^{n-1}$ si $n > 0$ et $D(a) = 0$ si $a \in K$.

On doit vérifier que l'application D a bien la même propriété que la dérivation pour le produit de deux fonctions.

Proposition 1 : dérivation d'un produit. $D(PQ) = D(P)Q + PD(Q)$.

Démonstration. Comme D est linéaire, il suffit de vérifier cette formule pour $P = \sum_{n \geq 0} a_n X^n$, avec les a_n presque tous nuls, et $Q = X^k$, avec $k > 0$. On a :

$$\begin{aligned} D(P)X^k + P D(X^k) &= \\ &= \left(\sum_{n \geq 1} n a_n X^{n-1} \right) X^k + \left(\sum_{n \geq 0} a_n X^n \right) (k X^{k-1}) \\ &= \sum_{n \geq 0} (n+k) a_n X^{n+k-1} \\ &= D(PX^k). \end{aligned}$$

□

On note comme d'habitude $D(P)$ par P' (qu'on appelle polynôme dérivé de P) et $D^k(P)$ par $P^{(k)}$ (qu'on appelle polynôme dérivé k -ième de P) ; pour $k = 0$, on pose $D^0(P) = P$.

Cas des corps finis. La dérivation que nous venons de définir n'a pas exactement les propriétés de la dérivation des fonctions de variables réelles. Si le corps est fini, un polynôme non constant peut avoir un polynôme dérivé nul : c'est le cas de $X^p, X^{2p}, \dots$ dans $\frac{\mathbb{Z}}{p\mathbb{Z}}[X]$. Ceci ne peut se produire si le corps de base est $\mathbb{R}$ ou $\mathbb{C}$ puisque, si P n'est pas constant, la dérivation du terme de plus haut degré de P donne un terme non nul de $D(P)$.

Proposition 2 : formule de Taylor. On suppose que le corps K est un sous-corps de $\mathbb{C}$ (un corps contenu dans $\mathbb{C}$), comme $\mathbb{R}$ ou $\mathbb{C}$, par exemple. Soit $P(X)$ un polynôme de $K[X]$.

1) Si $P(X) = a_n X^n + \dots + a_1 X + a_0$, alors, pour $k = 0, \dots, n$, on a :

$$a_k = \frac{P^{(k)}(0)}{k!}, \text{ autrement dit :}$$

$$P(X) = \sum_{k \leq n} \frac{P^{(k)}(0)}{k!} X^k.$$

2) Plus généralement, si a est un élément de K , on a :

$$P(X+a) = \sum_{k \leq n} \frac{P^{(k)}(a)}{k!} X^k.$$

3) Si a est un élément de K , on a également :

$$P(X+a) = \sum_{k \leq n} a^k \frac{P^{(k)}(X)}{k!}.$$

Démonstration. 1) Il suffit de calculer les $P^{(k)}(0)$ en dérivant k fois $a_n X^n + \dots + a_1 X + a_0$ et en calculant la valeur en 0 du résultat.

2) Appliquer la formule précédente au polynôme $P(X + a)$.

3) Raisonnons par récurrence sur n . L'hypothèse de récurrence donne $P'(X + a) =$

$$\sum_{k \leq n-1} a^k \frac{P^{(k+1)}(X)}{k!}. \text{ On en déduit } P(X + a) = \sum_{k \leq n-1} a^k \frac{P^{(k)}(X)}{k!} + C \text{ où } C \in K; \text{ pour } X = 0, \text{ le 1) donne le } C \text{ attendu et on peut conclure. } \square$$

Terminons ce paragraphe en étudiant la multiplicité des racines des polynômes.

Proposition 3. *On suppose que le corps K est un sous-corps de $\mathbb{C}$.*

1) *Si P est un polynôme de $K[X]$ admettant a comme racine de multiplicité $m \geq 1$, alors P' admet a comme racine de multiplicité $m - 1$ et, plus généralement, pour $k = 1, \dots, m - 1$, $P^{(k)}$ admet a comme racine de multiplicité $m - k$ et $P^{(m)}(a) \neq 0$.*

2) *Réciproquement, si $P(a) = P'(a) = \dots = P^{(m-1)}(a) = 0$ et si $P^{(m)}(a) \neq 0$, alors a est racine de multiplicité m de P .*

Démonstration. 1) On a $P = (X - a)^m Q$ avec $Q(a) \neq 0$ d'où $P' = m(X - a)^{m-1} Q + (X - a)^m Q' = (X - a)^{m-1} [mQ + (X - a)Q']$. Posons $Q_1 = mQ + (X - a)Q'$; comme $Q_1(a) = mQ(a) \neq 0$, on voit que a est une racine de multiplicité $m - 1$ de P' . En appliquant ce résultat à P' , on voit que a est une racine de multiplicité $m - 2$ de P'' , etc.

2) Raisonnons par récurrence. Le résultat est vrai pour $m = 0$. Supposons-le vrai pour un entier m et soit P un polynôme tel que $P(a) = P'(a) = \dots = P^{(m)}(a) = 0$ et $P^{(m+1)}(a) \neq 0$. L'hypothèse de récurrence montre que P' s'écrit $P' = (X - a)^m P_1$ avec $P_1(a) \neq 0$. Notons r la multiplicité de a comme racine de P . On a donc $P(X) = (X - a)^r P_2$ avec $P_2(a) \neq 0$; d'où $P'(X) = (X - a)^{r-1} [rP_2 + (X - a)P_2']$. La comparaison des deux expressions de P' montre que $r = m + 1$. $\square$

Par exemple, si P admet a comme racine de multiplicité 3, P' et P'' admettent a comme racine de multiplicité 2 et 1 respectivement et $P'''(a) \neq 0$.

Proposition 4. *On suppose que le corps K est un sous-corps de $\mathbb{C}$. Un polynôme P de $K[X]$ a une racine multiple dans $\mathbb{C}$ si et seulement si $D = \text{pgcd}(P; P')$ est non constant.*

Démonstration. Si P a une racine multiple a dans $\mathbb{C}$, $X - a$ est un facteur commun de P et de P' donc divise D . Réciproquement, si D n'est pas constant, le théorème de d'Alembert montre qu'il possède une racine a et on a $P(a) = P'(a) = 0$, ce qui prouve que a est une racine de multiplicité ≥ 2 de P . $\square$

13.7 ÉLÉMENTS IRRÉDUCTIBLES

Définition : éléments irréductibles. Un élément a d'un anneau intègre A est appelé irréductible s'il n'est pas inversible et si une relation de la forme $a = bc$ implique que b ou c est inversible.

On remarque que 0 n'est pas irréductible car $0 = 0 \times 0$ et 0 n'est pas inversible.

Les éléments irréductibles de $\mathbb{Z}$ sont de la forme $\pm p$ avec p premier.

L'ensemble des éléments irréductibles de $K[X]$ dépend du corps K ; nous allons l'étudier pour $K = \mathbb{R}$ et pour $K = \mathbb{C}$.

Proposition 1 : polynômes irréductibles de $\mathbb{C}[X]$. *Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes du premier degré.*

Démonstration. On a vu que les éléments inversibles de $K[X]$ sont les polynômes constants non nuls. Comme un polynôme du premier degré de $K[X]$ est clairement irréductible pour tout corps K , il faut examiner la réciproque.

Soit A un polynôme de $\mathbb{C}[X]$ de degré $n \geq 2$. D'après le théorème de d'Alembert, A a une racine a dans $\mathbb{C}$ donc $A = (X - a)B$ avec $X - a$ et B , de degré ≥ 1 , non inversibles dans $\mathbb{C}[X]$; A n'est donc pas irréductible dans $\mathbb{C}[X]$. $\square$

Lemme. *Soit A un polynôme de $\mathbb{R}[X]$ admettant une racine a complexe non réelle. Alors A admet aussi $\bar{a}$ comme racine.*

Démonstration. Si $A(X) = \sum_{0 \leq k \leq n} a_k X^k$, on a $A(\bar{a}) = \sum_{0 \leq k \leq n} a_k \bar{a}^k = \overline{A(a)}$ car les a_k sont réels, donc $A(\bar{a}) = 0$. $\square$

Proposition 2 : polynômes irréductibles de $\mathbb{R}[X]$. *Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes du premier degré et ceux du second degré sans racine réelle.*

Démonstration. Les polynômes de $\mathbb{R}[X]$ indiqués sont irréductibles : c'est clair pour les polynômes du premier degré. D'autre part, si un polynôme du second degré de $\mathbb{R}[X]$ est réductible dans $\mathbb{R}[X]$, il se décompose en un produit de deux polynômes du premier degré dans $\mathbb{R}[X]$, donc a deux racines réelles.

Examinons la réciproque.

Soit A un polynôme irréductible de $\mathbb{R}[X]$ de degré $n \geq 2$. D'après le théorème de d'Alembert, A a une racine a dans $\mathbb{C}$.

- Si a est réel, $A = (X - a)B$ avec $X - a$ et B non inversibles dans $\mathbb{R}[X]$; A n'est donc pas irréductible dans $\mathbb{R}[X]$.
- Si a est non réel, le conjugué $\bar{a}$ de a est une racine de A distincte de a donc A est divisible par $(X - a)(X - \bar{a}) = X^2 - 2\operatorname{Re}(a)X + |a|^2$ qui est irréductible sur $\mathbb{R}$; donc $A = u(X^2 - 2\operatorname{Re}(a)X + |a|^2)$ avec u réel non nul. $\square$

Pour continuer le parallélisme avec le chapitre précédent, nous énonçons (la démonstration, dans un cadre plus général, sera donnée au chapitre 20) une propriété analogue à celle sur la décomposition d'un entier en produit de facteurs premiers.

Proposition 3 : décomposition en produit de facteurs irréductibles. *Soit K un corps et soit E l'ensemble des polynômes unitaires irréductibles de $K[X]$. Alors, tout polynôme P de $K[X]$ s'écrit de manière unique :*

$$P = u \prod_{p \in E} p^{k_p}$$

où u est un élément de K^* et où les k_p sont des entiers nuls sauf un nombre fini d'entre eux.

Par exemple, le polynôme $2X^7 - 8X^6 - 18X^5 + 70X^4 + 80X^3 - 126X^2 - 216X - 216$ s'écrit : $2(X - 3)^3(X + 2)^2(X^2 + X + 1)$ comme produit de polynômes irréductibles de $\mathbb{R}[X]$ et $2(X - 3)^3(X + 2)^2(X - j)(X - j^2)$ comme produit de polynômes irréductibles de $\mathbb{C}[X]$.

13.8 LA STRUCTURE DE K -ALGÈBRE DE $K[X]$

Étant donné un corps K , nous avons rencontré à plusieurs reprises une structure que nous n'avons pas encore étudiée pour elle-même : la structure de K -algèbre.

Définition 1 : K -algèbre. Un ensemble A est dit muni d'une structure de K -algèbre s'il existe :

- 1) une loi notée additivement $+$: $A \times A \rightarrow A$;
- 2) une loi notée multiplicativement $\times$: $A \times A \rightarrow A$;
- 3) une loi de multiplication par les scalaires : $K \times A \rightarrow A$.

L'ensemble A a une structure d'anneau avec les deux premières lois ; avec la première et la troisième loi, il a une structure de K -espace vectoriel. On demande en outre que la troisième loi soit compatible avec la seconde de la façon suivante : pour tout x, y de A et tout a de K , on a $a(xy) = (ax)y = x(ay)$.

Nous avons rencontré plusieurs exemples de structures de K -algèbres.

Exemples.

- 1) Dans ce chapitre, la K -algèbre commutative $K[X]$ des polynômes à coefficients dans K .
- 2) Dans le chapitre 7, la K -algèbre $L(E, E)$ des endomorphismes d'un espace vectoriel E . La loi notée multiplicativement est ici la loi de composition des applications linéaires : elle n'est pas commutative.

- 3) Dans le chapitre 8, la K -algèbre $M_n(K)$ des matrices carrées d'ordre n à coefficients dans K . La loi notée multiplicativement est ici la loi de multiplication des matrices : elle n'est pas commutative.
- 4) Si K est un corps, l'ensemble $\text{Fonct}(K, K)$ des applications de K dans lui-même est une K -algèbre commutative pour les opérations d'addition des fonctions, de multiplication des fonctions, de multiplication par un scalaire des fonctions.

Définition 2 : homomorphisme de K -algèbres. Étant données deux K -algèbres A et A' , on appelle homomorphisme de K -algèbres une application linéaire $f : A \rightarrow A'$ qui est aussi un homomorphisme d'anneaux.

Proposition : propriété universelle de $K[X]$ en tant que K -algèbre. Soient A une K -algèbre et x un élément de A . Il existe un homomorphisme de K -algèbres unique $\varphi : K[X] \rightarrow A$ tel que $\varphi(X) = x$.

Démonstration. Soit $P = \sum_{0 \leq k \leq n} a_k X^k$ un polynôme de $K[X]$. Si φ existe, on a nécessairement $\varphi(P) = \varphi(\sum_{0 \leq k \leq n} a_k X^k) = \sum_{0 \leq k \leq n} a_k \varphi(X)^k = \sum_{0 \leq k \leq n} a_k x^k$. Il reste à vérifier que cette définition de φ donne bien un homomorphisme, c'est-à-dire que φ est une application linéaire et que $\varphi(P_1 P_2) = \varphi(P_1) \varphi(P_2)$ pour tous polynômes P_1, P_2 de $K[X]$. $\square$

Cette propriété universelle indique que, pour définir un homomorphisme φ de K -algèbres de source $K[X]$, il suffit de se donner la valeur de $\varphi(X)$. Elle va nous permettre de définir sans effort des homomorphismes. En voici quatre exemples : les deux premiers, que nous avons déjà utilisés, le troisième que nous utiliserons dans le chapitre 15.

Exemple 1 : évaluation. Soit a un élément de K . L'évaluation en a que nous avons défini en 13.2 peut être définie avec cette propriété universelle : on définit $\text{ev}_a : K[X] \rightarrow K$ par $\text{ev}_a(X) = a$. Les propriétés de l'évaluation données en 13.2 sont celles de l'homomorphisme ainsi défini.

Exemple 2 : fonction polynomiale associée. L'homomorphisme $\varphi : K[X] \rightarrow \text{Fonct}(K, K)$ défini par $\varphi(X) = \text{id}_K$ associe à un polynôme $P = \sum_{0 \leq k \leq n} a_k X^k$ la fonction polynomiale $x \mapsto \sum_{0 \leq k \leq n} a_k x^k$. Les propriétés de l'homomorphisme φ expriment que la fonction polynomiale associée à la somme de deux polynômes est la somme des fonctions polynomiales associées à chacun d'eux et des propriétés semblables pour le produit de deux polynômes ou la multiplication d'un polynôme par un scalaire.

Exemple 3 : polynômes d'endomorphismes ou de matrices. Soient E un K -espace vectoriel et $f : E \rightarrow E$ un endomorphisme de E . La propriété universelle donne un homomorphisme $\varphi : K[X] \rightarrow L(E, E)$ tel que $\varphi(X) = f$. On écrira $\varphi(P) = P(f)$. Par exemple, si $P(X) = X^2$ on a $\varphi(P) = P(f) = f \circ f = f^2$ et si $P(X) = 5X^3 - 3X^2 + 5X - 1$, on a $\varphi(P) = 5f^3 - 3f^2 + 5f - \text{id}$.

Les propriétés de l'homomorphisme φ s'écrivent alors :

a) $P_1(f) + P_2(f) = (P_1 + P_2)(f)$;

b) $P_1(f) \circ P_2(f) = (P_1 P_2)(f)$;

c) $(aP)(f) = a(P(f))$.

Bien que le but de φ ne soit pas une algèbre commutative, on vérifie une propriété de commutativité :

a) $P_1(f) \circ P_2(f) = P_2(f) \circ P_1(f)$

On définit de même des expressions polynomiales d'une matrice.

Exemple 4 : substitution. On a souvent besoin de substituer un polynôme S à X dans des polynômes de $K[X]$. Cette substitution est une application de la propriété universelle de $K[X]$ en considérant l'homomorphisme $\varphi : K[X] \rightarrow K[X]$ défini par $\varphi(X) = S$. Par exemple, si $S = 2X - 3$ et $P = X^3 - 3X + 1$, on a $P(2X - 3) = (2X - 3)^3 - 3(2X - 3) + 1$; plus généralement, si $P = \sum_{0 \leq k \leq n} a_k X^k$, on a $P(S) = \sum_{0 \leq k \leq n} a_k S^k$. On ne confondra pas $P(2X - 3)$ avec le produit de P par $2X - 3$.

Les propriétés de l'homomorphisme φ s'écrivent comme pour l'exemple précédent :

a) $P_1(S) + P_2(S) = (P_1 + P_2)(S)$;

b) $P_1(S)P_2(S) = (P_1 P_2)(S)$;

c) $(aP)(S) = a(P(S))$.

La fonction polynomiale associée à $P(S)$ est la fonction composée $x \mapsto P(S(x))$.

➤ Vers le chapitre 14

Les chapitres 11, 12, 13 ont mis en place des outils d'algèbre qui vont nous servir pour développer de nouveaux chapitres d'algèbre linéaire : la théorie des déterminants et la réduction des endomorphismes d'un espace vectoriel de dimension finie.

Exercices

13.1 Division euclidienne

- a) Soient K un corps, et a et b deux éléments distincts de K . Calculer le reste de la division euclidienne d'un polynôme P de $K[X]$ par $(X - a)(X - b)$ en fonction de $P(a)$ et de $P(b)$.
- b) Calculer le reste de la division euclidienne de $(\cos a + X \sin a)^n$ par $X^2 + 1$ dans $\mathbb{R}[X]$.

13.2 Divisibilité

- a) Pour tout entier n , on pose $P_n(X) = 1 + \sum_{0 \leq k \leq n} \frac{X(X+1)\dots(X+k)}{(k+1)!}$.

Factoriser $P_n(X)$ sur $\mathbb{R}$ (on commencera par les petites valeurs de n).

- b) Pour quelles valeurs de l'entier n le polynôme $A = (X + 1)^n - X^n - 1$ de $\mathbb{R}[X]$ est-il divisible par $X^2 + X + 1$?
- c) Pour quelles valeurs de l'entier n le polynôme $A = (X^2 - X + 1)^n - X^{2n} + X^n - 1$ de $\mathbb{R}[X]$ est-il divisible par $X^3 - X^2 + X - 1$?

13.3 pgcd

- 1) On considère les polynômes $A = X^4 + X^3 + X + 1$ et $B = X^3 + X^2 + X + 1$ de $\mathbb{R}[X]$.
- a) Calculer $D = \text{pgcd}(A, B)$.
- b) Trouver des polynômes U et V de $\mathbb{R}[X]$ tels que $UA + VB = D$.
- c) Décomposer A et B en produit de facteurs irréductibles dans $\mathbb{R}[X]$ et dans $\mathbb{C}[X]$.
- 2) Trouver le ou les polynômes P de degré ≤ 5 tels que $(X - 1)^3$ divise $P(X) + 1$, $(X + 1)^3$ divise $P(X) - 1$, en utilisant le polynôme dérivé P' .

13.4 Multiplicité

Soient K un corps, et P et Q deux polynômes de $K[X]$. On suppose que a est une racine commune de P et Q dans K de multiplicité r dans P et s dans Q . Que peut-on dire de la multiplicité de a comme racine de :

- a) $P + Q$;
- b) $P^2 + Q^3$;
- c) $P^2 Q^3$?

13.5 Irréductibilité sur $\mathbb{Q}$ et sur $\mathbb{R}$

Irréductibilité sur $\mathbb{Q}$. Soit $P = a_0 + a_1 X + \dots + a_n X^n$ un polynôme à coefficients entiers.

1) Montrer que si P admet une racine rationnelle $\frac{p}{q}$ avec p et q entiers et $\text{pgcd}(p, q) = 1$, alors p divise a_0 et q divise a_n .

2) Factoriser en produit de facteurs irréductibles dans $\mathbb{Q}[X]$ les polynômes (les calculs sont assez longs) :

a) $P_1 = 4X^4 - 28X^3 + 45X^2 - 6X - 18$;

b) $P_2 = 5X^3 - 4X^2 + 6$.

Irréductibilité sur $\mathbb{R}$

3) Étudier l'irréductibilité sur $\mathbb{R}$ des polynômes :

a) P_1 et P_2 ci-dessus ;

b) $X^4 + 1$, $X^4 + X^2 + 1$.

13.6 Extension des polynômes de Lagrange

Soient K un corps et $a, b, c, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2$ des éléments de K avec a, b, c distincts deux à deux. On cherche l'ensemble E des polynômes P de $K[X]$ tels que :

$$\begin{aligned} P(a) &= \alpha_1 & P(b) &= \beta_1 & P(c) &= \gamma_1 \\ P'(a) &= \alpha_2 & P'(b) &= \beta_2 & P'(c) &= \gamma_2 \\ P''(a) &= \alpha_3 & P''(b) &= \beta_3 \end{aligned}$$

a) Montrer que le problème se ramène à une équation linéaire de la forme $\varphi(P) = u$ où $\varphi : K[X] \rightarrow K^8$ est une application linéaire et u un élément de K^8 qu'on précisera.

b) Déterminer le noyau de φ .

c) Montrer que $\varphi(P)$ induit un isomorphisme de l'espace $K_7[X]$ des polynômes de $K[X]$ de degré ≤ 7 sur K^8 et en déduire que l'équation $\varphi(P) = u$ a une solution unique S dans $K_7[X]$ (on ne cherchera pas à expliciter S).

d) Donner une description de E .

13.7 Équation fonctionnelle

On pose $j = \exp^{2\pi i/3}$. On rappelle que $j^3 = 1$ et que $j^2 + j + 1 = 0$.

1) Soit $a \in \mathbb{C}$. Quelles sont les valeurs de a pour lesquelles :

a) la suite (a^n) ne prend qu'un nombre fini de valeurs ;

b) les suites (a^n) et $(a - 1)^n$ ne prennent qu'un nombre fini de valeurs ?

2) On cherche l'ensemble E des polynômes P de $\mathbb{C}[X]$ satisfaisant la relation :

$$P(X^2) = P(X)P(X+1)$$

a) Montrer que si a est une racine de P , alors a^{2^n} et $(a-1)^{2^n}$ sont des racines de P pour tout entier strictement positif n .

b) En déduire $a \in \{0, 1, -j, -j^2\}$ puis $a \in \{0, 1\}$.

3) Déterminer l'ensemble E .

Solutions

13.1 a) La division euclidienne par un polynôme de degré 2 donne un reste de degré 1. Posons $P(X) = (X-a)(X-b)Q(X) + \alpha X + \beta$. Les évaluations en a et en b de cette égalité conduisent à un système linéaire en α et β . On trouve le reste $\frac{P(b)-P(a)}{b-a}X + \frac{bP(a)-aP(b)}{b-a}$.

b) La division euclidienne donne un résultat de la forme : $(\cos a + X \sin a)^n = (X^2 + 1)Q(X) + \alpha X + \beta$. Cette égalité entre polynômes de $\mathbb{R}[X]$ peut être considérée comme une égalité entre polynômes de $\mathbb{C}[X]$. Évaluons cette égalité en ι : on trouve $\cos na + \iota \sin na = \alpha \iota + \beta$, d'où $\alpha = \sin na$ et $\beta = \cos na$.

13.2 a) On montre par récurrence que $P_n(X) = \frac{(X+1)\dots(X+n+1)}{(n+1)!}$.

b) Si A est divisible par $X^2 + X + 1$ dans $\mathbb{C}[X]$, le quotient et le reste sont dans $\mathbb{R}[X]$ comme le montre l'algorithme de division euclidienne, donc A est divisible par $X^2 + X + 1$ dans $\mathbb{R}[X]$. Les racines de $X^2 + X + 1$ sont $j = \exp^{2\pi i/3}$ et $j^2 = \exp^{-2\pi i/3}$. Si $A(j) = 0$, le polynôme A , étant à coefficients réels, admet le conjugué j^2 de j comme racine. La condition de divisibilité équivaut donc à l'égalité $A(j) = 0$ soit à $(j+1)^n - j^n - 1 = 0$, donc à $(-1)^{n+1}j^{2n} + j^n + 1 = 0$.

En raisonnant sur n modulo 6, on trouve $n = \pm 1 \pmod{6}$.

c) Comme $X^3 - X^2 + X - 1 = (X^2 + 1)(X - 1)$ et que $A(1) = 0$, la condition équivaut à $A(\iota) = 0$. En raisonnant sur n modulo 4, on voit que $n = 0, 1, 3 \pmod{4}$ conviennent et que $n = 2 \pmod{4}$ ne convient pas.

13.3 1) a) On remarque d'abord que A et B sont divisibles par $X + 1$, ce qui ramène à la recherche du pgcd de $X^3 + 1$ et de $X^2 + 1$. L'algorithme donne $X^3 + 1 = (X^2 + 1)X + (-X + 1)$, $X^2 + 1 = (-X + 1)(-X - 1) + 2$, donc $D = u(X + 1)$ avec u réel non nul qu'on prendra égal à 1.

b) On trouve $U = \frac{X+1}{2}$ et $V = \frac{-X^2 - X + 1}{2}$ en remontant les calculs.

c) On trouve $A = (X+1)^2(X^2 - X + 1)$, $B = (X+1)(X^2 + 1)$ dans $\mathbb{R}[X]$ et $A = (X+1)^2(X+j)(X+j^2)$, $B = (X+1)(X+i)(X-i)$ dans $\mathbb{C}[X]$.

2) Comme P' est divisible par $(X-1)^2$ et par $(X+1)^2$, il est divisible par leur produit ; on a donc : $P' = a(X^4 - 2X^2 + 1)$ donc $P = a(\frac{X^5}{5} - \frac{2}{3}X^3 + X) + b$. Les deux conditions de divisibilité donnent $b = 0$ et $a = -\frac{15}{8}$.

On peut aussi utiliser l'identité de Bézout : on remarque que les deux conditions de la question demandent l'existence de polynômes A et B tels que $P = -1 + (X-1)^3$

$A = 1 + (X+1)^3 B$, d'où $(X-1)^3 \frac{A}{2} - (X+1)^3 \frac{B}{2} = 1$; l'algorithme de recherche des

U et V de l'identité de Bézout est un peu long, mais on retrouve le même résultat.

13.4 a) On a $m(P+Q, a) \geq \inf(r, s)$. On a l'égalité si $r \neq s$; si $r = s$ on peut avoir l'inégalité stricte : prendre $P = X(X-a)$, $Q = -a(X-a)$ par exemple.

b) $m(P^2 + Q^3, a) \geq \inf(2r, 3s)$ avec l'égalité si $2r \neq 3s$.

c) $m(P^2 Q^3, a) = 2r + 3s$.

13.5 1) Si $\frac{p}{q}$ est racine de P , on a $a_0 + a_1 \frac{p}{q} + \dots + a_n \frac{p^n}{q^n} = 0$, donc :

a) $q(a_0 q^{n-1} + a_1 p q^{n-2} + \dots + a_{n-1} p^{n-1}) = -a_n p^n$;

c) $p(a_1 q^{n-1} + \dots + a_n p^{n-1}) = -a_0 q^n$

Comme p et q sont premiers entre eux, on déduit de la première égalité que q divise a_n et de la seconde que p divise a_0 .

2) a) La propriété précédente permet de limiter les essais. Si $\frac{p}{q}$ est racine de P_1 , on voit que p divise 18 et que q divise 4. Les seules racines rationnelles possibles de P_1 sont donc :

$$\pm 18 ; \pm 9, \pm 6, \pm 3, \pm 2, \pm 1, \pm \frac{1}{2}, \pm \frac{1}{4}, \pm \frac{3}{2}, \pm \frac{3}{4}, \pm \frac{9}{2}, \pm \frac{9}{4}.$$

On trouve que $-\frac{1}{2}$ et $\frac{3}{2}$ sont racines, ce qui conduit après division par $(X + \frac{1}{2})(X - \frac{3}{2})$ à : $P_1 = (X + \frac{1}{2})(X - \frac{3}{2})(X^2 - 6X + 6)$. Ce dernier facteur n'admettant

aucun des rationnels précédents comme racine est irréductible sur $\mathbb{Q}$ (on peut aussi calculer les racines de $X^2 - 6X + 6$ qui ne sont pas rationnelles).

b) Si P_2 est réductible sur $\mathbb{Q}$, il a un facteur du premier degré, donc une racine rationnelle. Les seules possibilités sont :

$\pm 6, \pm 3, \pm 2, \pm 1, \pm \frac{6}{5}, \pm \frac{3}{5}, \pm \frac{2}{5}, \pm \frac{1}{5}$. Comme aucun de ces nombres ne convient,

P_2 est irréductible sur $\mathbb{Q}$.

3) a) $P_1 = (X + \frac{1}{2})(X - \frac{3}{2})(X - (3 + \sqrt{3}))(X - (3 - \sqrt{3}))$.

b) L'étude des variations de P_2 montre qu'il n'a qu'une racine réelle. Si on la note a , on a $P_2 = (X - a)(5X^2 + (5a - 4)X + 5a^2 - 4a)$.

c) Les nombres complexes tels que $z^4 = -1 = e^{i\pi}$ sont $\pm e^{i\pi/4}, \pm e^{3i\pi/4}$. On a :
 $(X - e^{i\pi/4})(X + e^{i\pi/4}) = X^2 - \sqrt{2}X + 1$, $(X - e^{3i\pi/4})(X + e^{3i\pi/4}) = X^2 + \sqrt{2}X + 1$,
 d'où la factorisation de $X^4 + 1$.

d) On a : $X^4 + X^2 + 1 = (X^2 + 1)^2 - X^2 = (X^2 + X + 1)(X^2 - X + 1)$.

En général, on ne pourra factoriser si aisément !

13.6 a) On définit $\varphi : K[X] \rightarrow K^8$ par :

$$\varphi(P) = (P(a), P'(a), P''(a), P(b), P'(b), P''(b), P(c), P'(c))$$

Le problème se ramène à la résolution de l'équation linéaire :

$$\varphi(P) = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2). \quad (R)$$

b) Le noyau de φ est l'ensemble des polynômes qui admettent a et b comme racines triples et c comme racine double. C'est donc l'idéal des multiples de $(X - a)^3(X - b)^3(X - c)^2$.

c) Notons f la restriction de φ à $K_7[X]$. Le 2) montre que $\ker(f)$ est nul, donc f est injective, donc f est bijective à cause des dimensions, ce qui prouve l'existence d'un unique polynôme S de degré ≤ 7 vérifiant la relation (R).

d) $E = \{S + A(X - a)^3(X - b)^3(X - c)^2 \mid A \in K[X]\}$. On pourrait chercher à décrire les polynômes de $K_7[X]$ dont les images par f sont les vecteurs de la base canonique de K^8 . Certains sont faciles à trouver, d'autres moins. Par exemple, on cherchera les polynômes d'images e_1, e_2, e_3 sous les formes respectives
 $[\alpha + \beta(X - a) + \gamma(X - a)^2](X - b)^3(X - c)^2$,
 $[\beta(X - a) + \gamma(X - a)^2](X - b)^3(X - c)^2$,
 $[\gamma(X - a)^2](X - b)^3(X - c)^2$.

13.7 1) a) $a = 0$ convient. Si $|a| \neq 1$, la suite $|a|^n$ est strictement croissante ou strictement décroissante et elle prend une infinité de valeurs. Si $|a| = 1$ et si la suite (a^n) prend un nombre fini de valeurs, a est d'ordre fini dans le groupe multiplicatif des nombres complexes de module 1 ; par conséquent, il existe n tel que $a^n = 1$ (on dit que a est une racine de l'unité).

b) $a = 0$ convient, ainsi que $a = 1$. Sinon, on doit avoir $|a| = |a - 1| = 1$; l'intersection des deux cercles donne $a = e^{i\pi/3} = -j$ ou $a = e^{-i\pi/3} = -j^2$.

2) Si a est racine de P , on a $P(a^2) = P(a)P(a + 1)$ donc a^2 est racine de P ; par suite, $a^4, a^8, \dots$ sont racines de P ; on a aussi $P((a - 1)^2) = P(a - 1)P(a) = 0$ donc $(a - 1)^2, (a - 1)^4, \dots$ sont racines de P . Or P n'a qu'un nombre fini de racines. Le même raisonnement qu'au **1)** montre que a doit appartenir à $\{0, 1, -j, -j^2\}$.

Si $-j$ (resp : $-j^2$) était racine de P , l'équation fonctionnelle montre que j^2 (resp : j) serait racine de P , ce qui est impossible.

3) Si P satisfait la relation, P est donc de la forme $uX^k(X - 1)^l$ et on doit avoir : $uX^{2k}(X^2 - 1)^l = u^2 X^k (X - 1)^l (X + 1)^k X^l = u^2 X^{k+l} (X - 1)^l (X + 1)^k$.

La comparaison des deux décompositions en produits de facteurs irréductibles donne $u = 0$ ou $u = 1$ et $k = l$; les polynômes satisfaisant la relation sont donc le polynôme nul et les polynômes de la forme $X^k(X - 1)^k$.

14.1 INTRODUCTION HISTORIQUE

Ce chapitre est difficile. Ce n'est peut-être pas étonnant ; il a fallu bien plus de 100 ans aux mathématiciens pour réussir à donner un exposé satisfaisant de la théorie des déterminants.

Dans les premiers paragraphes de ce chapitre, nous allons aborder les déterminants en suivant les idées des premiers mathématiciens qui ont défriché ce sujet, ce qui permettra de mieux comprendre, nous l'espérons, la suite.

Cramer en 1750. En 1750, Cramer est le premier à donner une règle générale pour calculer les solutions d'un système linéaire de n équations à n inconnues. Nous avons déjà dit (voir 4.2) que sa notation se présente à peu de chose près comme une notation indicielle. Cramer donne les cas de systèmes à une, deux, trois inconnues. Il induit alors de l'examen de ses formules une règle générale qu'il énonce mais ne démontre pas. Cette règle est exacte, mais elle pose des problèmes et donne donc du travail à ses successeurs¹.

Les formules de Cramer. Nous utilisons des doubles indices pour avoir une notation qui s'adapte au cas général.

Commençons par rappeler le résultat du chapitre 4 pour le cas $n = 2$. Si le système s'écrit :

1. Voir fac-similé du texte aux pages suivantes.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

FACSIMILÉ DE L'INTRODUCTION À L'ANALYSE DES LIGNES COURBES ALGÈBRIQUES, GABRIEL CRAMER (1750)

De l'évanouissement des inconnues.

Quand un Problème renferme plusieurs inconnues, dont les relations sont tellement compliquées qu'on se trouve obligé de former plusieurs équations; alors, pour découvrir les valeurs de ces inconnues, on les fait toutes évanouir, moins une, qui combinée seule avec les grandeurs connues donne, si le Problème est déterminé, une *Equation finale*, dont la résolution fait connoître d'abord cette inconnue, & ensuite par son moyen toutes les autres.

Soient plusieurs inconnues $z, y, x, v, \text{\textit{etc.}}$ & autant d'équations

$$\begin{aligned} A^1 &= Z^1z + T^1y + X^1x + V^1v + \text{\textit{etc.}} \\ A^2 &= Z^2z + T^2y + X^2x + V^2v + \text{\textit{etc.}} \\ A^3 &= Z^3z + T^3y + X^3x + V^3v + \text{\textit{etc.}} \\ A^4 &= Z^4z + T^4y + X^4x + V^4v + \text{\textit{etc.}} \end{aligned}$$

où les lettres $A^1, A^2, A^3, A^4, \text{\textit{etc.}}$ ne marquent pas, comme à l'ordinaire, les puissances d' A , mais le premier membre, supposé connu, de la première, seconde, troisième, quatrième &c. équation. De même $Z^1, Z^2, \text{\textit{etc.}}$ sont les coefficients de z ; $T^1, T^2, \text{\textit{etc.}}$ ceux de y ; $X^1, X^2, \text{\textit{etc.}}$ ceux de x ; $V^1, V^2, \text{\textit{etc.}}$ ceux de v ; &c. dans la première, seconde, &c. équation.

Cette Notation supposée, s'il n'y a qu'une équation & qu'une inconnue z ; on aura $z = \frac{A^1}{Z^1}$. S'il y a deux équations & deux inconnues z & y ; on trouvera $z = \frac{A^1T^2 - A^2T^1}{Z^1T^2 - Z^2T^1}$, & $y = \frac{Z^1A^2 - Z^2A^1}{Z^1T^2 - Z^2T^1}$. S'il y a trois équations & trois inconnues $z, y, \text{ \& } x$; on trouvera

$$\begin{aligned} z &= \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1} \\ y &= \frac{Z^1A^2X^3 - Z^1A^3X^2 - Z^2A^1X^3 + Z^2A^3X^1 + Z^3A^1X^2 - Z^3A^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1} \\ x &= \frac{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1} \end{aligned}$$

Introd. à l'Analyse des Lignes Courbes. Oooo L'é-

L'examen de ces Formules fournit cette Règle générale. Le nombre des équations & des inconnues étant n , on trouvera la valeur de chaque inconnue en formant n fractions dont le dénominateur commun a autant de termes qu'il y a de divers arrangements de n choses différentes. Chaque terme est composé des lettres $ZYXV$ &c. toujours écrites dans le même ordre, mais auxquelles on distribue, comme exposants, les n premiers chiffres rangés en toutes les manières possibles. Ainsi, lorsqu'on a trois inconnues, le dénominateur a $[1 \times 2 \times 3 =]$ 6 termes, composés des trois lettres ZYX , qui reçoivent successivement les exposants $123, 132, 213, 231, 312, 321$. On donne à ces termes les signes $+$ ou $-$, selon la Règle suivante. Quand un exposant est suivi dans le même terme, médiatement ou immédiatement, d'un exposant plus petit que lui, j'appellerai cela un *dérangement*. Qu'on compte, pour chaque terme, le nombre des dérangements: s'il est pair ou nul, le terme aura le signe $+$; s'il est impair, le terme aura le signe $-$. Par ex. dans le terme $Z^1 Y^2 V^3$ il n'y a aucun dérangement: ce terme aura donc le signe $+$. Le terme $Z^3 Y^1 X^2$ a aussi le signe $+$, parce qu'il a deux dérangements, 3 avant 1 & 3 avant 2. Mais le terme $Z^3 Y^2 X^1$, qui a trois dérangements, 3 avant 2, 3 avant 1, & 2 avant 1, aura le signe $-$.

Le dénominateur commun étant ainsi formé, on aura la valeur de z en donnant à ce dénominateur le numérateur qui se forme en changeant, dans tous ses termes, Z en A . Et la valeur d' y est la fraction qui a le même dénominateur & pour numérateur la quantité qui résulte quand on change Y en A , dans tous les termes du dénominateur. Et on trouve d'une manière semblable la valeur des autres inconnues.

Généralement parlant, le Problème est déterminé. Mais

il peut y avoir des Cas particuliers, où il reste indéterminé; & d'autres où il devient impossible. C'est lorsque le dénominateur commun se trouve égal à zéro; c'est-à-dire, s'il n'y a que deux équations, lorsque $Z^1 Y^2 - Z^2 Y^1 = 0$, s'il y en a trois, lorsque $Z^1 Y^2 X^3 - Z^1 Y^3 X^2 - Z^2 Y^1 X^3 + Z^2 Y^3 X^1 + Z^3 Y^1 X^2 - Z^3 Y^2 X^1 = 0$, &c. Alors, si les grandeurs $A^1, A^2, A^3, \dots$ sont telles que les numérateurs soient aussi égaux à zéro, le Problème est indéterminé; car les fractions $\frac{0}{0}$, qui devoient donner la valeur des inconnues, sont indéterminées. Mais si les grandeurs $A^1, A^2, A^3, \dots$ sont telles que, le dénominateur commun étant zéro, les numérateurs ou quelques-uns d'entr'eux ne soient pas zéro, le Problème est impossible, ou du moins les grandeurs inconnues qui peuvent le résoudre sont toutes, ou en partie, infinies. Par ex. si l'on a ces deux équations $2 = 3z - 2y$ & $5 = 6z - 4y$, on trouvera $z = \frac{2}{5}$ & $y = \frac{3}{5}$. Donc z & y sont des grandeurs infinies, qui sont l'une à l'autre en raison de 2 à 3. En dégageant les inconnues par les Méthodes ordinaires, on tomberoit dans cette équation absurde $\frac{2}{3} = \frac{5}{2}$. Car la première équation donne $z = \frac{2}{3}y + \frac{2}{3}$ & la seconde $z = \frac{5}{6}y + \frac{5}{6}$. Donc $\frac{2}{3}y + \frac{2}{3} = \frac{5}{6}y + \frac{5}{6}$, ou $\frac{2}{3} = \frac{5}{6}$: ce qui est absurde, si z & y sont des grandeurs finies. Mais si elles sont infinies, on peut dire sans absurdité que $z = \frac{2}{3}y + \frac{2}{3}$ & en même tems que $z = \frac{5}{6}y + \frac{5}{6}$; parce que les grandeurs finies $\frac{2}{3}$ & $\frac{5}{6}$ n'étant rien en comparaison des grandeurs infinies z & $\frac{2}{3}y$, les deux équations $z = \frac{2}{3}y + \frac{2}{3}$ & $z = \frac{5}{6}y + \frac{5}{6}$ se réduisent toutes deux à $z = \frac{2}{3}y$, qui n'a rien de contradictoire.

Cramer G. – *Introduction à l'analyse des lignes courbes algébriques*, Genève, Frères Cramer & Cl. Philibert, 1750.

on trouve :

$$x_1 = \frac{b_1 a_{22} - b_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}}$$

$$x_2 = \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}}$$

Pour le cas où $n = 3$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{cases}$$

on trouve, après des calculs évidemment un peu plus longs :

$$\begin{aligned}
 x_1 &= \frac{b_1 a_{22} a_{33} + b_2 a_{32} a_{13} + b_3 a_{12} a_{23} - b_3 a_{22} a_{13} - b_1 a_{32} a_{23} - b_2 a_{12} a_{33}}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{11} a_{32} a_{23} - a_{21} a_{12} a_{33}} \\
 x_2 &= \frac{a_{11} b_2 a_{33} + a_{21} b_3 a_{13} + a_{31} b_1 a_{23} - a_{31} b_2 a_{13} - a_{11} b_3 a_{23} - a_{21} b_1 a_{33}}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{11} a_{32} a_{23} - a_{21} a_{12} a_{33}} \\
 x_3 &= \frac{a_{11} a_{22} b_3 + a_{21} a_{32} b_1 + a_{31} a_{12} b_2 - a_{31} a_{22} b_1 - a_{11} a_{32} b_2 - a_{21} a_{12} b_3}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{11} a_{32} a_{23} - a_{21} a_{12} a_{33}}
 \end{aligned}$$

L'examen de ces formules inspire à Cramer un commentaire sur le cas général de n équations à n inconnues : « On trouvera la valeur de chaque inconnue en formant n fractions dont le dénominateur commun a autant de termes qu'il y a de divers arrangements de n choses différentes. Chaque terme est composé des lettres $X Y Z V \&c.$ toujours écrites dans le même ordre, mais auxquelles on distribue comme indices (dans la notation de Cramer, les chiffres sont placés en exposant) les n premiers chiffres rangés de toutes les manières possibles. »

Comme cela risque de ne pas être clair, Cramer revient donc au cas $n = 3$. En premier lieu, x_1 , x_2 et x_3 ont le même dénominateur. Ce dénominateur contient les 6 produits possibles de la forme $a_{i1} a_{j2} a_{k3}$; c'est-à-dire que (i, j, k) peut prendre les 6 valeurs $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$, $(3, 2, 1)$. Cramer appelle *dérangement dans la suite* (i, j, k) le fait qu'un terme soit suivi d'un terme plus petit que lui ; $a_{i1} a_{j2} a_{k3}$ sera précédé du signe + si le nombre de dérangements est pair, du signe - s'il est impair. Par exemple, dans $(1, 2, 3)$ il n'y a pas de dérangements, dans $(3, 1, 2)$ il y a deux dérangements, $3 > 1$, $3 > 2$, donc $a_{11} a_{22} a_{33}$ et $a_{31} a_{12} a_{23}$ seront précédés du signe + ; dans $(3, 2, 1)$ il y a 3 dérangements, $3 > 1$, $3 > 2$, $2 > 1$, donc $a_{31} a_{22} a_{13}$ sera précédé du signe -.

L'autre commentaire de Cramer est que le numérateur de x_j s'obtient à partir du dénominateur en changeant les a_{ij} en b_i , ce qui se constate facilement.

La formule générale de Cramer. Quelle est la forme générale des dénominateurs des formules de Cramer ? Pour $n = 3$, chaque terme est de la forme $a_{i1} a_{j2} a_{k3}$. Il est donc défini par la permutation σ de l'ensemble $\{1, 2, 3\}$ donnée par $\sigma(1) = i$, $\sigma(2) = j$, $\sigma(3) = k$. Le signe dont il sera affecté est + ou - suivant le nombre de ce que Cramer appelle des dérangements. Ces dérangements correspondent aux couples (i, j) tels que $i < j$ et $\sigma(i) > \sigma(j)$; on retrouve la définition de la signature d'une permutation.

Dans le cas général, le dénominateur comportera un nombre de termes égal au nombre de divers arrangements de n choses, c'est-à-dire $n!$. Chaque terme sera de la forme : $a_{\sigma(1)1} \dots a_{\sigma(n)n}$ et son signe sera $\varepsilon(\sigma)$. Le dénominateur des formules de Cramer peut donc s'écrire :

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$$

Déterminant d'une matrice carrée. Cramer considérait des systèmes linéaires avec autant d'équations que d'inconnues, ce qui revient à considérer une matrice carrée $A = (a_{ij})$ d'ordre n . Le déterminant de A est défini par la formule de Cramer ci-dessus. Nous le noterons $\text{Dét}(A)$ (avec un D majuscule).

Si A est une matrice donnée par ses coefficients, par exemple :

$$A = \begin{pmatrix} 2 & 1 & 0 & -1 \\ 1 & 3 & 3 & 4 \\ 0 & 2 & -5 & 1 \\ 4 & -1 & 2 & 0 \end{pmatrix}$$

nous noterons aussi :

$$\text{Dét}(A) = \begin{vmatrix} 2 & 1 & 0 & -1 \\ 1 & 3 & 3 & 4 \\ 0 & 2 & -5 & 1 \\ 4 & -1 & 2 & 0 \end{vmatrix}$$

Cette notation avec des barres verticales est due à Cayley, qui l'a utilisée en 1841. Elle est proche de celle qu'il proposera pour les matrices (voir 8.3).

Le but des paragraphes qui suivent est de donner des formules et des méthodes permettant le calcul des déterminants plus rapidement qu'avec la méthode de Cramer.

14.2 CALCUL DES DÉTERMINANTS : MÉTHODE DE BÉZOUT

Étienne Bézout (1730-1783). Étienne Bézout est né à Nemours en 1730 dans une famille de magistrats. La lecture des travaux d'Euler l'enthousiasme et il se consacre aux mathématiques. Le Duc de Choiseul, ministre de Louis XV, lui propose en 1763 de former les officiers de la marine. En 1768, il est également nommé responsable des officiers de l'artillerie. Pour ces officiers, il écrit des cours de mathématiques d'un style nouveau, cherchant à se mettre réellement à la portée de non-mathématiciens. Le succès de ces livres et de leurs traductions en anglais en feront les ouvrages de référence pendant une cinquantaine d'années en France, en Angleterre et aux États-Unis. D'autres travaux de Bézout sont également remarquables.

En 1764, Bézout reprend le travail de Cramer, parce que l'usage de la formule de Cramer lui « a fait connoître que quoiqu'elle soit assez simple, quant aux lettres, elle ne l'est pas de même à l'égard des signes lorsqu'on a au-delà d'un certain nombre d'inconnues à calculer ». Bézout remarque qu'on peut écrire le dénominateur pour $n = 3$ en faisant apparaître trois dénominateurs du cas $n = 2$:

$$\begin{aligned} & a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} \\ &= (a_{11}a_{22} - a_{21}a_{12})a_{33} + (a_{31}a_{12} - a_{11}a_{32})a_{23} + (a_{21}a_{32} - a_{31}a_{22})a_{31} \end{aligned}$$

ce qui lui permet de voir une loi de formation par récurrence des expressions des dénominateurs.

Matrice extraite A_{ij} . Nous utiliserons la notation suivante : étant donné une matrice carrée $A = (a_{ij})$ d'ordre n et un couple (i, j) d'entiers, $1 \leq i, j \leq n$, nous noterons A_{ij} la matrice carrée d'ordre $n - 1$ obtenue en supprimant la i -ième ligne et la j -ième colonne de A . Pour la matrice A précédente, on a, par exemple :

$$A_{23} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 4 & -1 & 0 \end{pmatrix}$$

Définition du déterminant par développement suivant la première ligne. Le calcul par récurrence du déterminant proposé par Bézout nous conduit à une seconde définition des déterminants des matrices carrées. Nous noterons $\det(A)$ (avec un d minuscule) le déterminant d'une matrice A ainsi défini et le mot de déterminant sera entendu dans ce sens dans la suite du chapitre. Nous verrons en 14.4 que la définition de Cramer et celle-ci coïncident. Dans cette définition, le déterminant est construit par récurrence sur l'ordre de la matrice $A = (a_{ij})$:

- si $n = 1$ par $\det(A) = a_{11}$
- si $n > 1$ par $\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \dots$

$$+ (-1)^{n+1} a_{1n} \det(A_{1n}) = \sum_{1 \leq j \leq n} (-1)^{1+j} a_{1j} \det(A_{1j}).$$

Quand on applique cette formule de récurrence, on dit qu'on développe le déterminant par rapport à sa première ligne.

Exemples.

- Dans le cas particulier de la matrice carrée A d'ordre 0, on pose $\det(A) = 1$.
- Si $n = 1$: $\det(a_{11}) = a_{11}$.
- Si $n = 2$:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

- Si $n = 3$:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + a_{13} \det(A_{13})$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

qui donne la formule de Cramer comme on l'a déjà vu.

- Pour tout $n \geq 1$, on a $\det I_n = 1$, par récurrence immédiate.
- Voici le calcul du déterminant de la matrice A définie ci-dessus :

$$\begin{vmatrix} 2 & 1 & 0 & -1 \\ 1 & 3 & 3 & 4 \\ 0 & 2 & -5 & 1 \\ 4 & -1 & 2 & 0 \end{vmatrix} = 2 \begin{vmatrix} 3 & 3 & 4 \\ 2 & -5 & 1 \\ -1 & 2 & 0 \end{vmatrix} - \begin{vmatrix} 1 & 3 & 4 \\ 0 & -5 & 1 \\ 4 & 2 & 0 \end{vmatrix} + 0 + \begin{vmatrix} 1 & 3 & 3 \\ 0 & 2 & -5 \\ 4 & -1 & 2 \end{vmatrix}$$

En développant les déterminants d'ordre 3 par rapport à leur première ligne, on trouve $-26 - 90 - 85 = -201$. Refaites les différents calculs pour vous entraîner.

Les résultats de Cramer et Bézout ne pouvaient donner complète satisfaction ; il restait beaucoup à faire pour clarifier la notion de déterminant.

14.3 LE CARACTÈRE ALTERNÉ

Corps de base. Jusqu'ici, les espaces vectoriels que nous avons considérés étaient des $\mathbb{R}$ -espaces vectoriels. Nous avons expliqué que les résultats obtenus n'utilisaient pas de propriétés particulières de $\mathbb{R}$ et valaient pour tout corps. À partir de ce chapitre, nous noterons K le corps de base pour exprimer les résultats en toute généralité ; nous préciserons, quand cela sera nécessaire, ce qui dépend du choix de K .

Déterminant de n vecteurs. Plaçons-nous dans un espace vectoriel E de dimension n muni d'une base $B = (e_1, \dots, e_n)$ et soit $(u_1, \dots, u_n)$ une famille de n vecteurs de E . Notons $A = (a_{ij})$ la matrice carrée d'ordre n dont les colonnes sont formées par les coordonnées des u_i dans la base B . Le déterminant de la matrice A peut être vu comme une quantité définie par les vecteurs u_i : on définit ainsi une fonction $\det_B : E^n \rightarrow K$ en posant $\det_B(u_1, \dots, u_n) = \det(A)$.

On a $\det_B(e_1, \dots, e_n) = \det(I_n) = 1$.

Vandermonde et Laplace. C'est cet aspect du déterminant comme fonction des coordonnées de n vecteurs que découvrent et étudient Alexandre Vandermonde (1735-1796) et Pierre-Simon de Laplace en 1771 (mais ils ne s'expriment pas comme cela et raisonnent à partir des systèmes linéaires). Le travail de Vandermonde est probablement antérieur à celui de Laplace. Vandermonde propose une notation pour ces fonctions et montre, en particulier, qu'elles :

- 1) s'annulent lorsque les coefficients de deux inconnues sont les mêmes dans toutes les équations ;
- 2) sont opposées quand on permute les coefficients de deux inconnues dans toutes les équations.

Laplace appelle ces fonctions *résultantes* (Augustin Cauchy (1789-1857) imposera, en 1812, le nom de *déterminant* déjà utilisé par Gauss ; ses résultats sont proches de ceux de Vandermonde, mais ses arguments sont plus nets).



Augustin-Louis Cauchy (1789-1857)
Belhoste – Cauchy, Paris, Berlin, 1985.

Ce que disent Vandermonde et Laplace peut être traduit avec nos notations par les propriétés suivantes.

Proposition. Reprenons les notations ci-dessus.

1) Soit σ une permutation de S_n . On a :

$$\det_B(u_{\sigma(1)}, \dots, u_{\sigma(n)}) = \varepsilon(\sigma) \det_B(u_1, u_2, \dots, u_n)$$

2) En particulier, le déterminant change de signe quand on échange deux colonnes.

3) Si deux des vecteurs u_i et u_j , $i \neq j$, sont égaux, alors le déterminant est nul.

Démonstration. Rappelons que $\varepsilon(\sigma)$, la signature de la permutation σ , a été défini en 11.7.

1,2) Comme le groupe S_n est engendré par les transpositions de la forme $(k \ k+1)$ d'après l'exercice 11.4, il suffit de montrer la formule quand on échange deux colonnes adjacentes. Notons A' la matrice obtenue en échangeant les colonnes d'indice k et $k+1$ de A . Raisonnons par récurrence sur n .

On a :

$$\det(A') = \begin{vmatrix} a_{11} & \cdots & a_{1,k+1} & a_{1,k} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k+1} & a_{2,k} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

$$\begin{aligned} \text{donc } \det(A') &= a_{11} \det(A'_{11}) - a_{12} \det(A'_{12}) + \cdots + (-1)^{1+k} a_{1,k+1} \det(A'_{1k}) \\ &+ (-1)^{1+k+1} a_{1k} \det(A'_{1,k+1}) + \cdots + (-1)^{1+n} a_{1,n} \det(A'_{1n}) \end{aligned}$$

nous plaçons des virgules entre les indices quand c'est nécessaire pour la clarté). Si $j \neq k, k+1$, l'hypothèse de récurrence montre que $\det(A'_{1j}) = -\det(A_{1j})$, puisque A'_{1j} et A_{1j} ont deux colonnes échangées ; d'autre part, $A'_{1k} = A_{1,k+1}$ et $A'_{1,k+1} = A_{1k}$. Tous les termes de la somme donnant $\det(A)$ par développement suivant la première ligne ont donc changé de signe.

- 3) Posons $\tau = (i j)$. Si $u_i = u_j$, on a $\det_B(u_1, u_2, \dots, u_n) = \det_B(u_{\tau(1)}, \dots, u_{\tau(n)})$. Mais on vient de montrer que ces deux quantités sont opposées. Par conséquent, elles sont nulles. $\square$

14.4 MULTILINÉARITÉ

Définition 1 : forme multilinéaire. Soient E un K -espace vectoriel et p un entier. On appelle forme p -linéaire sur E une application $\varphi : E^p \rightarrow K$ (rappelons que E^p note le produit $E \times \dots \times E$ où E figure p fois) telle que, pour tout $i, 1 \leq i \leq p$, et toute famille $(u_j)_{1 \leq j \leq p, j \neq i}$, l'application

$$u \mapsto \varphi(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_p)$$

soit linéaire.

Par exemple, si $p = 3$, $\varphi : E^3 \rightarrow K$ est une forme 3-linéaire si, pour tout u_1, u_2, u_3 de E les applications $u \mapsto \varphi(u, u_2, u_3)$, $u \mapsto \varphi(u_1, u, u_3)$, $u \mapsto \varphi(u_1, u_2, u)$ sont linéaires. Plus généralement, on parle de formes multilinéaires sans préciser le p .

Exemples. Si $p = 1$, on retrouve la notion de forme linéaire étudiée au chapitre 10. Si $p = 2$, φ est appelée forme bilinéaire.

Proposition 1 : multilinéarité de $\det_B$. Soit B une base de E . L'application $\det_B : E^n \rightarrow K$ est n -linéaire.

Démonstration. Raisonnons par récurrence sur n .

Soit A une matrice carrée d'ordre n dont les colonnes définissent des vecteurs $u_1, \dots, u_n$. On suppose que pour un indice $k, 1 \leq k \leq n$, on a $u_k = u'_k + u''_k$. La formule définissant le déterminant de A s'écrit :

$$\det(A) = \begin{vmatrix} a_{11} & \cdots & a'_{1k} + a''_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a'_{2k} + a''_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Posons :

$$A' = \begin{pmatrix} a_{11} & \cdots & a'_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a'_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad A'' = \begin{pmatrix} a_{11} & \cdots & a''_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a''_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

et raisonnons par récurrence. On a :

$$\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \cdots + (-1)^{1+k} (a'_{1k} + a''_{1k}) \det(A_{1k}) \\ + \cdots + (-1)^{1+n} a_{1n} \det(A_{1n}).$$

L'hypothèse de récurrence montre que : $\det(A_{1j}) = \det(A'_{1j}) + \det(A''_{1j})$ pour $j \neq k$; d'autre part, $a_{1k} = a'_{1k} + a''_{1k}$ et $A_{1k} = A'_{1k} + A''_{1k}$, d'où la compatibilité de $\det_B$ avec la somme pour chaque k . On prouve de même la compatibilité avec le produit par un scalaire pour chaque k .

Attention. On vient de montrer que $\det_B(\lambda u_1, u_2, \dots, u_n) = \lambda \det_B(u_1, u_2, \dots, u_n)$, etc. Par contre, on remarquera que :

$$\det_B(\lambda(u_1, u_2, \dots, u_n)) = \det_B(\lambda u_1, \lambda u_2, \dots, \lambda u_n) = \lambda^n \det_B(u_1, u_2, \dots, u_n).$$

Commentaire. Les mathématiciens depuis Cramer s'étaient concentrés sur les propriétés de symétrie et d'alternance des déterminants. Il faut attendre un petit article d'Eugène Catalan (1814-1894) en 1846 (96 ans après Cramer) pour voir signaler la propriété précédente.

Définition 2 : formes p-linéaires alternées. Une forme p -linéaire $\varphi : E^p \rightarrow E$ est dite alternée si, pour tout $(u_1, \dots, u_p)$ de E^p , $\varphi(u_1, \dots, u_p)$ est nul dès qu'il existe deux indices i et j distincts tels que $u_i = u_j$.

La notion n'a pas de sens pour une forme linéaire. Une forme bilinéaire $\varphi : E^2 \rightarrow K$ est alternée si $\varphi(u, u) = 0$ pour tout u de E .

On a vu au paragraphe précédent que le déterminant est une forme n -linéaire alternée. On va généraliser la proposition vue alors.

Proposition 2. Soit $\varphi : E^p \rightarrow K$ une forme p -linéaire alternée.

1) Pour tout $u_1, \dots, u_p$ de E et toute transposition $(i j)$ de $\{1, \dots, p\}$, on a :

$$\varphi(u_1, \dots, u_i, \dots, u_j, \dots, u_p) = -\varphi(u_1, \dots, u_j, \dots, u_i, \dots, u_p)$$

2) Pour tout $u_1, \dots, u_p$ de E et toute permutation σ de $\{1, \dots, p\}$, on a :

$$\varphi(u_{\sigma(1)}, \dots, u_{\sigma(p)}) = \varepsilon(\sigma) \varphi(u_1, \dots, u_p)$$

Démonstration. 1) Examinons d'abord le cas où $p = 2$. Soient u_1 et u_2 dans E . La bilinéarité de φ donne : $\varphi(u_1 + u_2, u_1 + u_2) = \varphi(u_1, u_1) + \varphi(u_1, u_2) + \varphi(u_2, u_1) + \varphi(u_2, u_2)$; comme φ est alternée, on a donc $0 = \varphi(u_1, u_2) + \varphi(u_2, u_1)$, d'où le résultat. Examinons maintenant le cas où p est quelconque. Pour toute famille $(u_k)_{1 \leq k \leq p, k \neq i, j}$ l'application $(u, v) \mapsto \varphi(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_{j-1}, v, u_{j+1}, \dots, u_p)$ est bilinéaire alternée. Sa valeur en (u, v) est donc l'opposée de sa valeur en (v, u) , d'où la formule.

2) Dans le cas général, on démontre par récurrence qu'un produit de transpositions vérifie la formule. Comme toute permutation de $\{1, \dots, p\}$ est un produit de transpositions, on conclut. $\square$

Proposition 3. Soit $\varphi : E^p \rightarrow K$ une forme p -linéaire alternée.

1) Pour toute famille $(u_1, \dots, u_p)$ de p vecteurs linéairement dépendants de E , on a :

$$\varphi(u_1, \dots, u_p) = 0.$$

2) On ne change pas la valeur de φ en ajoutant à un vecteur une combinaison linéaire des autres vecteurs.

Démonstration. 1) On peut toujours se ramener par la formule précédente au cas où u_1 est une combinaison linéaire des autres vecteurs : $u_1 = \sum_{2 \leq i \leq p} \lambda_i u_i$. On a alors $\varphi(u_1, \dots, u_p) = \varphi\left(\sum_{2 \leq i \leq p} \lambda_i u_i, u_2, \dots, u_p\right) = \sum_{2 \leq i \leq p} \lambda_i \varphi(u_i, u_2, \dots, u_p) = 0$.

2) Soit $v = \sum_{2 \leq i \leq p} \lambda_i u_i$ une combinaison linéaire de $u_2, \dots, u_p$. On a $\varphi(u_1 + v, u_2, \dots, u_p) = \varphi(u_1, u_2, \dots, u_p) + \varphi(v, u_2, \dots, u_p) = \varphi(u_1, u_2, \dots, u_p)$. $\square$

Formes n -linéaires alternées. On suppose désormais que E est un espace vectoriel de dimension n et on étudie les formes n -linéaires alternées sur E .

Proposition 4. Soient E un espace vectoriel de dimension n et $B = (e_1, \dots, e_n)$ une base de E . Soit $\varphi : E^n \rightarrow K$ une forme n -linéaire alternée. On a :

$$\varphi(u_1, u_2, \dots, u_n) = \text{Dét}(u_1, u_2, \dots, u_n) \varphi(e_1, \dots, e_n)$$

Démonstration. Soit $(u_1, \dots, u_n)$ une famille de n vecteurs de E et notons (a_{ij}) la matrice formée par les coordonnées des vecteurs $u_1, \dots, u_n$ dans la base B . On a

donc : $\varphi(u_1, u_2, \dots, u_n) = \varphi\left(\sum_{1 \leq i \leq n} a_{i1} e_i, \dots, \sum_{1 \leq i \leq n} a_{in} e_i\right)$. Si on développe cette expression, les termes où figurent deux fois le même vecteur de base sont nuls puisque φ est alternée. Par conséquent, une fois ces termes nuls retirés, il reste des termes de la forme $a_{\sigma(1)1} \dots a_{\sigma(n)n} \varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)})$ où σ est une permutation de S_n . D'après la proposition 2 ci-dessus, un tel terme est $\varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \varphi(e_1, \dots, e_n)$, d'où la conclusion.

Égalité des deux définitions de déterminants. En prenant $\varphi = \det_B$ dans la proposition 4, on voit que $\det_B(u_1, \dots, u_n) = \text{Dét}(u_1, \dots, u_n)$ pour toute famille $u_1, \dots, u_n$. Les fonctions $\det_B$ et Dét sont donc égales ; adieu la notation Dét ! La proposition 4 donne alors la proposition suivante.

Proposition 5 : formule principale. Pour toute base B de E et toute forme n -linéaire alternée $\varphi : E^n \rightarrow K$, on a :

$$\varphi(u_1, u_2, \dots, u_n) = \det_B(u_1, u_2, \dots, u_n) \varphi(e_1, \dots, e_n). \quad \square$$

Si $E = \{0\}$, le déterminant de id_E est 1.

14.5 FORMULES ET CALCULS

Règle de Sarrus. Cette règle, valable uniquement pour le calcul de déterminants d'ordre 3, a été proposée vers 1840 par Pierre Sarrus (1798-1861). Rappelons que le déterminant d'une matrice carrée d'ordre 3 est donné par la formule :

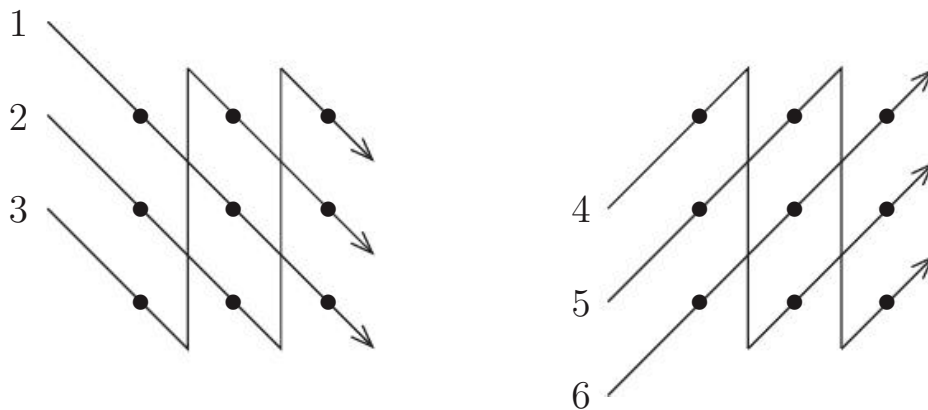
$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}$$

L'astuce de Sarrus a été de voir que les trois premiers termes, avec le signe +, correspondent aux produits des coefficients de la matrice situés sur les droites numérotées 1, 2, 3 descendant de la gauche vers la droite du dessin (voir ci-dessous), et que les trois termes suivants, avec le signe –, correspondent aux produits des coefficients de la matrice situés sur les droites numérotées 4, 5, 6 montant de la gauche vers la droite ; il faut compter les indices des neuf coefficients modulo 3, ce qui explique que les droites 2, 3, 4, 5 aient cet aspect brisé.

Proposition 1 : déterminant de la transposée. Soit $A = (a_{ij})$ une matrice carrée d'ordre n ; sa transposée est ${}^tA = (b_{ij})$ avec $b_{ij} = a_{ji}$ pour $1 \leq i, j \leq n$. On a :

$$\det({}^tA) = \det(A)$$



Démonstration. Utilisons la formule de Cramer. On a :

$$\begin{aligned} \det({}^t A) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1)1} \dots b_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}. \end{aligned}$$

Puisque σ est une permutation de $\{1, \dots, n\}$, $\sigma(1), \dots, \sigma(n)$ forment aussi l'ensemble $\{1, \dots, n\}$. Considérons un terme de la somme ci-dessus, comme $a_{1\sigma(1)} \dots a_{n\sigma(n)}$. Le terme a_{ij} de ce produit est tel que $j = \sigma(i)$, donc $i = \sigma^{-1}(j)$; le produit peut donc encore s'écrire $a_{\sigma^{-1}(1)1} \dots a_{\sigma^{-1}(n)n}$. Par conséquent :

$$\det({}^t A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma^{-1}(1)1} \dots a_{\sigma^{-1}(n)n} = \sum_{\sigma'^{-1} \in S_n} \varepsilon(\sigma'^{-1}) a_{\sigma'(1)1} \dots a_{\sigma'(n)n}$$

comme la somme pour $\sigma'^{-1} \in S_n$ est la même que la somme pour $\sigma' \in S_n$ et que la signature d'une permutation et de son inverse sont égales, on a $\det(A') = \det(A)$. $\square$

Cette formule permet de transposer aux lignes les propriétés vues pour les colonnes :

- le déterminant change de signe quand on échange deux lignes ;
- le déterminant est linéaire par rapport à chaque ligne ;
- le déterminant ne change pas si on ajoute à une ligne une combinaison linéaire des autres lignes ;

Proposition 2 : déterminant d'une matrice triangulaire. Soit $A = (a_{ij})$ une matrice triangulaire supérieure. On a $\det(A) = \prod_{1 \leq i \leq n} a_{ii}$.

Démonstration. Il suffit de faire une récurrence sur l'ordre de la matrice. $\square$

La formule est identique pour les matrices triangulaires inférieures : on peut la montrer par récurrence ou en utilisant la formule sur la transposée.

Méthode du pivot. Soit A une matrice d'ordre n dont les colonnes définissent des vecteurs $u_1, \dots, u_n$ par rapport à la base canonique. Pour calculer le déterminant de A ,

on peut utiliser la méthode du pivot sur les vecteurs $u_1, \dots, u_n$ puisque le déterminant change de signe quand on échange deux colonnes et qu'il ne change pas quand on ajoute à une colonne une combinaison linéaire des autres. Quand on s'est ramené à un système triangulaire, on applique le résultat précédent.

Par transposition, on peut également appliquer la méthode du pivot pour calculer sur les lignes.

Ces méthodes sont recommandées pour tous les calculs de déterminants numériques. Cependant, on doit souvent calculer des déterminants de matrices particulières où les calculs peuvent être conduits en suivant l'une des possibilités suivantes.

Développement suivant une ligne quelconque. Soit i , $1 \leq i \leq n$. Faisons remonter la i -ième ligne en l'échangeant successivement avec la $(i-1)$ -ième ligne, puis la $(i-2)$ -ième ligne, ..., enfin avec la première ligne. On obtient une matrice A' telle que $\det(A) = (-1)^{i-1} \det(A')$. En développant $\det(A')$ par rapport à la première ligne de A' , on a :

$$\det(A) = (-1)^{i-1} \sum_{1 \leq j \leq n} (-1)^{1+j} a_{ij} \det(A'_{ij})$$

donc :

$$\det(A) = \sum_{1 \leq j \leq n} (-1)^{1+j} a_{ij} \det(A_{1j})$$

qui permet le calcul de $\det(A)$ en développant par rapport à la i -ème ligne.

Développement suivant une colonne quelconque. Par transposition, la formule précédente donne une formule de développement par rapport à une colonne quelconque :

$$\det(A) = \sum_{1 \leq i \leq n} (-1)^{1+j} a_{ij} \det(A_{ij})$$

Pour se souvenir des signes de ces deux formules, on peut remarquer que la distribution des signes + et - avec la formule $(-1)^{i+j}$ est analogue à la distribution des cases noires et blanches sur un damier :

$$\begin{vmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Ces formules sont utiles quand une ligne ou une colonne possède beaucoup de 0. Par exemple, il est conseillé de calculer le déterminant suivant en développant par rapport à la troisième colonne puis le résultat par rapport à la seconde colonne :

$$\begin{vmatrix} 1 & 2 & 0 & 5 \\ -1 & 1 & 2 & 1 \\ 4 & 0 & 0 & 6 \\ 1 & 0 & 0 & 4 \end{vmatrix} = (-1)^{2+3} \times 2 \begin{vmatrix} 1 & 2 & 5 \\ 4 & 0 & 6 \\ -1 & 0 & 4 \end{vmatrix} = -2(-1)^{1+2} \times 2 \begin{vmatrix} 4 & 6 \\ -1 & 4 \end{vmatrix} = 88.$$

14.6 DÉTERMINANT D'UN ENDOMORPHISME

Proposition 1 : caractérisation des bases. Soient E un espace vectoriel de dimension n et $B = (e_1, \dots, e_n)$ une base de E . Une famille $C = (\varepsilon_1, \dots, \varepsilon_n)$ de n vecteurs de E est une base de E si et seulement si $\det_B(\varepsilon_1, \dots, \varepsilon_n) \neq 0$ et alors : $\det_B(\varepsilon_1, \dots, \varepsilon_n) \det_C(e_1, \dots, e_n) = 1$.

Démonstration. On a vu que si $C = (\varepsilon_1, \dots, \varepsilon_n)$ est une famille liée de E , alors $\varphi(\varepsilon_1, \dots, \varepsilon_n) = 0$ pour toute forme n -linéaire alternée $\varphi : E^n \rightarrow K$, en particulier pour $\varphi = \det_B$.

Si maintenant C est une base de E , pour toute famille $(u_1, \dots, u_n)$ de vecteurs de E , la formule de la proposition 5 de 14.4 appliquée à la forme multilinéaire alternée $\det_C$ donne :

$$\det_C(u_1, \dots, u_n) = \det_B(u_1, \dots, u_n) \det_C(e_1, \dots, e_n) ;$$

Comme $\det_C(\varepsilon_1, \dots, \varepsilon_n) = 1$, on a : $\det_B(\varepsilon_1, \dots, \varepsilon_n) \det_C(e_1, \dots, e_n) = 1$, ce qui montre en particulier $\det_B(\varepsilon_1, \dots, \varepsilon_n) \neq 0$. $\square$

Commentaire. Ce critère est apparemment très commode pour reconnaître si une famille de vecteurs forme une base. Mais il demande le calcul d'un déterminant, calcul qui se fait en appliquant l'algorithme du pivot.

Proposition 2. Soient E un espace vectoriel de dimension n , f un endomorphisme de E , et $B = (e_1, \dots, e_n)$ et $C = (\varepsilon_1, \dots, \varepsilon_n)$ deux bases de E . On a :

$$\det_B(f(e_1), \dots, f(e_n)) = \det_C(f(\varepsilon_1), \dots, f(\varepsilon_n))$$

Démonstration. La formule de la proposition 5 de 14.4 appliquée à la forme n -linéaire $\det_B$ par rapport à la base C donne :

$$\det_B(f(u_1), \dots, f(u_n)) = \det_C(f(u_1), \dots, f(u_n)) \det_B(\varepsilon_1, \dots, \varepsilon_n).$$

En particulier :

$$\det_B(f(\varepsilon_1), \dots, f(\varepsilon_n)) = \det_C(f(\varepsilon_1), \dots, f(\varepsilon_n)) \det_B(\varepsilon_1, \dots, \varepsilon_n).$$

La même formule appliquée à la forme n -linéaire alternée :

$$(u_1, \dots, u_n) \mapsto \det_B(f(u_1), \dots, f(u_n))$$

par rapport à la base B donne :

$$\det_B(f(u_1), \dots, f(u_n)) = \det_B(u_1, \dots, u_n) \det_B(f(e_1), \dots, f(e_n)).$$

En particulier :

$$\det_B(f(\varepsilon_1), \dots, f(\varepsilon_n)) = \det_B(\varepsilon_1, \dots, \varepsilon_n) \det_B(f(e_1), \dots, f(e_n)).$$

Comme $\det_B(\varepsilon_1, \dots, \varepsilon_n)$ est non nul d'après la proposition précédente, on conclut. $\square$

Déterminant d'un endomorphisme. On voit que $\det_B(f(e_1), \dots, f(e_n))$ ne dépend pas de la base B de E . On appellera déterminant de l'endomorphisme f et on notera $\det(f)$ cette valeur commune. Pour toute base $B = (e_1, \dots, e_n)$ de E , on a donc :

$$\det(f) = \det_B(f(e_1), \dots, f(e_n)).$$

La proposition suivante, dont on justifiera le nom en 14.9, a été obtenue dans la démonstration de la proposition précédente.

Proposition 3 : formule du volume. Soit E un espace vectoriel de dimension finie n et f un endomorphisme de E . Pour toute base $B = (e_1, \dots, e_n)$ de E , on a :

$$\det_B(f(u_1), \dots, f(u_n)) = \det(f) \det_B(u_1, \dots, u_n).$$

Proposition 4 : critère d'inversibilité. Soit E un espace vectoriel de dimension finie n et f un endomorphisme de E . L'endomorphisme f est inversible si et seulement si $\det(f) \neq 0$.

Démonstration. Soit $B = (e_1, \dots, e_n)$ une base de E . On a $\det(f) = \det_B(f(e_1), \dots, f(e_n))$. On sait que f est inversible si et seulement si $(f(e_1), \dots, f(e_n))$ est une base de E . La caractérisation des bases donnée au début de ce paragraphe permet de conclure. $\square$

Proposition 5 : déterminant du composé de deux endomorphismes. Soit E un espace vectoriel de dimension finie n et f, g deux endomorphismes de E . On a :

$$\det(g \circ f) = \det(g) \det(f).$$

Démonstration. Soit $B = (e_1, \dots, e_n)$ une base de E . La formule de la proposition 5 de 14.4 appliquée à la forme n -linéaire et alternée :

$$(u_1, \dots, u_n) \mapsto \det_B(g(u_1), \dots, g(u_n))$$

par rapport à la base B donne :

$$\det_B (g(f(u_1)), \dots, g(f(u_n))) = \det_B (f(u_1), \dots, f(u_n)) \times \det_B (g(e_1), \dots, g(e_n))$$

On conclut en appliquant cette formule à la famille $(e_1, \dots, e_n)$. $\square$

Déterminant de la matrice d'un endomorphisme. Soient $f : E \rightarrow E$ une application linéaire, $B = (e_1, \dots, e_n)$ une base de E et $A = M(f, B, B)$ la matrice de f par rapport à la base B . On a $\det(f) = \det(f(e_1), \dots, f(e_n))$, donc le déterminant de f est égal au déterminant de la matrice A .

14.7 DÉTERMINANT D'UNE MATRICE CARRÉE

les propriétés des déterminants d'applications linéaires donnent des propriétés des déterminants de matrices :

- $\det(A A') = \det(A) \det(A') = \det(A') \det(A)$;
- A est inversible si et seulement si $\det(A) \neq 0$;
- si A est inversible, on a $\det(A^{-1}) = \frac{1}{\det(A)}$;
- si A et A' sont semblables, on a $\det(A) = \det(A')$.

Commentaire. C'est en 1812, qu'indépendamment l'un de l'autre, Cauchy et Jacques Binet (1786-1856, d'origine rennaise) obtiennent la formule de multiplication des déterminants au bout de longs calculs.

Nous allons maintenant établir une formule donnant l'inverse d'une matrice carrée inversible.

Cofacteurs. Soit $A = (a_{ij})$ une matrice carrée d'ordre n . Pour chaque couple (i, j) , $1 \leq i, j \leq n$, on a défini en 14.2 ce qu'est la matrice extraite A_{ij} de A . On définit le cofacteur $\text{cof}_{ij}(A)$ par :

$$\text{cof}_{ij}(A) = (-1)^{i+j} \det(A_{ij})$$

La matrice carrée d'ordre n définie par $\text{cof}(A) = (\text{cof}_{ij}(A))$ est appelée matrice des cofacteurs de A . Notons que ${}^t \text{cof}(A) = \text{cof}({}^t A)$.

Proposition : formule de l'inverse. Soit $A = (a_{ij})$ une matrice carrée d'ordre n . On a :

$$A \text{cof}({}^t A) = \text{cof}({}^t A)A = \det(A)I_n$$

$$({}^t A^{-1}) = \frac{1}{\det(A)} \text{cof}({}^t A)$$

Démonstration. D'après la proposition de 8.7, il suffit de vérifier $A \operatorname{cof}({}^t A) = \det(A)I_n$. Pour tout couple (i, j) , $1 \leq i, j \leq n$, le terme d'indice (i, j) de $A \operatorname{cof}({}^t A)$ est $\sum_{1 \leq k \leq n} a_{ik} \operatorname{cof}_{jk}(A) = \sum_{1 \leq k \leq n} a_{ik} (-1)^{j+k} \det(A_{jk})$.

Distinguons deux cas :

➤ Premier cas : $i = j$

On voit que $\sum_{1 \leq k \leq n} a_{ik} (-1)^{i+k} \det(A_{ik}) = \det(A)$ en reconnaissant la formule de développement du déterminant de A suivant la i -ième ligne.

➤ Second cas : $i \neq j$

On voit que $\sum_{1 \leq k \leq n} a_{ik} (-1)^{j+k} \det(A_{jk})$ est la formule du développement suivant sa j -ième ligne de la matrice A' obtenue à partir de A en remplaçant la j -ième ligne de A par la i -ième ligne de A . Comme A' a deux lignes égales, son déterminant est nul.

Exemple. Calculons l'inverse de la matrice

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 3 & -1 & 1 \\ -2 & 1 & -3 \end{pmatrix}$$

avec cette formule. On trouve $\det(A) = 22$, par exemple en appliquant la règle de Sarrus. La matrice des cofacteurs s'obtient en n'oubliant pas de tenir compte des signes $(-1)^{i+j}$:

$$\operatorname{cof}(A) = \begin{pmatrix} 2 & 7 & 1 \\ 8 & -5 & -7 \\ 2 & -4 & -10 \end{pmatrix}$$

En n'oubliant pas de transposer, on retrouve le résultat de 8.8 :

$$A^{-1} = \frac{1}{22} \begin{pmatrix} 2 & 8 & 2 \\ 7 & -5 & -4 \\ 1 & -7 & -10 \end{pmatrix}$$

Système de Cramer. On appelle système de Cramer un système linéaire $AX = Y$ où A est une matrice carrée inversible.

On a donc :

$$X = A^{-1}Y = \frac{1}{\det(A)} \operatorname{cof}({}^t A)Y.$$

Soit i , $1 \leq i \leq n$. La i -ième ligne x_i de X est égale à $\frac{1}{\det(A)} \sum_{1 \leq j \leq n} y_j \operatorname{cof}_{ji}(A)$
 $= \frac{1}{\det(A)} \sum_{1 \leq j \leq n} (-1)^{i+j} y_j \det(A_{ji})$. C'est le déterminant de la matrice obtenue à partir de A en changeant la i -ième colonne de A en Y . C'est exactement la description que donne Cramer des numérateurs de ses formules (voir 14.1) : les valeurs des inconnues ont un dénominateur commun $\det(A)$ et, pour obtenir le numérateur de x_i , on substitue les y_j aux a_{ji} ; on expliquera cette formule à la fin de 14.9.

Faut-il souligner encore que ces belles formules n'ont que peu souvent un intérêt pratique : le nombre de multiplications qu'elles nécessitent croît rapidement avec l'ordre de la matrice.

Exemple. Reprenons le système de 8.8 :

$$\begin{cases} x + 3y - z = y_1 \\ 3x - y + z = y_2 \\ -2x + y - 3z = y_3 \end{cases}$$

La formule de Cramer donne $x = \frac{1}{22} \begin{vmatrix} y_1 & 3 & -1 \\ y_2 & -1 & 1 \\ y_3 & 1 & -3 \end{vmatrix} = \frac{1}{11} (y_1 + 4y_2 + y_3)$, etc.

14.8 RETOUR SUR LE RANG

Matrice extraite. Généralisons la notion de matrice extraite définie en 14.2. Soit $A = (a_{ij})$ une matrice à n lignes et p colonnes. On appelle matrice extraite de A , toute matrice A_{IJ} obtenue en choisissant des sous-ensembles $I \subset \{1, \dots, n\}$, $J \subset \{1, \dots, p\}$ et en supprimant dans A les lignes et les colonnes dont les indices ne sont pas dans I ou dans J .

Proposition. Une matrice A est de rang r si et seulement si on peut extraire de A une matrice carrée A_{IJ} d'ordre r et de déterminant non nul et si toute matrice extraite $A_{I'J'}$ carrée d'ordre $r+1$ a un déterminant nul.

Démonstration. Supposons A de rang r . L'algorithme du pivot permet de trouver dans la famille $C = (u_1, \dots, u_p)$ des p vecteurs colonnes $u_1, \dots, u_p$ de A une sous-famille $C' = (v_1, \dots, v_r)$ de r vecteurs linéairement indépendants. On peut supposer, à permutation des colonnes près, que $v_1 = u_1, \dots, v_r = u_r$ et, à permutation des lignes près, que l'algorithme du pivot appliqué à $v_1, \dots, v_r$ donne une famille triangulaire

$(w_1, \dots, w_r)$ de vecteurs non nuls. Posons $I = J = \{1, \dots, r\}$. Le déterminant de A_{IJ} est égal au déterminant de la matrice triangulaire dont les colonnes sont les r premières coordonnées des $w_1, \dots, w_r$. Il est donc non nul.

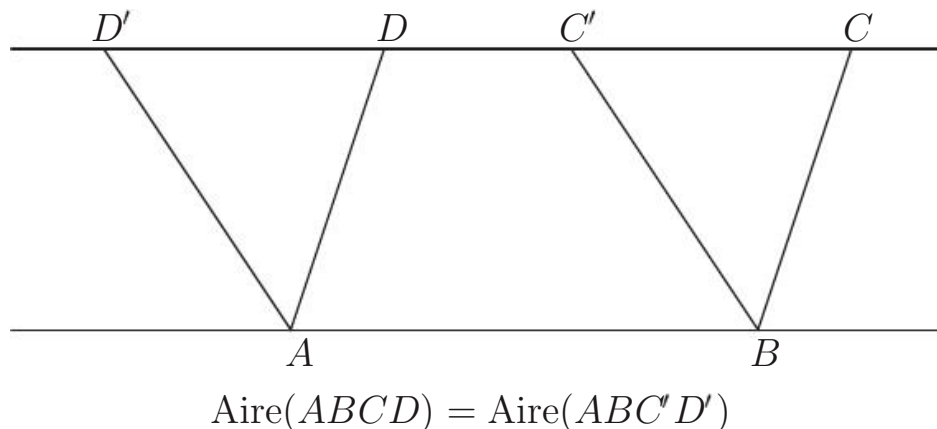
Soient maintenant I', J' des sous-ensembles de $r + 1$ éléments de $\{1, \dots, n\}$. On peut supposer que $I' = J' = \{1, \dots, r + 1\}$. La famille $(u_1, \dots, u_{r+1})$ est liée par une relation linéaire non triviale. Cette relation montre que les vecteurs définis par les composantes à indice dans J' de $u_1, \dots, u_{r+1}$ sont liés. On a donc $\det(A_{I'J'}) = 0$.

Montrons la réciproque. Soient I et J des sous-ensembles de r éléments de $\{1, \dots, n\}$ tels que $\det(A_{IJ}) \neq 0$. Les vecteurs $u_1, \dots, u_r$ ne sont pas liés par une relation non triviale, car on aurait $\det(A_{IJ}) = 0$. On peut supposer $I = \{1, \dots, r\}$ et $J = \{1, \dots, r + 1\}$. La méthode du pivot appliquée à $u_1, \dots, u_r$ et à u_{r+1} donne une famille triangulaire $v_1, \dots, v_r$ de vecteurs non nuls et un vecteur v_{r+1} . Si u_{r+1} est indépendant de $u_1, \dots, u_r$, alors v_{r+1} , dont les r premières coordonnées sont nulles, a une coordonnée d'indice $s > r$ non nulle et on a $\det A_{I'J \cup \{s\}} \neq 0$, ce qui est contradictoire. Par conséquent, le rang de A est r . $\square$

Commentaire. Le concept de rang d'un système linéaire est un concept difficile qui s'élabore à l'aide de la théorie des déterminants dans de nombreux travaux de mathématiciens de la seconde moitié du XIX^e siècle. C'est Georg Frobenius (1849-1917) qui donne le nom de rang dans des articles écrits entre 1875 et 1879 où il en dégage les principales propriétés.

14.9 DÉTERMINANT ET VOLUME

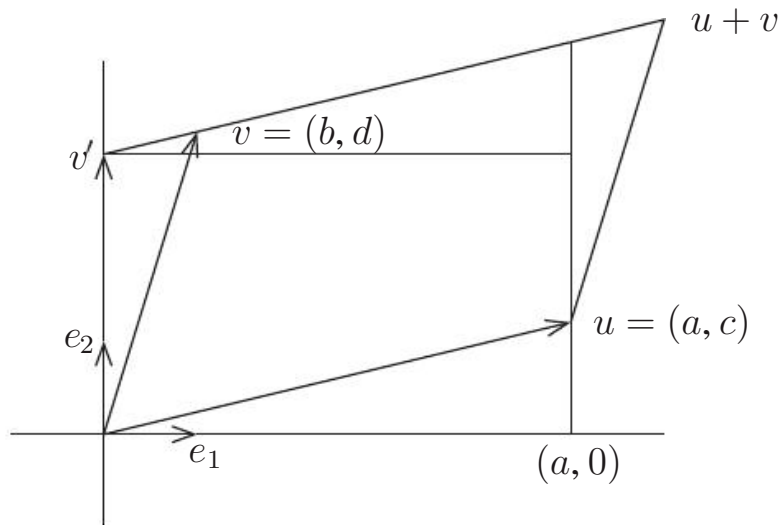
Aire des parallélogrammes. Rappelons d'abord un résultat du livre 1 des *Éléments* d'Euclide. Deux parallélogrammes qui définissent la même bande et ont des bases égales ont des aires égales.



Soient $u = (a, c)$ et $v = (b, d)$ deux vecteurs du plan muni d'une base (e_1, e_2) . On appelle parallélogramme construit sur u, v l'ensemble $\{su + tv \mid 0 \leq s, t \leq 1\}$. On pose

que l'aire du parallélogramme construit sur e_1, e_2 est 1. Quelle est l'aire $A(u, v)$ du parallélogramme construit sur u, v ?

Reprenons la méthode du pivot.



Si $a = b = 0$, u et v sont colinéaires et $A = 0$.

Sinon, on peut supposer $a \neq 0$ et la méthode du pivot conduit à poser $v' = v - \frac{b}{a}u = \left(0, d - \frac{b}{a}c\right)$. Le théorème de la bande rappelé ci-dessus montre que l'aire du parallélogramme construit sur u, v est égale à celle du parallélogramme construit sur u, v' . Cette seconde aire est égale à l'aire du parallélogramme construit sur $(a, 0)$ et v' . Par conséquent, $A(u, v) = \left|a \left(d - \frac{b}{a}c\right)\right| = |ad - bc| = |\det(u, v)|$.

Volumes dans l'espace. On vérifierait de même que le volume du parallélépipède défini par trois vecteurs u_1, u_2, u_3 de l'espace muni d'une base $B = (e_1, e_2, e_3)$ est $|\det_B(u_1, u_2, u_3)|$ en posant le volume du parallélépipède défini par e_1, e_2, e_3 égale à 1.

On peut généraliser ce qui précède à $\mathbb{R}^n$ muni de sa base canonique ou à un $\mathbb{R}$ -espace vectoriel muni d'une base B . On définira le *parallélépipède* construit sur les n vecteurs $u_1, \dots, u_n$ comme l'ensemble $\{s_1u_1 + \dots + s_nu_n \mid 0 \leq s_1, \dots, s_n \leq 1\}$ et son volume sera $|\det(u_1, \dots, u_n)|$, le volume du parallélépipède construit sur la base étant 1.

On peut également définir des volumes algébriques en supprimant les valeurs absolues.

Volume et application linéaire. Le nom de formule du volume de la proposition 3 de 14.6 :

$$\det_B(f(u_1), \dots, f(u_n)) = \det(f) \det_B(u_1, \dots, u_n).$$

s'explique, puisque $\det(f)$ s'interprète comme le facteur par lequel on multiplie les volumes des parallélépipèdes quand on applique f .

C'est Cauchy qui a fait cette remarque en 1815 à propos d'un problème de déplacement de molécules. Quelques années plus tard, Carl Jacobi (1804-1851) définira de façon générale le jacobien qui apparaît dans les formules de changement de variables dans les intégrales multiples.

Interprétation géométrique des formules de Cramer. Soit $AX = Y$ un système de Cramer. Les vecteurs $\varepsilon_i, 1 \leq i \leq n$ définies par les colonnes de A forment une base de $\mathbb{R}^n$. Notons $x = (x_1, \dots, x_n)$ et y les vecteurs définis par X et Y . On a $y = \sum_{1 \leq i \leq n} x_i \varepsilon_i$; le calcul du volume algébrique formé par les ε_j avec y remplaçant le i -ième vecteur donne immédiatement :

$$\det(\varepsilon_1, \dots, \varepsilon_{i-1}, y, \varepsilon_{i+1}, \dots, \varepsilon_n) = x_i \det(\varepsilon_1, \dots, \varepsilon_n).$$

On retrouve la formule donnée à la fin de 14.7 : le membre de gauche est le déterminant de la matrice obtenue à partir de A en changeant la i -ième colonne de A en Y et le membre de droite est $\det(A)$.

14.10 DÉTERMINANT ET ORIENTATION

Soient B et B' deux bases d'un $\mathbb{R}$ -espace vectoriel E de dimension finie. Posons $P = M(\text{id}, B, B')$; on a $P^{-1} = M(\text{id}, B', B)$.

Définition 1 : bases de même orientation. On dit que B' a même orientation que B si $\det(P) > 0$.

On a alors $\det(P^{-1}) > 0$, donc B a même orientation que B' . Si une troisième base B'' a même orientation que B' , il est facile de voir qu'elle a même orientation que B .

La relation *avoir même orientation* est donc une relation d'équivalence entre bases de E . Étant donnée une base B de E , il y a deux classes d'équivalence pour cette relation dans l'ensemble des bases : les bases de même orientation que B qu'on appelle *bases directes* (pour cette orientation) et les bases d'orientation opposée à celles de B qu'on appelle *bases indirectes*.

Exemples. Dans $\mathbb{R}$, les bases définies par un réel strictement positif sont de même orientation et les bases définies par un réel strictement négatif sont d'une autre orientation.

Dans $\mathbb{R}^2$, la base canonique (e_1, e_2) est de même orientation que la base $(e_1 + e_2, -e_1 + e_2)$ mais elle n'a pas la même orientation que la base $(e_1, -e_2)$.

On sait qu'on recherche dans certains problèmes de physique à construire des bases de $\mathbb{R}^3$ de même orientation que la base initialement choisie.

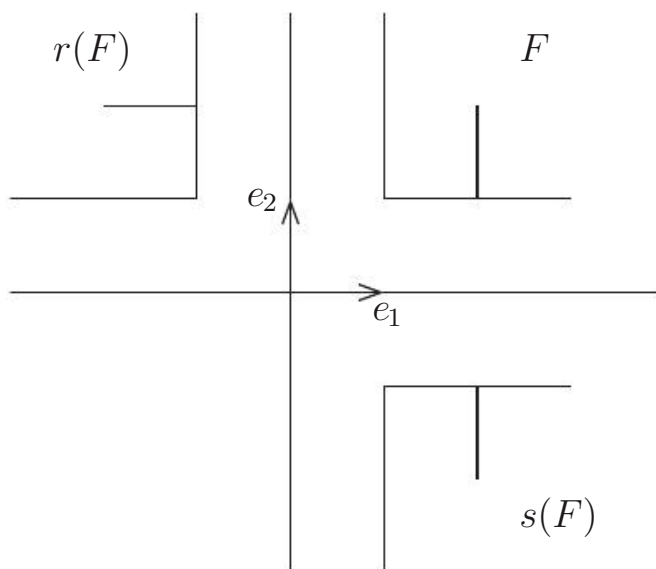
Puisqu'on ne peut définir de relation d'ordre dans $\mathbb{C}$, la notion d'orientation que nous venons de définir ne peut s'étendre aux $\mathbb{C}$ -espaces vectoriels.

Définition 2 : endomorphismes directs et indirects. Soit E un $\mathbb{R}$ -espace vectoriel et $B = (e_1, \dots, e_n)$ une base de E . La base B définit une orientation sur E . On dit qu'un automorphisme f de E est direct (ou conserve l'orientation) si la base $(f(e_1), \dots, f(e_n))$ est directe. Dans le cas contraire, on dit que f est indirect (ou ne conserve pas l'orientation).

La formule du volume montre que f est direct si et seulement si $\det(f) > 0$.

On voit alors que si f est un automorphisme direct, il conserve l'orientation des bases de E ; d'autre part, le composé de deux automorphismes directs est direct.

Exemples. Les rotations de $\mathbb{R}^2$ sont des automorphismes directs alors qu'une symétrie par rapport à une droite est un automorphisme indirect. On peut faire une figure de géométrie affine pour visualiser cela ; r désigne la rotation d'angle $\pi/2$ et s la symétrie par rapport à l'axe des x ; $r(F)$ et $s(F)$ sont les images de la figure F par r et s .



➤ Vers le chapitre 15

La notion de déterminant permet d'aborder un problème important : celui de la diagonalisation des matrices ou, plus généralement, la recherche de bases adaptées à des applications linéaires.

Exercices

14.1 Calculs

a) Calculer les déterminants des matrices suivantes :

$$A = \begin{pmatrix} 1 & 3 \\ -7 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 3 & 3 \\ 5 & 15 & 7 \\ 0 & -2 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 2 & -1 & 3 & -4 \\ 2 & 0 & 4 & -5 \\ -2 & 4 & 3 & 1 \\ 0 & -3 & 1 & -1 \end{pmatrix}$$

b) En déduire que ces matrices sont inversibles.

c) Calculer les inverses des matrices carrées d'ordre 2 et 3 précédentes en utilisant les matrices de cofacteurs (faites la même chose pour la matrice d'ordre 4 si vous le voulez : le nombre de calculs à faire augmente rapidement avec l'ordre de la matrice).

d) Calculer les déterminants des matrices :

$$D = \begin{pmatrix} 1 & \cos a & \cos 2a \\ 1 & \cos b & \cos 2b \\ 1 & \cos c & \cos 2c \end{pmatrix} \quad E = \begin{pmatrix} \sin a & \sin 2a & \sin 3a \\ \sin b & \sin 2b & \sin 3b \\ \sin c & \sin 2c & \sin 3c \end{pmatrix}$$

14.2 Matrices antisymétriques

Soit $A = (a_{ij})$ une matrice carrée d'ordre n antisymétrique, c'est-à-dire telle que $a_{ij} = -a_{ji}$ pour $1 \leq i, j \leq n$.

a) Que peut-on dire des a_{ii} pour $1 \leq i \leq n$?

b) Montrer que $\det(A) = 0$ si n est impair.

c) Calculer $\det(A)$ pour $n = 2, 4$; vérifier que ce déterminant est un carré quand on le considère comme un polynôme en les coefficients de la matrice (cette propriété se généralise aux matrices antisymétriques d'ordre pair).

d) Calculer $\det(A)$ si A est une matrice carrée antisymétrique d'ordre pair dont les coefficients de la partie triangulaire supérieure de A sont nuls, à l'exception de ceux de la diagonale allant de $(1, 2)$ à $(n-1, n)$.

Ces résultats ont été obtenus par Jacobi en 1827.

14.3 Matrice des cofacteurs

Soit A une matrice carrée d'ordre n . On note C la matrice des cofacteurs de A . Exprimer $\det(C)$ en fonction de $\det(A)$.

14.4 Calculs de déterminants d'ordre n

Les coefficients non indiqués sont des 0.

1) Calculer, pour $n \geq 2$: $D_n = \begin{vmatrix} a & b & & \\ & \ddots & \ddots & \\ & & a & b \\ b & & & a \end{vmatrix}$

2) Pour $n \geq 1$ on pose : $D_n = \begin{vmatrix} z & x & \cdots & x \\ y & z & & \\ \vdots & & \ddots & \\ y & & & z \end{vmatrix}$

a) Trouver une relation de récurrence entre D_n et D_{n-1} pour $n \geq 0$.

b) En déduire D_n .

3) Pour $n \geq 1$ on pose : $D_n = \begin{vmatrix} z & x & & \\ y & z & \ddots & \\ & \ddots & \ddots & x \\ & & y & z \end{vmatrix}$

a) Trouver une relation de récurrence entre D_n, D_{n-1} et D_{n-2} pour $n \geq 2$ (on posera $D_0 = 1$).

b) En déduire D_n .

14.5 Identités de Lagrange

Lagrange a obtenu en 1773 l'identité suivante :

$$\begin{aligned} & (xy'z'' + yz'x'' + zx'y'' - xz'y'' - yx'z'' - zy'x'')^2 \\ &= (x^2 + y^2 + z^2)(x'^2 + y'^2 + z'^2)(x''^2 + y''^2 + z''^2) \\ &+ 2(xx' + yy' + zz')(xx'' + yy'' + zz'')(x'x'' + y'y'' + z'z'') \\ &- (x^2 + y^2 + z^2)(x'x'' + y'y'' + z'z'')^2 \\ &- (x'^2 + y'^2 + z'^2)(xx'' + yy'' + zz'')^2 \\ &- (x''^2 + y''^2 + z''^2)(xx' + yy' + zz')^2. \end{aligned}$$

En donner une justification à l'aide des déterminants.

14.6 Formes multilinéaires

1) Soit E un K -espace vectoriel de dimension finie n muni d'une base $B = (e_1, \dots, e_n)$ et soit f un endomorphisme de E .

a) Montrer que l'application $\varphi : E^n \rightarrow K$ définie par $\varphi(u_1, \dots, u_n) = \sum_{1 \leq i \leq n} \det_B(u_1, \dots, f(u_i), \dots, u_n)$ est multilinéaire alternée.

b) En déduire que $\varphi(u_1, \dots, u_n) = \text{Tr}(f) \det_B(u_1, \dots, u_n)$.

2) a) Soit E un K -espace vectoriel de dimension finie n muni d'une base B et soit $(u_1, \dots, u_n)$ une famille de n vecteurs de E . Montrer que :

$$\det_B(u_1 + \dots + u_n, u_1 - u_2, \dots, u_{n-1} - u_n) = n(-1)^{n-1} \det_B(u_1, \dots, u_n).$$

Cette formule a été donnée par Catalan en 1846.

Calculer :

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{vmatrix}$$

14.7 Déterminant de Vandermonde

Soient K un corps, n un entier et $x_1, \dots, x_n$ des éléments de K . On note $V(x_1, \dots, x_n)$ le déterminant de la matrice carrée $M(x_1, \dots, x_n) = (x_i^{j-1})$. Par exemple, pour $n = 4$, on a :

$$M(x_1, x_2, x_3, x_4) = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{pmatrix}$$

1) Calculer $V(x_1, \dots, x_n)$ (on pourra retrancher de chaque colonne la colonne précédente multipliée par x_1 , en commençant par la dernière colonne).

2) À quelle condition un déterminant de Vandermonde s'annule-t-il ?

Remarque. Vandermonde ne semble pas avoir considéré particulièrement les déterminants qui portent son nom.

14.8 Équation d'un hyperplan

Soit $(u_1, \dots, u_{n-1})$ une famille libre d'un espace de dimension n . Donner, à l'aide d'un déterminant, une équation de l'hyperplan engendré par la famille.

14.9 Orientation

Soient E un $\mathbb{R}$ -espace vectoriel de dimension finie et F un hyperplan de E .

1) On choisit un vecteur v de E non dans F . Montrer que deux bases $B = (e_1, \dots, e_{n-1})$ et $B' = (e'_1, \dots, e'_{n-1})$ de F , tels que les bases $B_1 = (e_1, \dots, e_{n-1}, v)$ et $B'_1 = (e'_1, \dots, e'_{n-1}, v)$ ont même orientation dans E , définissent la même orientation de F .

2) Soit $B = (e_1, \dots, e_n)$ une base de E . Montrer que si v est un vecteur de E dont la n -ième coordonnée est > 0 , alors les bases B et $B' = (e_1, \dots, e_{n-1}, v)$ définissent la même orientation de E .

Solutions

14.1 a) On trouve $\det(A) = 26$; $\det(B) = (-1)^{3+2}(-2) \begin{vmatrix} 2 & 3 \\ 5 & 7 \end{vmatrix} = -2$ en développant

par rapport à la dernière ligne qui contient deux zéros et en n'oubliant pas la puissance de -1 convenable ; $\det(C)$ se calcule par la méthode du pivot en prenant la première ligne pour faire apparaître des zéros dans les deuxièmes et troisièmes coefficients de la première colonne, etc. ; on trouve $\det(C) = -24$.

b) Les matrices sont inversibles car leurs déterminants sont non nuls.

$$\text{c) } A^{-1} = \frac{1}{26} \begin{pmatrix} 5 & -3 \\ 7 & 1 \end{pmatrix} ; \quad B^{-1} = \begin{pmatrix} -7 & 3 & 12 \\ 0 & 0 & -\frac{1}{2} \\ 5 & -2 & -\frac{15}{2} \end{pmatrix} ;$$

$$C^{-1} = \frac{1}{24} \begin{pmatrix} 61 & -45 & 4 & -15 \\ -6 & 6 & 0 & -6 \\ 32 & -24 & 8 & 0 \\ 50 & -42 & 8 & -6 \end{pmatrix}.$$

d) On trouve $\det(D) = 2(\cos b - \cos a)(\cos c - \cos a)(\cos c - \cos b)$ en remarquant que $\cos 2x = 2 \cos^2 x - 1$, ce qui permet d'éliminer les -1 de la dernière colonne et de mettre un 2 en facteur. On peut reconnaître un déterminant de Vandermonde (voir exercice 14.8) ou continuer les calculs en soustrayant la première ligne des deux autres, etc.

On trouve $\det(E) = 4 \sin a \sin b \sin c \det(D)$.

14.2 a) La condition $a_{ii} = -a_{ii}$ implique que les coefficients diagonaux des matrices antisymétriques sont nuls.

b) On a $A = -{}^tA$, donc $\det(A) = \det(-{}^tA) = (-1)^n \det(A) = -\det(A)$, la dernière égalité parce que n est impair. D'où $\det(A) = 0$.

c) On trouve $\begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} = a^2$;

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = a^2 f^2 + b^2 e^2 + c^2 d^2 + 2acdf - 2abef - 2bcde \\ = (af - be + cd)^2.$$

d) On trouve $\det(A) = (\prod_{1 \leq k \leq n-1} a_{2k-1, 2k})^2$.

14.3 On a $A^{-1} = \frac{1}{\det(A)} C$, d'où $\det(A^{-1}) = \frac{1}{(\det(A))^n} \det(C)$; on en déduit : $\det(C) = (\det(A))^{n-1}$.

14.4 a) En développant par rapport à la première colonne, on trouve $D_n = a^n + (-1)^{n+1} b^n$.

b) En développant par rapport à la dernière colonne, on trouve

$$D_n = zD_{n-1} + (-1)^{n+1} x \begin{vmatrix} y & z & & \\ \vdots & & \ddots & \\ y & & & z \\ y & & & \end{vmatrix} = zD_{n-1} + (-1)^{n+1} x(-1)^n yz^{n-2}.$$

Par conséquent, $D_n = zD_{n-1} - xyz^{n-2}$ pour $n \geq 2$.

On est ramené à la détermination d'une suite définie par une relation de récurrence linéaire avec second membre. On reprend la méthode du chapitre 2. On recherche d'abord les solutions de $D_n = zD_{n-1}$, ce qui donne $D_n = Cz^n$. Comme le second membre est aussi de cette forme, on recherche une solution de la relation avec second membre de la forme $D_n = Knz^n$; on trouve $K = -\frac{xy}{z^2}$; enfin on cherche C tel que $D_2 = z^2 - xy$, ce qui donne $C = \frac{xy}{z^2} + 1$. On trouve finalement $D_n = z^n - (n-1)xyz^{n-2}$.

c) On trouve, en développant par rapport à la première ligne $D_n = zD_{n-1} - xyD_{n-2}$, qui est une relation de récurrence linéaire qu'on peut étudier avec les méthodes du chapitre 2. Elle est valable pour $n \geq 2$.

Si $z^2 - 4xy \neq 0$, l'équation caractéristique $r^2 - zr + xy = 0$ a deux racines réelles ou complexes r_1 et r_2 et $D_n = \frac{1}{r_1 - r_2} (r_1^{n+1} - r_2^{n+1})$ compte tenu des valeurs de D_0 et D_1 .

Si $z^2 - 4xy = 0$, l'équation caractéristique $r^2 - zr + xy = 0$ a une racine double $r = \frac{z}{2}$ et $D_n = (n+1) \left(\frac{z}{2}\right)^n$.

14.5 Posons $X = \begin{pmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix}$. Le membre de gauche de l'identité de Lagrange

est $(\det(X))^2$; le membre de droite est $\det({}^tXX)$ qui lui est égal.

14.6 1) a) La linéarité de f et la multilinéarité de $\det_B$ justifient la multilinéarité de φ . Si $1 \leq i < j \leq n$ et si $u_i = u_j$, tous les termes de la somme définissant φ sont nuls sauf deux : $\det_B(\dots, f(u_i), \dots, u_j, \dots)$ et $\det_B(\dots, u_i, \dots, f(u_j), \dots)$ qui sont opposés ; φ est donc alternée.

b) On en déduit que $\varphi(u_1, u_2, \dots, u_n) = \det_B(u_1, u_2, \dots, u_n) \varphi(e_1, \dots, e_n)$ (proposition 5 de 14.4). Si on pose $M(f, B, B) = (a_{ij})$, on a

$$\begin{aligned} \varphi(e_1, \dots, e_n) &= \sum_{1 \leq i \leq n} \det_B(e_1, \dots, f(e_i), \dots, e_n) \\ &= \sum_{1 \leq i \leq n} \det_B(e_1, \dots, a_{ii}e_i, \dots, e_n) = \sum_{1 \leq i \leq n} a_{ii} = \text{Tr}(f). \end{aligned}$$

2) a) Posons $v = u_1 + \dots + u_n$. On a

$$v + (n-1)(u_1 - u_2) + (n-2)(u_2 - u_3) + \dots + (u_{n-1} - u_n) = nu_1, \text{ d'où :}$$

$$\begin{aligned} \det_B(v, \dots, u_{n-1} - u_n) &= \det_B(nu_1, u_1 - u_2, \dots, u_{n-1} - u_n) \\ &= \det_B(nu_1, -u_2, \dots, -u_n) = n(-1)^{n-1} \det_B(u_1, \dots, u_n). \end{aligned}$$

b) On reconnaît, à transposition près, le déterminant du n -uplet $(e_1 + \dots + e_n, e_1 - e_2, \dots, e_{n-1} - e_n)$. On trouve donc -6 .

14.7 1) Retranchons de chaque colonne la colonne précédente multipliée par x_1 , en commençant par la dernière colonne. La première ligne devient $1 \dots 0$. Les lignes suivantes deviennent $1 \quad x_i - x_1 \dots x_i^{n-2} - x_1 x_i^{n-3} x_i^{n-1} - x_1 x_i^{n-2}$; on peut mettre $x_i - x_1$ en facteur. On obtient alors $V(x_1, \dots, x_n) = \prod_{2 \leq i \leq n} (x_i - x_1) V(x_2, \dots, x_n)$. D'où, par récurrence :

$$V(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

2) La formule précédente montre qu'un déterminant de Vandermonde est nul si et seulement s'il existe i et j distincts tels que $x_i = x_j$.

14.8 Un vecteur v est dans l'hyperplan si et seulement si la famille $(u_1, \dots, u_{n-1}, v)$ est liée, c'est-à-dire si et seulement si $\det(u_1, \dots, u_{n-1}, v) = 0$. Par exemple, la droite

du plan engendrée par le vecteur (a, b) a pour équation $\begin{vmatrix} a & x \\ b & y \end{vmatrix} = 0$ et le plan de $\mathbb{R}^3$

engendré par les vecteurs (a, b, c) et (a', b', c') a pour équation : $\begin{vmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{vmatrix} = 0$.

14.9 a) Posons $P = M(\text{id}_F, B, B')$ et $Q = M(\text{id}_E, B_1, B_1')$ On voit que $\det(Q) = \det(P)$.

b) Il est facile de voir que le déterminant de $M(\text{id}_E, B, B')$ est > 0 .

AUTOUR DE LA DIAGONALISATION

15

15.1 INTRODUCTION

Soit K un corps et E un K -espace vectoriel de dimension finie. Un moyen de calculer plus facilement avec une application linéaire $f: E \rightarrow E$ est de trouver une base de E dans laquelle la matrice de f ait une forme simple : diagonale si possible, ou triangulaire. Cela s'appelle diagonaliser, triangulariser (on dit aussi trigonaliser), ou plus généralement réduire l'endomorphisme f . On dit aussi qu'on effectue une diagonalisation, une triangularisation, une réduction. On dit de même qu'on diagonalise, triangularise, réduit une matrice A quand on diagonalise, triangularise, réduit l'endomorphisme associé à A dans la base canonique de K^n . Si on peut diagonaliser f ou A , on dit que f est diagonalisable ou que A est diagonalisable. Nous verrons en 15.7 que cette possibilité dépend du corps K .

Reprenons l'exemple du changement de base présenté en 8.11, celui de l'endomorphisme $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ défini par la matrice :

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{pmatrix}$$

par rapport à la base canonique $B = (e_1, e_2, e_3)$ de $\mathbb{R}^3$.

En considérant la base $B' = (v_1, v_2, v_3)$ de $\mathbb{R}^3$ définie par :

$$v_1 = (1, 1, 1)$$

$$v_2 = (1, 3, 2)$$

$$v_3 = (1, 2, 1)$$

Chapitre 15 • Autour de la diagonalisation

on a vu que la matrice $A' = M(f, B', B')$ est :

$$A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Si on pose :

$$P = M(\text{id}, B', B) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

on a :

$$P^{-1} = M(\text{id}, B, B') = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

Les relations entre A et A' sont données par le diagramme commutatif suivant (dans lequel on peut écrire si besoin, comme en 8.9 dans le cas général, les matrices des applications linéaires) :

$$\begin{array}{ccc} (\mathbb{R}^3, B) & \xrightarrow{f, A} & (\mathbb{R}^3, B) \\ \text{id}, P \uparrow & & \downarrow \text{id}, P^{-1} \\ (\mathbb{R}^3, B') & \xrightarrow{f, A'} & (\mathbb{R}^3, B') \end{array}$$

Ce sont :

$$A' = P^{-1} A P$$

$$A = P A' P^{-1}$$

À quelles conditions une base telle que B' , par rapport à laquelle la matrice de l'application linéaire a une forme simple, existe-t-elle ? Comment peut-on la trouver ? Que peut-on faire si elle n'existe pas ? Ce sont les questions auxquelles nous allons répondre, partiellement, dans ce chapitre. Nous en donnerons quelques applications.

15.2 ÉTUDE DU PROBLÈME

Soient donc E un K -espace vectoriel de dimension finie n , $f : E \rightarrow E$ un endomorphisme de E , $B = (e_1, \dots, e_n)$ une base de E et $A = M(f, B, B) = (a_{ij})$.

On cherche, si c'est possible, une base $B' = (v_1, \dots, v_n)$ de E telle que la matrice $A' = M(f, B', B')$ soit diagonale :

$$A' = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Cela implique que, pour $1 \leq i \leq n$, les vecteurs v_i vérifient :

$$f(v_i) = \lambda_i v_i$$

ce qui s'écrit encore :

$$(f - \lambda_i I)(v_i) = 0.$$

Autrement dit, v_i est un vecteur non nul de $\ker(f - \lambda_i I)$ et $f - \lambda_i I$ n'est pas injective.

Nous venons de voir, dans la proposition 4 de 14.6, une manière de caractériser un endomorphisme non injectif : c'est un endomorphisme dont le déterminant est nul. On a donc $\det(f - \lambda_i I) = 0$ pour $1 \leq i \leq n$, ce qui signifie que $\lambda_1, \dots, \lambda_n$ sont solutions de l'équation :

$$\det(f - \lambda I) = 0,$$

ce qui s'écrit, en utilisant la base B :

$$\begin{vmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Cette équation est une équation polynomiale à coefficients dans K de degré n en λ (ce que montre une récurrence évidente sur n). C'est elle qui va permettre de construire, quand c'est possible, la base B' . L'étude précédente conduit à poser les définitions suivantes.

15.3 DÉFINITIONS

Soient toujours E un K -espace vectoriel de dimension finie n , $f : E \rightarrow E$ un endomorphisme de E , B une base de E et $A = M(f, B, B) = (a_{ij})$.

Valeur propre. On appelle *valeur propre de f* tout élément λ de K tel que $\det(f - \lambda I) = 0$, autrement dit tel que $f - \lambda I$ ne soit pas injective.

On appelle *valeur propre de A* tout élément λ de K tel que $\det(A - \lambda I) = 0$, autrement dit tel que $A - \lambda I$ ne soit pas inversible¹.

1. On dit *valor propio* en espagnol, *autovalor* en Amérique latine et au Brésil, *Eigenwert* en allemand, *eigenvalue* en anglais...

Vecteur propre. On appelle *vecteur propre de f* (ou de A) associé à la valeur propre λ tout vecteur non nul de $\ker(f - \lambda I)$, c'est-à-dire tout vecteur v non nul de E tel que $f(v) = \lambda v$ ou $AV = \lambda V$ si V est la matrice de v dans la base B .

Si f est diagonalisable, alors E admet une base formée de vecteurs propres de f .

Le spectre d'un endomorphisme f ou d'une matrice A est l'ensemble de ses vecteurs propres. On les note $\text{spec}(f)$ ou $\text{spec}(A)$.

Espace propre. On appelle *espace propre de f* associé à la valeur propre λ le sous-espace vectoriel $\ker(f - \lambda I)$ de E , c'est-à-dire le sous-espace formé par les vecteurs v de E tels que $f(v) = \lambda v$.

Polynôme caractéristique. On appelle polynôme caractéristique de f ou de A , le polynôme

$$\chi(f) = \det(f - \lambda I) = \chi(A) = \det(A - \lambda I)$$

C'est un polynôme de degré n en λ .

Deux matrices semblables ont même polynôme caractéristique, car si $B = P^{-1}AP$, on a $B - \lambda I = P^{-1}(A - \lambda I)P$, donc

$$\chi(B) = \det(P^{-1})\det(A - \lambda I)\det(P) = \chi(A).$$

Proposition 1. Soient λ et μ deux valeurs propres distinctes de l'endomorphisme $f : E \rightarrow E$. Les espaces propres $\ker(f - \lambda I)$ et $\ker(f - \mu I)$ ont une intersection réduite à $\{0\}$.

Démonstration. Si $v \in \ker(f - \lambda I) \cap \ker(f - \mu I)$, on a $f(v) = \lambda v = \mu v$ donc $(\lambda - \mu)v = 0$. Comme $\lambda - \mu$ est non nul, on a $v = 0$. $\square$

Proposition 2. Soit $f : E \rightarrow E$ un endomorphisme de E . On suppose que le polynôme $\chi(f)$ a k racines distinctes $\lambda_1, \dots, \lambda_k$ et on note $E_1, \dots, E_k$ les sous-espaces propres associés. Si $E = E_1 \oplus \dots \oplus E_k$, alors f est diagonalisable.

Démonstration. Pour $i = 1, \dots, k$, choisissons une base B_i de E_i . La base B' de E réunion des B_i est formée de vecteurs propres de f associés aux valeurs propres $\lambda_1, \dots, \lambda_k$ donc la matrice $M(f, B', B')$ est diagonale. $\square$

15.4 EXEMPLE

Voyons comment retrouver B' et A' dans l'exemple donné ci-dessus. On a $K = \mathbb{R}$. Le polynôme caractéristique de f est de degré 3 en λ ; c'est :

$$\begin{vmatrix} 1-\lambda & 2 & -3 \\ 1 & 4-\lambda & -5 \\ 0 & 2 & -2-\lambda \end{vmatrix} = -\lambda(\lambda-1)(\lambda-2).$$

Ce polynôme a trois racines dans $\mathbb{R}$: $\lambda_1 = 0$, $\lambda_2 = 1$ et $\lambda_3 = 2$, ce qui donne trois valeurs propres distinctes pour f .

Pour déterminer $\ker(f - \lambda_1 I)$, on cherche les vecteurs $v = (x, y, z)$ tels que $(f - \lambda_1 I)(v) = 0$, c'est-à-dire tels que $f(v) = 0$. À l'aide de la base B cette équation donne le système :

$$\begin{cases} x + 2y - 3z = 0 \\ x + 4y - 5z = 0 \\ 2y - 2z = 0 \end{cases}$$

Ces trois équations ne sont pas indépendantes et la résolution du système conduit à $\ker(f) = \text{Vect}(v_1)$ avec $v_1 = (1, 1, 1)$. On peut prendre comme vecteur propre associé à la valeur propre $\lambda_1 = 0$ le vecteur v_1 ou tout vecteur non nul colinéaire à v_1 .

On détermine de même $\ker(f - \lambda_2 I) = \ker(f - I)$ avec le système :

$$\begin{cases} 2y - 3z = 0 \\ x + 3y - 5z = 0 \\ 2y - 3z = 0 \end{cases}$$

qui donne $\ker(f - I) = \text{Vect}(v_2)$ avec $v_2 = (1, 3, 2)$.

Enfin, de même, on trouve $\ker(f - \lambda_3 I) = \ker(f - 2I) = \text{Vect}(v_3)$ avec $v_3 = (1, 2, 1)$.

La matrice A' de f dans la base $B' = (v_1, v_2, v_3)$ est donnée par $f(v_1) = 0$, $f(v_2) = v_2$ et $f(v_3) = 2v_3$; on a bien retrouvé ce qu'on voulait.

15.5 CONDITION SUFFISANTE DE DIAGONALISABILITÉ

Proposition. Soient E un K -espace vectoriel de dimension n et soit $f : E \rightarrow E$ un endomorphisme. Si le polynôme caractéristique de f a n racines distinctes deux à deux dans K , alors f est diagonalisable.

Comme les calculs portent souvent sur des matrices, on pourra préférer l'énoncé suivant.

Proposition. Soit A une matrice carrée d'ordre n . Si le polynôme caractéristique de A a n racines distinctes deux à deux dans K , alors A est diagonalisable.

Démonstration. Soient $\lambda_1, \dots, \lambda_n$ les n racines distinctes du polynôme $\chi(f) = \det(f - \lambda I)$. Soient $v_1, \dots, v_n$ des vecteurs propres associés à ces valeurs propres, c'est-à-dire tels que $f(v_i) = \lambda_i v_i$ pour $1 \leq i \leq n$. Il suffit de montrer que $B' = (v_1, \dots, v_n)$ est une base de E .

Pour cela, montrons, par récurrence sur k , que $(v_1, \dots, v_k)$ est une famille libre de E . C'est vrai pour $k = 1$. Supposons que ce soit vrai pour un entier k , $1 \leq k < n$ et montrons que $(v_1, \dots, v_{k+1})$ est une famille libre de E .

Si on a l'écriture :

$$0 = \alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1} \quad (l)$$

on a, en prenant l'image des deux membres par f :

$$0 = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k+1} \lambda_{k+1} v_{k+1} \quad (l')$$

La combinaison $l' - \lambda_{k+1} l$ s'écrit :

$$0 = \alpha_1 (\lambda_1 - \lambda_{k+1}) v_1 + \cdots + \alpha_k (\lambda_k - \lambda_{k+1}) v_k$$

Par hypothèse de récurrence, chacun des coefficients de cette combinaison linéaire est nul et, comme les λ_i sont deux à deux distincts, cela implique $\alpha_1 = \dots = \alpha_k = 0$, ce qui, reporté dans l , donne $\alpha_{k+1} = 0$.

Par conséquent, B' est une famille libre de n éléments de E , donc une base de E . □

Pratique de la diagonalisation. Mettons en forme la démarche donnée au paragraphe 4 sur un exemple.

Pour diagonaliser $f: E \rightarrow E$, on connaît en général une base B de E et la matrice $A = M(f, B, B)$.

- 1) On calcule le polynôme caractéristique de f : $\chi(f) = \det(A - \lambda I)$.
- 2) On résout l'équation polynomiale $\chi(f) = 0$.

Si cette équation a n racines distinctes deux à deux dans K : $\lambda_1, \dots, \lambda_n$, on résout les systèmes linéaires donnés par les équations $(f - \lambda_i I)(v_i) = 0$, $1 \leq i \leq n$, pour obtenir des vecteurs propres $v_1, \dots, v_n$ associés aux valeurs propres $\lambda_1, \dots, \lambda_n$. La famille $B' = (v_1, \dots, v_n)$ est une base de E et la matrice $A' = M(f, B', B')$ est diagonale.

Pour obtenir les relations entre A et A' , on calcule l'inverse de la matrice $P = M(\text{id}, B', B)$. C'est la matrice $P^{-1} = M(\text{id}, B, B')$. On a alors :

$$A' = P^{-1} A P$$

$$A = P A' P^{-1}$$

Répetons que le meilleur moyen de ne pas se tromper dans ces deux formules est de revenir au diagramme de changement de base (voir 8.9 ou 15.1).

15.6 CONDITION NÉCESSAIRE ET SUFFISANTE DE DIAGONALISABILITÉ

Le polynôme caractéristique de f peut ne pas toujours avoir n racines distinctes dans K .

Proposition 1. Soit $f : E \rightarrow E$ un endomorphisme et soit λ_1 une valeur propre de f dans K . Si $\dim(\ker(f - \lambda_1 I)) = r$, alors λ_1 est une racine d'ordre $\geq r$ de $\chi(f)$.

Démonstration. Choisissons une base $(e_1, \dots, e_r)$ de $\ker(f - \lambda_1 I)$ et complétons-la en une base $B = (e_1, \dots, e_n)$ de E . La matrice $A = M(f, B, B)$ est de la forme :

$$\left(\begin{array}{ccc|c} \lambda_1 & \dots & 0 & | \\ \vdots & \ddots & \vdots & | \\ 0 & \dots & \lambda_1 & | \\ \hline & & 0 & | \end{array} \begin{array}{c} A_1 \\ A_2 \end{array} \right)$$

et on a : $\chi(f) = (\lambda_1 - \lambda)^r \det(A_2 - \lambda I)$. Il est possible que λ_1 soit aussi racine de $\det(A_2 - \lambda I)$. $\square$

Corollaire.

- 1) Si $\chi(f)$ n'est pas scindé dans K , alors f n'est pas diagonalisable.
- 2) S'il existe une valeur propre λ_i racine d'ordre m_i de $\chi(f)$ telle que la dimension de l'espace propre associé $\dim(\ker(f - \lambda_i I))$ soit strictement inférieur à m_i , alors f n'est pas diagonalisable.

Démonstration. Le raisonnement est le même dans les deux cas.

Soient $\lambda_1, \dots, \lambda_k$ les racines de $\chi(f)$ dans K et $m_1, \dots, m_k$ leurs ordres de multiplicité respectifs.

La proposition précédente montre que $\dim(\ker(f - \lambda_i I)) \leq m_i$ pour $1 \leq i \leq k$. On a donc : $\sum_{1 \leq i \leq k} \dim(\ker(f - \lambda_i I)) \leq \sum_{1 \leq i \leq k} m_i \leq n$. Dans les deux cas, l'une des inégalités est stricte. Par conséquent, on ne peut trouver de base de E formée de vecteurs propres puisque les sous-espaces propres de E ne peuvent engendrer E .

Le corollaire précédent donne deux conditions nécessaires de diagonalisation. La proposition suivante montre que ce sont des conditions suffisantes pour pouvoir diagonaliser.

Proposition 2. Soient $f : E \rightarrow E$ un endomorphisme. Une condition nécessaire et suffisante pour que f soit diagonalisable est que les deux conditions suivantes soient vérifiées :

- 1) le polynôme caractéristique $\chi(f)$ de f est scindé dans $K[X]$;
- 2) si on note $\lambda_1, \dots, \lambda_k$ les racines distinctes de $\chi(f)$ et $m_1, \dots, m_k$ leurs ordres de multiplicité respectifs, $E_i = \ker(f - \lambda_i I)$, $1 \leq i \leq k$, les espaces propres associés, alors $\dim(E_i) = m_i$ pour $1 \leq i \leq k$.

Commentaire. La condition de diagonalisabilité de 15.5 peut être vue comme un cas particulier de cette proposition avec $k = n$ et $m_i = 1$ pour tout i : nous en avons donné un énoncé à part car c'est le cas le plus fréquemment rencontré.

Démonstration. Pour montrer que ces conditions sont suffisantes, on va montrer par récurrence que les sous-espaces propres E_i , $1 \leq i \leq j$ sont en somme directe pour $1 \leq j \leq k$.

C'est vrai pour $j = 1$. Supposons que la propriété soit vraie pour un entier j , $1 \leq j < k$ et montrons qu'elle est vraie pour l'entier $j + 1$. Notons $S_j = \bigoplus_{1 \leq i \leq j} E_i$. Il suffit de montrer que $S_j \cap E_{j+1} = \{0\}$.

Soit donc v dans cette intersection. On a, d'une part :

$$f(v) = \lambda_{j+1}v$$

et, d'autre part, v s'écrit :

$$v = v_1 + \dots + v_j \quad (I)$$

avec $v_i \in E_i$ pour $1 \leq i \leq j$. En prenant l'image des deux membres par f , on obtient :

$$\lambda_{j+1}v = \lambda_1v_1 + \dots + \lambda_jv_j \quad (I')$$

La différence $I' - \lambda_{j+1}I$ donne :

$$0 = (\lambda_1 - \lambda_{j+1})v_1 + \dots + (\lambda_j - \lambda_{j+1})v_j$$

L'hypothèse de récurrence implique que les termes de la somme du second membre sont tous nuls. On en déduit que $v_1, \dots, v_j$ sont nuls, d'où $v = 0$.

Par conséquent, les E_i , $1 \leq i \leq k$, sont en somme directe dans E . Comme $\sum_{1 \leq i \leq k} \dim(E_i) = \sum_{1 \leq i \leq k} m_i \leq n$, on voit que $E = \bigoplus_{1 \leq i \leq k} E_i$, donc E est somme directe de ses espaces propres et f est diagonalisable. $\square$

Pratique de la diagonalisation dans ce cas. La seule différence avec la méthode de la section précédente est que les systèmes linéaires donnés par les équations $(f - \lambda_i I)(v_i) = 0$, $1 \leq i \leq k$, définissent des sous-espaces propres de dimension éventuellement > 1 et dont on devra chercher des bases.

Exemple 1. Considérons l'application linéaire $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie dans la base canonique de $\mathbb{R}^3$ par la matrice :

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Le polynôme caractéristique de A est :

$$\chi(A) = \begin{vmatrix} -\lambda & 1 & -1 \\ -1 & 2-\lambda & -1 \\ -1 & 1 & -\lambda \end{vmatrix} = -\lambda(\lambda-1)^2$$

L'espace propre E_0 associé à la valeur propre 0 est défini par $A(v) = 0$ qui donne le système :

$$\begin{cases} y - z = 0 \\ -x + 2y - z = 0 \\ -x + y = 0 \end{cases}$$

On voit que E_0 est engendré par $v_1 = (1, 1, 1)$.

L'espace propre E_1 associé à la valeur propre 1 est défini par $(A - I)(v) = 0$ qui donne un système se réduisant à l'unique équation :

$$-x + y - z = 0$$

L'espace propre E_1 est donc de dimension 2, ce qui prouve que f est diagonalisable. On trouve que E_1 est engendré, par exemple, par $v_2 = (1, 3, 2)$ et $v_3 = (1, 2, 1)$.

Par rapport à la base $B' = (v_1, v_2, v_3)$, on a :

$$A' = M(f, B', B') = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et on peut vérifier que $A = P A' P^{-1}$, $A' = P^{-1} A P$ avec P et P^{-1} comme en 15.1.

Exemple 2. Considérons maintenant l'application linéaire $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie dans la base canonique de $\mathbb{R}^3$ par la matrice :

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{pmatrix}$$

Reprenons les vecteurs v_1, v_2, v_3 et la matrice P de l'exemple 1.

Le polynôme caractéristique de A est encore : $-\lambda(\lambda - 1)^2$, et v_1 est encore un vecteur engendrant le noyau de f . L'espace propre associé à la valeur propre 1 est défini par les équations :

$$\begin{cases} 2y - 3z = 0 \\ 2x + 4y - 7z = 0 \end{cases}$$

On trouve $\ker(f - I) = \text{Vect}((1, 3, 2))$. Cet espace propre est donc de dimension 1, ce qui prouve que f n'est pas diagonalisable. On choisit alors un vecteur complétant la famille libre (v_1, v_2) en une base, par exemple le vecteur v_3 . Par rapport à la base B' , on a :

$$A' = M(f, B', B') = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

et $A = P A' P^{-1}$, $A' = P^{-1} A P$.

Nous n'expliquerons pas dans ce livre comment on peut se ramener à une matrice d'une telle forme. Ces résultats remontent à des travaux de Jordan et demanderaient quelques pages plus délicates d'explications.

15.7 CHANGEMENT DE CORPS DE BASE

Nous avons écrit au début de ce chapitre que la possibilité de diagonaliser dépendait du corps de base. En voici un exemple ; la matrice :

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

représente une rotation dans le plan vectoriel $\mathbb{R}^2$ muni d'une base orthonormée. Si $\theta \neq 0 \bmod \pi$, tout vecteur v non nul de $\mathbb{R}^2$ est transformé en un vecteur non colinéaire à v . Par conséquent, f n'a pas de vecteur propre. Si on calcule le polynôme caractéristique de f , on trouve :

$$\chi(A) = \begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{vmatrix} = \lambda^2 - 2\lambda \cos \theta + 1 = (\lambda - \cos \theta)^2 + \sin^2 \theta$$

qui n'a pas de racine réelle.

La matrice A peut aussi être considérée comme la matrice d'une application $g : \mathbb{C}^2 \rightarrow \mathbb{C}^2$. Son polynôme caractéristique est toujours $\lambda^2 - 2\lambda \cos \theta + 1$, mais doit être considéré comme un polynôme à coefficients dans $\mathbb{C}$. Il admet deux racines distinctes dans $\mathbb{C}$: $e^{i\theta}$ et $e^{-i\theta}$, ce qui prouve que g est diagonalisable, autrement dit que A est diagonalisable quand on la considère comme matrice à coefficients complexes.

L'espace propre associé à $e^{i\theta}$ est donné par l'équation :

$$(\cos \theta - e^{i\theta})x - \sin \theta y = 0$$

soit $\iota x + y = 0$. On peut prendre comme vecteur propre $v_1 = (1, -\iota)$; on peut aussi prendre $(\iota, 1) = \iota v_1$ qui lui est colinéaire dans $\mathbb{C}^2$.

De même, on peut prendre $v_2 = (1, \iota)$ comme vecteur propre associé à la valeur propre $e^{-i\theta}$. Si on pose :

$$P = \begin{pmatrix} 1 & 1 \\ -\iota & \iota \end{pmatrix}$$

on trouve

$$P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & \iota \\ 1 & -\iota \end{pmatrix}$$

et la matrice A' de g dans la base (v_1, v_2) est la matrice diagonale :

$$A' = P^{-1}AP = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

15.8 SECONDE CONDITION NÉCESSAIRE ET SUFFISANTE DE DIAGONALISABILITÉ

Cette condition est de nature différente des précédentes.

Polynômes en f . Soient E un K -espace vectoriel et $f: E \rightarrow E$ un endomorphisme de E . On sait qu'on peut composer des endomorphismes, en faire la somme et le produit par un scalaire. On peut donc construire $f^2 = f \circ f$, $f^3 = f \circ f \circ f$ et par récurrence f^n pour tout entier $n \geq 0$ (en posant $f^0 = I$) et plus généralement toute expression de la forme $a_0I + a_1f + \dots + a_nf^n$ où les coefficients a_i sont dans le corps K . Posons $P = a_0 + a_1X + \dots + a_nX^n$; P est un polynôme de $K[X]$ et l'expression $a_0I + a_1f + \dots + a_nf^n$ peut être notée $P(f)$.

Cette construction faite à la main a toutes les propriétés souhaitables :

- 1) si P_1 et P_2 sont deux polynômes de $K[X]$, alors :
 $(P_1 + P_2)(f) = P_1(f) + P_2(f)$;
 $(P_1 P_2)(f) = P_1(f) \circ P_2(f) = P_2(f) \circ P_1(f)$;
- 2) si P est un polynôme de $K[X]$ et a un élément de K , alors :
 $(aP)(f) = aP(f)$.

La propriété universelle présentée en 13.8 permet d'éviter toutes les vérifications précédentes ; comme elle est peut-être un peu abstraite, nous avons préféré reprendre complètement la construction de $P(f)$.

Pour tout polynôme $P = a_0 + a_1X + \dots + a_nX^n$ de $K[X]$ et toute matrice carrée A d'ordre n , on définit de même $P(A) = a_0I_n + a_1A + \dots + a_nA^n$ avec les mêmes propriétés de compatibilité avec sommes et produits que ci-dessus.

Définition : polynôme annulateur. On dira qu'un polynôme annule f si $P(f) = 0$; on dira aussi que P est un polynôme annulateur de f .

Considérons l'homomorphisme d'anneaux $\varphi: K[X] \rightarrow L(E)$ défini par $\varphi(X) = f$. L'ensemble des polynômes annulant f est le noyau de φ . C'est donc un idéal de $K[X]$. Comme $K[X]$ est un anneau principal, cet idéal est principal. Le polynôme unitaire engendrant cet idéal est un polynôme annulant f et de degré minimal parmi les polynômes annulateurs non nuls annulant f . On l'appelle le polynôme minimal de f . Nous ne l'étudions pas ici.

Proposition 1. Soit E un K -espace vectoriel de dimension n et f un endomorphisme de E .

- 1) Condition nécessaire de diagonalisabilité : si f est diagonalisable, il existe un polynôme scindé P de $K[X]$ n'ayant que des racines simples et tel que $P(f) = 0$ (0 désigne ici l'endomorphisme nul de E).
- 2) S'il existe un polynôme P de $K[X]$ annulant f , alors toute valeur propre λ de f est racine de P .

Démonstration. 1) Notons $B = (v_1, \dots, v_n)$ une base de vecteurs propres de E et $\lambda_1, \dots, \lambda_k$ les valeurs propres distinctes de f (plusieurs vecteurs de B peuvent être associés à la même valeur propre). Notons P le polynôme scindé défini par : $P(X) = (X - \lambda_1) \dots (X - \lambda_k)$. L'endomorphisme $P(f)$ est le produit des $f - \lambda_j I$, $1 \leq j \leq k$. Ce produit est commutatif : $(f - \lambda_j I) \circ (f - \lambda_j I) = (f - \lambda_j I) \circ (f - \lambda_j I)$. Soit v_i , $1 \leq i \leq n$, un vecteur de B et soit λ_j , $1 \leq j \leq k$, la valeur propre associée. Pour évaluer $P(f)(v_i)$, on peut commencer par appliquer $f - \lambda_j I$, ce qui donne $f(v_i) - \lambda_j v_i = 0$; donc $P(f)(v_i) = 0$. Comme $P(f)$ est nul sur chacun des vecteurs de la base B , on a $P(f) = 0$.

2) Posons $P = a_0 + a_1 X + \dots + a_n X^n$. Soient λ une valeur propre de f et v un vecteur propre de f associé à λ . On a $0 = P(f)(v) = (a_0 I + a_1 f + \dots + a_n f^n)(v) = (a_0 + a_1 \lambda + \dots + a_n \lambda^n)(v) = P(\lambda)v$. Comme $v \neq 0$, c'est que $P(\lambda) = 0$. Il est bien sûr possible que P ait des racines qui ne soient pas valeurs propres de f . $\square$

Proposition 2 : théorème des noyaux. Soient E un K -espace vectoriel et $f: E \rightarrow E$ un endomorphisme de E .

- 1) On suppose qu'il existe un polynôme P de $K[X]$ de la forme $P = ST$ avec S, T polynômes de $K[X]$ premiers entre eux, tel que $P(f) = 0$.
Alors $E = \ker S(f) \oplus \ker T(f)$.
- 2) On suppose qu'il existe un polynôme P de $K[X]$ de la forme $P = P_1 \dots P_k$ avec $P_1, \dots, P_k$ polynômes de $K[X]$ premiers entre eux deux à deux, tel que $P(f) = 0$.
Alors $E = \ker P_1(f) \oplus \dots \oplus \ker P_k(f)$.

Démonstration. 1) D'après l'identité de Bézout (voir 13.4), on sait qu'il existe des polynômes U et V de $K[X]$ tels que $US + VT = 1$. On a donc

$$U(f) \circ S(f) + V(f) \circ T(f) = I.$$

Si v est un vecteur de $\ker S(f) \cap \ker T(f)$, on a

$$v = I(v) = U(f)(S(f)(v)) + V(f)(T(f)(v)) = 0,$$

ce qui prouve : $\ker S(f) \cap \ker T(f) = \{0\}$.

Montrons maintenant que $\ker S(f) + \ker T(f) = E$. Soit $v \in E$.

On a, comme ci-dessus : $v = I(v) = U(f)(S(f)(v)) + V(f)(T(f)(v))$.

Posons $v_1 = U(f)(S(f)(v))$; on a $T(f)(v_1) = U(f)(P(f)(v_1)) = 0$ donc v_1 est dans $\ker T(f)$. De même $v_2 = V(f)(T(f)(v)) \in \ker S(f)$.

On a $v = v_1 + v_2$ et on peut alors conclure.

2) Il s'agit de généraliser ce qui précède. Pour $i = 1, \dots, k$, posons $S_i = \frac{P}{P_i}$; le polynôme S_i est le produit des polynômes $P_1, \dots, P_k$ sans le polynôme P_i . Les polynômes $S_1, \dots, S_k$ sont premiers entre eux, donc il existe, d'après l'identité de Bézout (voir 13.4), des polynômes $U_1, \dots, U_k$ de $K[X]$ tels que $U_1 S_1 + \dots + U_k S_k = 1$. On a donc $U_1(f) \circ S_1(f) + \dots + U_k(f) \circ S_k(f) = I$. Tout vecteur v de E s'écrit donc $v = U_1(f)(S_1(f)(v)) + \dots + U_k(f)(S_k(f)(v))$. Comme $U_i(f)(S_i(f)(v)) \in \ker P_i(f)$ pour $1 \leq i \leq k$, on voit que $E = \ker P_1(f) + \dots + \ker P_k(f)$.

Pour montrer que cette somme est directe, il suffit, d'après la proposition 1 de 9.3, de montrer que $\ker P_1(f) \cap (\ker P_2(f) + \dots + \ker P_k(f)) = \{0\}$, les autres égalités à montrer se déduisant de celle-ci par permutation des indices. Si v est un vecteur de cette intersection, on a $P_1(f)(v) = 0$ d'une part et, d'autre part, v peut s'écrire $v = v_2 + \dots + v_k$ avec $v_i \in \ker P_i(f)$ pour $2 \leq i \leq k$. Il existe des polynômes U et V de $K[X]$ tels que $UP_1 + VS_1 = 1$ d'après l'identité de Bézout. D'où $v = U(f)(P_1(f)(v)) + V(f)(S_1(f)(v)) = V(f)(S_1(f)(v_2 + \dots + v_k)) = 0$. On conclut que la somme est directe. $\square$

Proposition 3 : condition suffisante de diagonalisabilité. Soient E un K -espace vectoriel de dimension finie et $f : E \rightarrow E$ un endomorphisme de E . On suppose qu'il existe un polynôme scindé P de $K[X]$ n'ayant que des racines simples et tel que $P(f) = 0$.

Alors f est diagonalisable.

Démonstration. Posons $P(X) = (X - \lambda_1) \dots (X - \lambda_k)$. Les polynômes $X - \lambda_1, \dots, X - \lambda_k$ sont premiers entre eux deux à deux et on peut appliquer la proposition précédente. Donc E est somme directe des noyaux des endomorphismes $f - \lambda_i I$ (certains d'entre eux peuvent être nuls), ce qui implique que f est diagonalisable. $\square$

Les propositions 1 et 3 de ce paragraphe donnent une condition nécessaire et suffisante de diagonalisation d'un endomorphisme f : l'existence d'un polynôme scindé à racines simples annulant f .

Exemples. Connaître un polynôme annulant f est fréquent. Il suffit de penser aux symétries qui sont annulées par le polynôme $X^2 - 1$ ou encore aux projections qui sont annulées par le polynôme $X^2 - X$. Dans ces deux cas, une étude directe permet de réduire ces endomorphismes à la forme diagonale.

15.9 TRIANGULARISATION

Que faire quand on ne peut diagonaliser un endomorphisme f ? Une réduction qui paraît intéressante est la réduction à la forme triangulaire. Si cette réduction est

possible, notons $A' = (a_{ij})$ la matrice triangulaire représentant f . Le polynôme caractéristique de f est $\chi(f) = \det(A' - \lambda I) = (a_{11} - \lambda) \dots (a_{nn} - \lambda)$. Il est donc scindé. Nous allons montrer que cette condition nécessaire pour pouvoir trianguliser est également suffisante.

Remarquons avant que les problèmes de réduction à des formes triangulaires supérieures et inférieures sont équivalents : si on a trouvé une base $(v_1, \dots, v_n)$ par rapport à laquelle la matrice de f est triangulaire supérieure, alors la matrice de f est triangulaire inférieure par rapport à la base $(v_n, \dots, v_1)$.

Proposition. Soient E un K -espace vectoriel de dimension finie et $f : E \rightarrow E$ un endomorphisme de E . On suppose que le polynôme caractéristique de f est scindé dans $K[X]$.

Alors f est triangularisable.

En particulier, si $K = \mathbb{C}$, tout endomorphisme de E est triangularisable.

Démonstration. Raisonnons par récurrence sur la dimension de E .

Le résultat est vrai pour les espaces de dimension 1. Supposons-le vrai pour les espaces de dimension $\leq n-1$ et soit E de dimension n .

Soit $(X - \lambda_1) \dots (X - \lambda_n)$ la factorisation du polynôme caractéristique de f dans $K[X]$; les valeurs propres λ_i ne sont pas nécessairement distinctes. Notons v_1 un vecteur propre associé à la valeur propre λ_1 . D'après le théorème de la base incomplète, il existe une base B de E de la forme $B = (v_1, e_2, \dots, e_n)$. Par rapport à cette base, la matrice de f est de la forme :

$$A = M(f, B, B) = \begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

La famille $B_1 = (e_2, \dots, e_n)$ est une base du sous-espace $F = \text{Vect}(B_1)$ de E . Notons $g : F \rightarrow F$ l'endomorphisme défini par :

$$A_1 = M(g, B_1, B_1) = \begin{pmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Comme $\chi(A) = (\lambda_1 - \lambda)\chi(A_1)$, le polynôme caractéristique de A_1 est scindé. Comme $\dim F = n-1$, on peut appliquer l'hypothèse de récurrence : il existe donc une base $B_2 = (v_2, \dots, v_n)$ de F telle que $M(g, B_2, B_2)$ soit triangulaire supérieure. Si on pose $B' = (v_1, \dots, v_n)$, $M(f, B', B')$ est triangulaire supérieure. $\square$

Pratique de la triangularisation.

- 1) On calcule le polynôme caractéristique de f .
- 2) On résout l'équation polynomiale $\chi(f) = 0$.

Si cette équation a ses n racines dans K , on peut trianguler.

- 3) On recherche l'espace propre associé à une valeur propre choisie de f .
- 4) On en détermine une base qu'on complète en une base de E .
- 5) On détermine la matrice de f par rapport à cette base.
- 6) On recommence la méthode sur la matrice d'ordre inférieur, etc.

15.10 THÉORÈME DE HAMILTON-CAYLEY

Dans l'article de 1858 dont nous avons déjà parlé, Cayley énonce la version matricielle du théorème suivant, en ne la vérifiant que pour les matrices d'ordre 2 et 3.

Vérification du Théorème pour les matrices carrées d'ordre 3

I have verified the theorem, in the next simplest case of a matrix of the order 3, viz if M be such a matrix, suppose

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

then the derived determinant vanishes, or we have

$$\begin{vmatrix} a-M & b & c \\ d & e-M & f \\ g & h & i-M \end{vmatrix} = 0$$

or expanding

$$M^3 - (a + e + i)M^2 + (ei + ia + ae - fh - cg - bd)M - (aei + bfg + cdh - afh - bdi - ceg) = 0$$

but I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree.

Cayley – *A memoir on the theory of matrices*, Philosophical Transactions of the Royal Society of London, 1858.

Théorème de Hamilton-Cayley. Soit E un K espace-vectoriel de dimension finie. Tout endomorphisme $f : E \rightarrow E$ est annulé par son polynôme caractéristique.

Démonstration. Notons P le polynôme caractéristique de f . Puisque $P(\lambda) = \det(f - \lambda I)$, on peut être tenté de remplacer λ par $f : P(f) = \det(f - f) = 0$ et conclure que le théorème est évident. Mais quel sens donner à ce calcul ?

À défaut de pouvoir donner une démonstration formelle générale, qui demande des connaissances plus avancées, nous allons raisonner dans le cas où $K = \mathbb{C}$.

Nous venons de voir qu'il existe une base $B' = (v_1, \dots, v_n)$ de E telle que $A' = M(f, B', B') = (a_{ij})$ soit triangulaire. On a donc $P(\lambda) = \det(A' - \lambda I) = (a_{11}I - \lambda) \dots (a_{nn}I - \lambda)$. Calculons $P(f)$ sur les vecteurs de la base B' .

Comme $(a_{11}I - f)(v_1) = 0$, on a $P(f)(v_1) = 0$.

On a $f(v_2) = a_{12}v_1 + a_{22}v_2$, donc :

$$(a_{11}I - f)(a_{22}I - f)(v_2) = (a_{11}I - f)(-a_{12}v_1) = 0,$$

d'où $P(f)(v_2) = 0$.

On peut montrer de même que $P(f)(v_3) = 0$. En effet :

$$\begin{aligned} (a_{11}I - f)(a_{22}I - f)(a_{33}I - f)(v_3) \\ = (a_{11}I - f)(a_{22}I - f)(-a_{13}v_1 - a_{23}v_2) = 0. \end{aligned}$$

Ces calculs suggèrent un raisonnement par récurrence que nous laissons aux lecteurs et lectrices. □

Application au calcul de l'inverse. Soit f un endomorphisme d'un espace vectoriel de dimension n ; notons $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0$ le polynôme caractéristique de f . Le théorème de Hamilton-Cayley montre que :

$$f^n + a_{n-1}f^{n-1} + \dots + a_1f + a_0I = 0,$$

d'où :

$$f \circ [f^{n-1} + a_{n-1}f^{n-2} + \dots + a_1I] = -a_0I.$$

Si f est inversible, on a $a_0 = \det(f) \neq 0$ et l'inverse de f est l'application linéaire définie par l'expression entre crochets ci-dessus divisée par $-a_0$.

15.11 QUELQUES APPLICATIONS

15.11.1 Calcul de puissances

Pour calculer une puissance élevée d'un endomorphisme ou d'une matrice, on peut d'abord chercher une base dans laquelle les calculs seront plus simples. En effet, si

on sait que $A = PA'P^{-1}$ avec A' diagonale, on a $A^2 = PA'P^{-1} PA'P^{-1} = PA'^2 P^{-1}$, $A^3 = P A'P^{-1} P A'P^{-1} P A'P^{-1} = P A'^3 P^{-1}$ et pour $n \geq 0$:

$$A^n = P A'^n P^{-1}.$$

Si A est inversible, on a $A^{-1} = P A'^{-1} P^{-1}$ et, pour $n < 0$, $A^n = P A'^n P^{-1}$.

15.11.2 Systèmes différentiels linéaires du premier ordre

Prenons un exemple où apparaît la matrice A de l'exemple 15.1. Il s'agit de trouver l'ensemble des fonctions dérivables $x_1, x_2, x_3 : \mathbb{R} \rightarrow \mathbb{R}$ telles que :

$$\begin{cases} x_1' &= x_1 + 2x_2 - 3x_3 \\ x_2' &= x_1 + 4x_2 - 5x_3 \\ x_3' &= \quad \quad 2x_2 - 2x_3 \end{cases}$$

Le système se met sous la forme $X' = AX$ avec $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $X' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$.

Les calculs se présentent sous forme matricielle. On a vu que $A = P A'P^{-1}$; le système s'écrit donc $X' = AX = PA'P^{-1} X$, soit $P^{-1}X' = A'P^{-1}X$. En posant

$Y = P^{-1}X$, on obtient l'équation $Y' = A'Y$. Si on pose $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, l'équation

$Y' = A'Y$ équivaut au système :

$$\begin{cases} y_1' &= 0 \\ y_2' &= y_2 \\ y_3' &= 2y_3 \end{cases}$$

On en déduit :

$$\begin{cases} y_1 &= C_1 \\ y_2 &= C_2 e^t \\ y_3 &= C_3 e^{2t} \end{cases}$$

où C_1, C_2, C_3 sont des réels quelconques.

Pour trouver les fonctions x_1, x_2, x_3 , il suffit de calculer $X = PY$.

On résoudra avec la même méthode de changement de base des systèmes de la forme $X' = AX + B$ (voir exercice 15.1 3b).

15.11.3 Retour aux équations différentielles linéaires du second ordre

Une méthode pour résoudre un système différentiel linéaire où apparaissent des dérivées d'ordre $1, \dots, k$ d'une fonction inconnue $x : \mathbb{R} \rightarrow \mathbb{R}$ est d'introduire de

nouvelles fonctions inconnues $u_1 = x'$, $u_2 = x''$, ..., $u_{k-1} = x^{(k-1)}$ et d'ajouter au système les équations $x' = u_1$, $u_1' = u_2$, ... Cette méthode permet d'aborder autrement les équations différentielles linéaires du second ordre à coefficients constants du chapitre 1, ce que nous allons montrer sur des exemples.

Exemple 1. L'équation différentielle :

$$x'' - 3x' + 2x = 0$$

est équivalente au système :

$$\begin{cases} x' &= & u \\ u' &= & -2x + 3u \end{cases}$$

obtenu en ajoutant une fonction inconnue $u = x'$. Ce système s'écrit $X' = AX$ en

posant $X = \begin{pmatrix} x \\ u \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$.

En diagonalisant A , on trouve que $A = P A' P^{-1}$ avec $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ et $A' = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

On remarquera que le polynôme caractéristique de A : $\lambda^2 - 3\lambda + 2$ correspond à l'équation caractéristique de l'équation différentielle : $r^2 - 3r + 2$.

Le système s'écrit alors $Y' = A'Y$ avec $Y = P^{-1}X = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, c'est-à-dire :

$$\begin{cases} y_1' &= & y_1 \\ y_2' &= & 2y_2 \end{cases}$$

On a donc $y_1(t) = C_1 e^t$, $y_2(t) = C_2 e^{2t}$, avec C_1 et C_2 réels quelconques. La relation $X = PY$ donne alors $x(t) = C_1 e^t + C_2 e^{2t}$. On retrouve le résultat de la proposition de 1.5. avec cette fois la certitude que l'ensemble des solutions de l'équation différentielle est exactement l'ensemble des fonctions de la forme $t \mapsto C_1 e^t + C_2 e^{2t}$.

Exemple 2. L'équation différentielle :

$$x'' - 2x' + x = 0$$

est équivalente au système :

$$\begin{cases} x' &= & u \\ u' &= & -x + 2u \end{cases}$$

Ce système s'écrit $X' = AX$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$.

Le polynôme caractéristique de A est $\lambda^2 - 2\lambda + 1$, ce qui donne la seule valeur propre 1. L'espace propre associé est engendré par le vecteur $v = (1, 1)$, ce qui montre que A n'est pas diagonalisable. On peut cependant la trianguler en prenant par exemple le vecteur $w = (0, 1)$ pour former la base (v, w) . On a $A = PA'P^{-1}$

avec $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ et $A' = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Le système s'écrit alors $Y' = A'Y$ avec $Y = P^{-1}X = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, c'est-à-dire :

$$\begin{cases} y_1' &= y_1 + y_2 \\ y_2' &= y_2 \end{cases}$$

On a donc $y_2(t) = C_1 e^t$, et la première équation donne, par exemple avec la méthode de la variation de la constante, $y_1(t) = (C_1 t + C_2) e^t$, avec C_1 et C_2 réels quelconques. La relation $X = PY$ donne alors $x(t) = (C_1 t + C_2) e^t$, le résultat du chapitre 1.

Exemple 3. Traitons enfin le cas où l'équation caractéristique a deux racines complexes conjuguées sur l'exemple de l'équation $x'' + x = 0$. Cette équation est équivalente au système :

$$\begin{cases} x' &= u \\ u' &= -x \end{cases}$$

qui s'écrit $X' = AX$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

Le polynôme caractéristique de A est $\lambda^2 + 1$ qui n'a pas de racine réelle. On peut considérer cette matrice comme une matrice à coefficients complexes. Son polynôme caractéristique a deux racines complexes conjuguées : i et $-i$. On a

$$A = P A' P^{-1} \text{ avec } P = \begin{pmatrix} 1 & 0 \\ i & -i \end{pmatrix} \text{ et } A' = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Le système s'écrit alors $Y' = A'Y$ avec $Y = P^{-1}X = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, c'est-à-dire :

$$\begin{cases} y_1' &= i y_1 \\ y_2' &= -i y_2 \end{cases}$$

On a donc $y_1(t) = C_1 e^{it}$ et $y_2(t) = C_2 e^{-it}$ avec C_1 et C_2 complexes quelconques. La relation $X = PY$ donne alors $x(t) = C_1 e^{it} + C_2 e^{-it}$. Pour trouver l'ensemble des solutions de l'équation initiale, on peut encore écrire $C_1 = a + ib$, $C_2 = c + id$, avec a, b, c, d réels. Pour que x soit à valeurs réelles, sa partie imaginaire $(b + d) \cos t + (a - c) \sin t$ doit être nulle. Comme les fonctions $t \mapsto \cos t$ et $t \mapsto \sin t$ sont

linéairement indépendantes, on doit avoir $a = c$ et $d = -b$, d'où $x(t) = 2a \cos t - 2b \sin t$, le résultat de 1.5.

15.11.4 Suites récurrentes linéaires

La situation est analogue à celle de la sous-section précédente.

Traisons l'exemple, déjà étudié en 2.5, des suites dites de Fibonacci définies par la relation de récurrence :

$$u_{n+2} = u_{n+1} + u_n$$

On introduit la suite v_n définie par $v_n = u_{n+1}$, ce qui nous ramène à l'étude du système :

$$\begin{cases} u_{n+1} &= & v_n \\ v_{n+1} &= & u_n + v_n \end{cases}$$

qui s'écrit $U_{n+1} = AU_n$ en posant $U_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

Le polynôme caractéristique de A est $\lambda^2 - \lambda - 1 = 0$. Notons $r_1 = \frac{1+\sqrt{5}}{2}$ et $r_2 = -\frac{1}{r_1} = \frac{1-\sqrt{5}}{2}$ ses racines.

La diagonalisation de A donne $A = P A' P^{-1}$ avec $A' = \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 \\ r_1 & r_2 \end{pmatrix}$.

En posant $S_n = \begin{pmatrix} s_n \\ t_n \end{pmatrix} = P^{-1}U_n$, on se ramène au système $S_{n+1} = A'S_n$ dont les solutions sont $s_n = C_1 r_1^n$ et $t_n = C_2 r_2^n$. Enfin, $U_n = P S_n$ donne les solutions trouvées au chapitre 2.

➤ Vers le chapitre 16

Pour terminer la partie consacrée à la deuxième année, nous allons aborder la notion d'orthogonalité, en présentant les bases de la théorie des espaces munis d'un produit scalaire : les espaces euclidiens et préhilbertiens réels.

Exercices

15.1 Diagonalisation

1) Déterminer le polynôme caractéristique, les valeurs propres, les espaces propres associés des matrices suivantes et écrire la formule de changement de base.

a) $A = \begin{pmatrix} 5 & -3 \\ 6 & -6 \end{pmatrix}$

b) $A = \begin{pmatrix} 5 & 8 & -7 \\ -2 & -3 & 2 \\ 2 & 4 & -4 \end{pmatrix}$

2) Déterminer les suites (u_n) et (v_n) satisfaisant le système de relations de récurrence :

a) $\begin{cases} u_{n+1} = 5u_n - 3v_n \\ v_{n+1} = 6u_n - 6v_n \end{cases}$

b) $\begin{cases} u_{n+1} = 5u_n - 3v_n + 2 \times 3^{n+1} + 1 \\ v_{n+1} = 6u_n - 6v_n + 4 \times 3^n + 3 \end{cases}$

3) Résoudre les systèmes différentiels :

a) $\begin{cases} u' = 5u - 3v \\ v' = 6u - 6v \end{cases}$

b) $\begin{cases} u' = 5u - 3v + 6e^{3t} + 1 \\ v' = 6u - 6v + 4e^{3t} + 3 \end{cases}$

15.2 Comparaison de $f \circ g$ et $g \circ f$

Soient E un espace vectoriel de dimension n et f, g deux endomorphismes de E

1) Montrer que l'ensemble des valeurs propres de $f \circ g$ est égal à l'ensemble des valeurs propres de $g \circ f$ (distinguer les cas $\lambda = 0$ et $\lambda \neq 0$).

2) On suppose que f et g commutent et que f a n valeurs propres distinctes.

a) Montrer que les vecteurs propres de f sont vecteurs propres de g .

b) En déduire qu'il existe une base B de E dans laquelle les matrices de f et g sont diagonales.

c) Montrer qu'il existe un polynôme P de $K_{n-1}[X]$ tel que $g = P(f)$.

3) On suppose que f et g commutent. Montrer que tout espace propre de f est stable par g . En déduire que le noyau de f est stable par g et le noyau de g stable par f .

15.3 Endomorphisme de $L(E)$

Soient E un espace vectoriel de dimension n , $L(E)$ l'espace vectoriel des endomorphismes de E et p un projecteur de rang r de E . On définit l'endomorphisme F de $L(E)$ par $F(f) = \frac{1}{2}(p \circ f + f \circ p)$

$$L(E) \text{ par } F(f) = \frac{1}{2}(p \circ f + f \circ p)$$

- a) Calculer F^2 et F^3 .
- b) Déterminer un polynôme de degré 3 annihilant F .
- c) En déduire que F est diagonalisable.
- d) Déterminer les valeurs propres, les espaces propres de F et leurs dimensions en fonction de $\ker(p)$, $\text{Im}(p)$ et r .

15.4 Espaces de fonctions 1

- a) On pose $E = C^\infty(\mathbb{R}, \mathbb{R})$. On note $L : E \rightarrow E$ l'application linéaire définie par $L(f) = f''$. Déterminer les valeurs propres de L et les espaces propres associés.
- b) On pose $E = \text{Cont}([0, 1], \mathbb{R})$. On note $L : E \rightarrow E$ l'application linéaire définie par $L(f)(x) = \int_0^1 \inf(x, t) f(t) dt$. Déterminer les valeurs propres de L et les espaces propres associés (on explicitera $L(f)(x)$, puis on calculera la dérivée seconde de $g = L(f)$).

15.5 Espaces de fonctions 2

Soit E l'espace des fonctions continues et bornées de $\mathbb{R}$ dans lui-même. On définit l'endomorphisme L de E par $L(f)(x) = f(x - 1)$.

- a) Soit λ un réel. On suppose $\lambda \neq \pm 1$. Montrer que λ n'est pas valeur propre de L (on distinguera les cas $\lambda = 0$, $|\lambda| > 1$ et $|\lambda| < 1$).
- b) Déterminer les valeurs propres de L et les espaces propres associés.

15.6 Diagonalisabilité

- a) Soit A une matrice carrée d'ordre $n > 1$ telle que $(A - 2I)^5 = 0$ et $(A - 2I)^2 \neq 0$. A est-elle diagonalisable ?
- b) Soit A une matrice nilpotente (exercice 8.5) non nulle. A est-elle diagonalisable ?
- c) **Endomorphisme de rang 1**

Soit f un K -endomorphisme de rang 1 d'un K -espace vectoriel de dimension finie $n \geq 2$. Montrer que f est diagonalisable si et seulement si $\text{Tr}(f) \neq 0$ (on pourra considérer l'ordre de la racine 0 du polynôme caractéristique de f).

15.7 Soient A et B deux matrices carrées d'ordre n à coefficients réels. On suppose que A et B sont semblables en tant que matrices à coefficients dans $\mathbb{C}$. Montrer qu'elles sont semblables en tant que matrices à coefficients dans $\mathbb{R}$ (si $A = PBP^{-1}$ avec P à coefficients complexes, on posera $P = P_1 + iP_2$ avec P_1, P_2 à coefficients réels et on cherchera une matrice de la forme $P_1 + xP_2$).

15.8 Calcul du polynôme caractéristique et de l'inverse : méthode de Le Verrier

On pose $A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & -1 \\ 2 & -4 & -1 \end{pmatrix}$. On note $\chi(A) = X^3 + aX^2 + bX + c$ le polynôme caractéristique de A et $\lambda_1, \lambda_2, \lambda_3$ les valeurs propres de A . Pour $k = 1, 2, 3$, on pose $p_k = \lambda_1^k + \lambda_2^k + \lambda_3^k$.

- Calculer A^2 et A^3 .
- Exprimer a, b, c en fonction des valeurs propres de A .
- Exprimer a, b, c en fonction de p_1, p_2, p_3 .
- En déduire A^{-1} .

Commentaire. Cette méthode de calcul de l'inverse d'une matrice est appelée méthode de Le Verrier ; elle peut paraître compliquée pour une matrice d'ordre 3, mais elle donne de très bons résultats, au moins théoriques, pour des matrices de tailles plus grandes. Urbain Le Verrier (1811-1877) est un astronome célèbre pour avoir découvert en 1846 par le calcul la planète Neptune à partir des irrégularités de l'orbite de la planète Uranus.

15.9 Décomposition de Jordan-Chevalley

Cet exercice propose la démonstration d'un résultat souvent appelé *décomposition de Dunford*, du nom de Nelson Dunford (1906-1986), sur la réduction des endomorphismes, un résultat qui remonte à Jordan.

Soient E un $\mathbb{C}$ -espace vectoriel de dimension n et $f : E \rightarrow E$ un endomorphisme. On suppose que le polynôme caractéristique P de f est scindé, et on propose de montrer qu'il existe un couple unique (d, n) d'endomorphismes de E tels que :

- $f = d + n$ (décomposition de Dunford) ;
- d est diagonalisable ;
- n est nilpotent ;
- d et n commutent ;
- d et n peuvent s'écrire comme des polynômes à coefficients dans $\mathbb{C}$ de f .

On suppose que P s'écrit $P(X) = \prod_{1 \leq i \leq r} P_i$ avec $P_i(X) = (X - \lambda_i)^{k_i}$, les λ étant des éléments distincts de $\mathbb{C}$ et les k_i des entiers strictement positifs. On note $E_i = \ker((f - \lambda_i)^{k_i})$.

- 1) Montrer que $E = \bigoplus_{1 \leq i \leq r} E_i$.
- 2) Montrer qu'il existe un polynôme T de $\mathbb{C}[X]$ tel que, pour $i = 1, \dots, r$, on ait $T(X) = \lambda_i \bmod P_i$ dans $\mathbb{C}[X]$.
- 3) Montrer que $d = T(f)$ est diagonalisable.
- 4) Montrer que le couple (d, n) avec $n = f - d$ satisfait toutes les conditions demandées.
- 5) Pour montrer l'unicité du couple (d, n) , on raisonne par l'absurde en supposant l'existence d'un couple (d', n') ayant les mêmes propriétés que (d, n) .
 - a) Montrer que d, d', n, n' commutent entre eux.
 - b) En utilisant la relation $dd' = d'd$, montrer que $d - d'$ est diagonalisable (utiliser l'exercice 15.2).
 - c) Montrer que $n' - n$ est nilpotent.
 - d) Conclure.

Commentaire. La démonstration précédente s'appuie essentiellement sur la connaissance des valeurs propres de l'endomorphisme f , une connaissance illusoire quand la dimension de l'espace vectoriel est supérieur à 3 ou 4. Une méthode due à Claude Chevalley (1909-1984), publiée dans ses livres fameux des années 1950 sur les groupes de Lie, évite cette difficulté. En s'appuyant sur la méthode de Newton de calcul des valeurs approchées d'une équation, elle permet de calculer en un nombre fini d'itérations (le logarithme en base 2 de $\dim(E)$) les endomorphismes d et n en donnant leur expression dans la base B dans laquelle f est donné. Ainsi, on connaît d et n par leurs matrices dans la base B sans pouvoir diagonaliser explicitement la matrice de d .

Solutions

15.1 1) a) Le polynôme caractéristique de A est :

$$\chi(A) = \det(A - \lambda I) = \begin{vmatrix} 5 - \lambda & -3 \\ 6 & -6 - \lambda \end{vmatrix} = \lambda^2 + \lambda - 12 = (\lambda + 4)(\lambda - 3).$$

Les racines de ce polynôme sont -4 et 3 . Ce sont les valeurs propres de A .

L'espace propre associé à -4 est défini par le système de deux équations à deux

inconnues $(A + 4I) \begin{pmatrix} x \\ y \end{pmatrix} = 0$, qui est de rang 1 et se réduit à $3x - y = 0$; on prend

comme vecteur propre $u = (1, 3)$.

L'espace propre associé à 3 est défini par le système de deux équations à deux inconnues $(A - 3I) \begin{pmatrix} x \\ y \end{pmatrix} = 0$, ce qui donne $2x - 3y = 0$; on prend comme vecteur propre $v = (3, 2)$.

Notons $B = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $B' = (u, v)$ la base formée des vecteurs propres précédents. On a $P = M(\text{id}, B', B) = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$. On calcule alors $P^{-1} = -\frac{1}{7} \begin{pmatrix} 2 & -3 \\ -3 & 1 \end{pmatrix}$.

On peut vérifier que $D = \begin{pmatrix} -4 & 0 \\ 0 & 3 \end{pmatrix} = P^{-1}AP$.

b) Ici $\chi(A) = \det(A - \lambda I) = -\lambda^3 - 2\lambda^2 + \lambda + 2 = -(\lambda - 1)(\lambda + 1)(\lambda + 2)$. Les valeurs propres sont 1, -1 et -2 .

L'espace propre associé à 1 est défini par le système d'équations $(A - I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$, ce qui donne le système de rang 2 :

$$\begin{cases} 4x + 8y - 7z = 0 \\ -2x - 4y + 2z = 0 \\ 2x + 4y - 5z = 0 \end{cases}$$

On prend comme vecteur propre $u = (2, -1, 0)$.

On trouve de même un vecteur propre $v = (1, 1, 2)$ associé à -1 et un vecteur propre $w = (1, 0, 1)$ associé à -2 .

Notons $B = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et $B' = (u, v, w)$ la base formée de

vecteurs propres. On a $P = M(\text{id}, B', B) = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$. On calcule alors

$$P^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & -1 \\ -2 & -4 & 3 \end{pmatrix}.$$

On peut vérifier que $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix} = P^{-1}AP$.

2) Nous utilisons les calculs et les notations du 1a ci-dessus et les méthodes du chapitre 2.

Posons $U_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$ et $W_n = \begin{pmatrix} w_n \\ x_n \end{pmatrix} = P^{-1}U_n$.

a) Le système proposé équivaut à $U_{n+1} = AU_n$, donc à $U_{n+1} = PD P^{-1}U_n$ donc à $W_{n+1} = DW_n$. On est ramené aux deux équations $w_{n+1} = -4w_n$ et $x_{n+1} = 3x_n$, dont les solutions sont de la forme $w_n = c(-4)^n$ et $x_n = d3^n$ avec c et d réels. Comme $U_n = P W_n$, on obtient $u_n = c(-4)^n + 3d3^n$ et $v_n = 3c(-4)^n + 2d3^n$.

b) Posons $B_n = \begin{pmatrix} 1 + 2 \times 3^{n+1} \\ 3 + 4 \times 3^n \end{pmatrix}$. Le système proposé équivaut à $W_{n+1} = DW_n + P^{-1} B_n$. Le calcul de P^{-1} est ici nécessaire. On est ramené aux deux équations $w_{n+1} = -4w_n + 1$ et $x_{n+1} = 3x_n + 2 \times 3^n$. On cherche des solutions particulières de ces équations de la forme $w_n = C_1$, $x_n = C_2 3^n$; on trouve $C_1 = \frac{1}{5}$ et $C_2 = \frac{2}{3}$. Les solutions du système initial sont de la forme :

$$u_n = c(-4)^n + 3d3^n + \frac{1}{5} + 2n \times 3^n$$

$$v_n = 3c(-4)^n + 2d3^n + \frac{3}{5} + \frac{4}{3}n \times 3^n.$$

3) On procède de façon analogue au 2), en utilisant les méthodes du chapitre 1. On trouve :

$$u(t) = ce^{-4t} + 3de^{3t} + 6te^{3t} + \frac{1}{4};$$

$$v(t) = 3ce^{-4t} + 2de^{3t} + 4te^{3t} + \frac{3}{4}.$$

15.2 1) Si $\lambda = 0$ est valeur propre de $f \circ g$, on a $\det(f \circ g) = 0$ donc $\det(g \circ f) = \det(f) \det(g) = \det(f \circ g) = 0$ donc $\lambda = 0$ est valeur propre de $g \circ f$.

Si $\lambda \neq 0$ est valeur propre de $f \circ g$, il existe un vecteur v non nul de E tel que $f(g(v)) = \lambda v$; on a $g(v) \neq 0$ et $(g \circ f)(g(v)) = g(f(g(v))) = \lambda g(v)$, ce qui prouve que λ est valeur propre de $g \circ f$.

2) a) Soit λ une valeur propre de f et v un vecteur propre associé. On a $f(g(v)) = g(f(v)) = \lambda g(v)$. Si $g(v) = 0$, alors 0 est valeur propre de g et v est un vecteur propre associé. Si $g(v) \neq 0$, alors $g(v)$ est un vecteur propre de f associé à la valeur propre λ . Comme les valeurs propres de f sont distinctes, les espaces propres de f sont de dimension 1; donc $g(v)$ est colinéaire à v , ce qui prouve que v est vecteur propre de g .

b) Les vecteurs propres de f forment une base B de E ; comme ce sont des vecteurs propres de g , la matrice de g par rapport à B est diagonale.

c) Posons $B = (v_1, \dots, v_n)$. Pour tout i , $1 \leq i \leq n$, notons λ_i et μ_i , les scalaires tels que $f(v_i) = \lambda_i v_i$ et $g(v_i) = \mu_i v_i$. On cherche un polynôme $P = \sum_{1 \leq k \leq n-1} x_k X^k$ tel que $g = P(f)$. Cette condition équivaut à $g(v_i) = P(f)(v_i)$ pour $1 \leq i \leq n$, ce qui donne le système :

$$\begin{cases} x_0 + \lambda_1 x_1 + \cdots + \lambda_1^{n-1} x_{n-1} = \mu_1 \\ \dots \\ x_0 + \lambda_n x_1 + \cdots + \lambda_n^{n-1} x_{n-1} = \mu_n \end{cases}$$

Le déterminant de ce système est un déterminant de Vandermonde ; il est non nul car les valeurs propres de f sont distinctes. On peut donc trouver $x_0, \dots, x_{n-1}$ et le polynôme P .

3) Si λ est une valeur propre de f et x un vecteur propre associé, on a $f(g(x)) = g(f(x)) = \lambda g(x)$, ce qui prouve que $g(x)$ est dans l'espace propre de f associé à λ . En prenant $\lambda = 0$, on résout la dernière question. Ce résultat est très utile dans tous les problèmes où apparaissent des endomorphismes commutant entre eux.

$$\mathbf{15.3\ a)} F^2(f) = \frac{1}{4} (p \circ f + f \circ p + 2p \circ f \circ p)$$

$$F^3(f) = \frac{1}{8} (p \circ f + f \circ p + 6p \circ f \circ p)$$

b) On trouve $2F^3 - 3F^2 + F = 0$.

c) Comme $2X^3 - 3X^2 + X = X(2X - 1)(X - 1)$ est un polynôme scindé avec des racines simples, F est diagonalisable.

d) L'espace propre associé à la valeur propre 0 est l'espace des f de $L(E)$ tels que $p \circ f + f \circ p = 0$. Si $u \in \ker(p)$, on a $p(f(u)) = 0$ donc $f(u) \in \ker(p)$. Si $u \in \text{Im}(p)$, on a $u = p(u)$ donc $f(u) = f(p(u)) = -p(f(u))$ donc $f(u) = 0$. On en déduit que :

$$f(\ker(p)) \subset \ker(p) \text{ et } f(\text{Im}(p)) = \{0\} \quad (P)$$

Réciproquement, un endomorphisme f qui vérifie ces deux propriétés est dans le noyau de F car tout u de E s'écrit $u = u_1 + u_2$ avec $u_1 \in \ker(p)$ et $u_2 \in \text{Im}(p)$ (décomposition de E en somme directe $\ker(p) \oplus \text{Im}(p)$) et $(p \circ f + f \circ p)(u_1 + u_2) = 0$. L'espace propre associé à 0 est donc l'espace des fonctions f vérifiant (P) . Cet espace est de dimension $(n - r)^2$.

On montre de même que l'espace propre associé à la valeur propre 1 est l'espace des endomorphismes f vérifiant :

$$f(\ker(p)) = \{0\} \text{ et } f(\text{Im}(p)) \subset \text{Im}(p)$$

et que l'espace propre associé à la valeur propre $\frac{1}{2}$ est l'espace des endomorphismes f vérifiant :

$$f(\ker(p)) \subset \text{Im}(p) \text{ et } f(\text{Im}(p)) \subset \ker(p).$$

Ces deux espaces sont de dimensions respectives r^2 et $2r(n-r)$.
On peut vérifier que la somme de ces trois dimensions est bien n^2 .

15.4 a) Si λ est valeur propre de L , il existe une fonction f non nulle de E telle que $f'' = L(f) = \lambda f$. Si $\lambda = 0$, l'espace propre associé est celui des fonctions affines. Si $\lambda > 0$, l'espace propre associé est $\text{Vect}(f_1, f_2)$ avec $f_1 : x \mapsto e^{\sqrt{\lambda}x}$ et $f_2 : x \mapsto e^{-\sqrt{\lambda}x}$. Si $\lambda < 0$, l'espace propre associé est $\text{Vect}(f_1, f_2)$ avec $f_1 : x \mapsto \cos \sqrt{-\lambda}x$ et $f_2 : x \mapsto \sin \sqrt{-\lambda}x$.

b) Posons $g = L(f)$. On a $g(x) = \int_0^x tf(t)dt + x \int_x^1 f(t)dt$. On en déduit $g''(x) = -f(x)$. Donc $g = L(f)$ équivaut à $g'' = -f$, $g(0) = 0$ et $g'(1) = 0$.

Si λ est valeur propre de L et si f est un vecteur propre associé, on a $L(f) = \lambda f$ donc $\lambda f'' + f = 0$.

Comme f est non nulle, on ne peut avoir $\lambda = 0$.

Pour $\lambda > 0$, $\lambda f'' + f = 0$ donne $f(x) = a \cos \frac{x}{\sqrt{\lambda}} + b \sin \frac{x}{\sqrt{\lambda}}$ avec a, b réels. Les conditions $g(0) = 0$ et $g'(1) = 0$ imposent $a = 0$ et $\lambda = \frac{4}{(2k+1)^2 \pi^2}$, $k \in \mathbb{Z}$.

Chacun de ces nombres est valeur propre de L et l'espace propre associé est engendré par la fonction $x \mapsto \sin \frac{x}{\sqrt{\lambda}}$.

Pour $\lambda < 0$, $\lambda f'' + f = 0$ donne $f(x) = a \text{ch} \frac{x}{\sqrt{-\lambda}} + b \text{sh} \frac{x}{\sqrt{-\lambda}}$ avec a et b réels.

Les conditions $g(0) = 0$ et $g'(1) = 0$ imposent $a = 0$ et $b = 0$. Par conséquent, λ n'est pas valeur propre de L .

15.5 a) Si f est un vecteur propre de L associé à la valeur propre λ , on a $L(f)(x) = f(x-1) = \lambda f(x)$ pour tout réel x . On ne peut avoir $\lambda = 0$ car alors f serait la fonction nulle.

On a $L^n(f)(x) = f(x-n) = \lambda^n f(x)$. Choisissons un réel x tel que $f(x)$ soit non nul. Si $|\lambda| > 1$, $f(x-n) = \lambda^n f(x)$ tend vers l'infini avec n , ce qui est impossible.

Si $|\lambda| < 1$, on raisonne de même avec l'égalité $f(x+n) = \frac{1}{\lambda^n} f(x)$.

b) Les deux seules valeurs propres possibles sont donc 1 et -1 . L'espace propre associé à 1 est l'espace des fonctions périodiques de période 1. L'espace propre associé à -1 est l'espace des fonctions périodiques de période 2 telles que $f(x) = -f(x-1)$ pour $x \in [1, 2]$.

15.6 a, b) Dans le premier cas, la seule valeur propre possible pour A est 2. Si A était diagonalisable, on aurait $A = 2I$, ce qui est faux. Dans le second, la seule valeur propre possible pour A est 0 et on raisonne de la même façon.

c) Le noyau de f est de dimension $n - 1$; il n'est donc pas réduit à $\{0\}$ et ses éléments non nuls sont vecteurs propres de f associés à la valeur propre 0. La valeur propre 0 est une racine d'ordre $r \geq n - 1$ de $\chi(f)$. Si $r = n$, 0 est la seule valeur propre de f donc $\text{Tr}(f) = 0$ (la trace est la somme des valeurs propres). Comme l'espace propre associé à 0 est de dimension $n - 1$, f n'est pas diagonalisable. Si $r = n - 1$, le polynôme caractéristique de f a une valeur propre non nulle simple. On a $\text{Tr}(f) \neq 0$ et f est diagonalisable puisque la somme des dimensions des espaces propres est n .

15.7 Soit P une matrice à coefficients complexes telle que $A = P B P^{-1}$, autrement dit $AP = PB$. La matrice P s'écrit $P = P_1 + i P_2$ où P_1 et P_2 sont à coefficients réels mais ne sont peut-être inversibles ni l'une ni l'autre. En séparant les parties réelles et imaginaires, on voit que $AP_1 = P_1 B$ et $AP_2 = P_2 B$. Pour tout x réel, on a donc $A(P_1 + xP_2) = (P_1 + xP_2)B$. La fonction polynomiale $x \mapsto \det(P_1 + xP_2)$ n'est pas nulle puisque, pour $x = i$, $P_1 + xP_2$ est inversible. Elle a donc au plus n racines réelles. Pour tout x réel distinct de ces racines, $R = P_1 + xP_2$ est inversible et $AR = RB$ donc $A = R B R^{-1}$.

$$\mathbf{15.8 a)} \quad A^2 = \begin{pmatrix} 5 & -6 & 2 \\ 0 & 3 & 2 \\ -2 & 2 & 7 \end{pmatrix} \quad A^3 = \begin{pmatrix} 8 & -13 & 9 \\ 7 & -8 & -5 \\ 12 & -26 & -11 \end{pmatrix}$$

b) Les relations entre coefficients et racines du polynôme s'obtiennent en écrivant $\chi(A) = X^3 + aX^2 + bX + c = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$ et en identifiant. On trouve :

$$a = -(\lambda_1 + \lambda_2 + \lambda_3) ; \quad b = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 ; \quad c = -\lambda_1\lambda_2\lambda_3.$$

c) On a $a = -p_1$, $b = \frac{1}{2}(p_1^2 - p_2)$. La valeur de c s'obtient en montrant que $p_3 = p_2 - bp_1 - 3c$ (ces formules sont des cas particuliers des formules appelées formules de Newton).

d) On a : $\text{Tr}(A) = p_1 = 1$, $\text{Tr}(A^2) = p_2 = 15$, $\text{Tr}(A^3) = p_3 = -11$.

On en déduit $a = -1$, $b = -7$, $c = 11$.

Le théorème de Cayley-Hamilton donne $A^3 - A^2 - 7A + 11I = 0$. On en déduit

$$A(A^2 - A - 7I) = -11I \text{ d'où } A^{-1} = \frac{1}{11}(-A^2 + A + 7I). \text{ Finalement :}$$

$$A^{-1} = \frac{1}{11} \begin{pmatrix} 4 & 5 & -1 \\ 1 & 4 & -3 \\ 4 & -6 & -1 \end{pmatrix}.$$

15.9 1) Les polynômes P_i sont premiers entre eux deux à deux dans l'anneau $\mathbb{C}[X]$; on applique le théorème des noyaux.

2) Pour la même raison, le théorème chinois appliqué dans $\mathbb{C}[X]$ montre l'existence d'un polynôme T de $\mathbb{C}[X]$ satisfaisant les relations indiquées. Il existe donc des polynômes Q_i , $1 \leq i \leq r$ tels que $T(X) = \lambda_i + Q_i(X)P_i(X)$ pour $i = 1, \dots, r$.

3) Montrons que, pour tout $i = 1, \dots, r$, la restriction de $d = T(f)$ à E_i est une homothétie. Soit $x \in E_i$; on a $d(x) = T(f)(x) = [\lambda_i \text{id} + Q_i(f)P_i(f)](x) = \lambda_i x$, puisque $P_i(f)(x) = 0$. On en déduit que d est diagonalisable, que ses valeurs propres sont les λ_i et ses espaces propres les E_i .

4) On a construit d comme la valeur d'un polynôme de $\mathbb{C}[X]$ en f ; par conséquent, $n = f - d$ est également la valeur d'un polynôme de $\mathbb{C}[X]$ en f . Soit $x \in E_i$. On a $n(x) = (f - d)(x) = (f - \lambda_i \text{id})(x)$. Donc $n^{k_i}(x) = 0$. On en déduit que, si $K \geq k_i$ pour $i = 1, \dots, r$, on a $n^K(x) = 0$ pour tout x de E . L'endomorphisme n est nilpotent.

5) Pour montrer l'unicité du couple (d, n) , on raisonne par l'absurde en supposant l'existence d'un couple (d', n') ayant les mêmes propriétés que (d, n) .

a) Les endomorphismes d, d', n, n' étant tous des polynômes en f , commutent entre eux.

b) Comme on a $dd' = d'd$, l'exercice 15.2 montre qu'une base vecteurs propres pour d est une base de vecteurs propres pour d' , donc $d - d'$ est diagonalisable.

c) Supposons que $n^K = (n')^{K'} = 0$. Posons $L = K + K'$ et élevons $n' - n$ à la puissance L , la formule du binôme montre que $(n' - n)^L$ est une somme de termes de la forme $C_L^k n'^k n^{L-k}$ avec $k = 0, \dots, L$. Dans cette somme, soit $k \geq K'$ et alors $n'^k = 0$, soit $k \leq K'$ et alors $L - k \geq K$ et $n^{L-k} = 0$. Donc $(n' - n)^L = 0$, ce qui montre que $n' - n$ est nilpotent.

d) Comme $d' - d = n' - n$, l'endomorphisme $d' - d$ est également nilpotent. Comme $d' - d$ est diagonalisable, on a $d' - d = 0$, ce qui montre que $d' = d$, d'où $n' = n$.

16.1 INTRODUCTION

L'orthogonalité est une notion première de la géométrie. Depuis un siècle, elle joue aussi un rôle important pour traiter les problèmes d'analyse fonctionnelle, par exemple ceux posés par les physiciens. La géométrie s'introduit ainsi dans un domaine inattendu, elle apporte des idées visuelles d'orthogonalité là où les mathématiciens d'avant 1900 ne voyaient que des calculs.

La définition d'espace vectoriel, qui date de Peano (1888) n'a pas eu d'influence à l'époque. Quand on regarde ce qui s'est passé, on peut comprendre que parler de côtés perpendiculaires dans un triangle et envisager que des fonctions puissent entretenir une relation analogue sont deux idées très différentes et qu'il a fallu du temps, même à des mathématiciens de génie, pour apercevoir les liens entre les deux situations et l'intérêt de les penser de façon unifiée.

16.2 ORTHOGONALITÉ DANS LE PLAN ET L'ESPACE ORDINAIRES

Revenons d'abord au plan et à l'espace de la géométrie euclidienne en dimension 2 et 3 étudiée au lycée à partir de notions intuitives. Dans ce cadre, on définit les notions suivantes : distance de deux points, norme d'un vecteur, orthogonalité de

vecteurs, angle de vecteurs. Une base orthonormée du plan (de l'espace) est formée de deux (trois) vecteurs orthogonaux de norme 1.

Norme d'un vecteur. Dans le plan muni d'une base orthonormée, la norme d'un vecteur $u = (x, y)$ est $\|u\| = \sqrt{x^2 + y^2}$, d'après le théorème de Pythagore. Dans l'espace muni d'une base orthonormée, la norme d'un vecteur $u = (x, y, z)$ est $\|u\| = \sqrt{x^2 + y^2 + z^2}$.

Proposition 1 : condition d'orthogonalité.

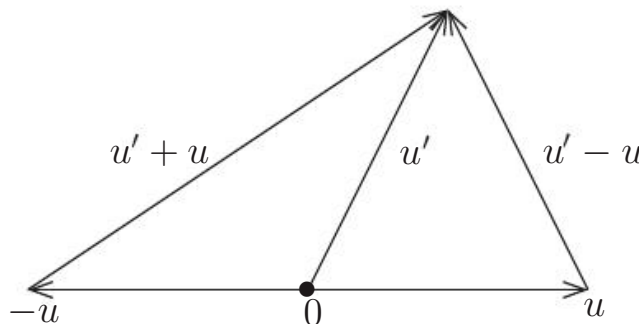
1) Dans le plan euclidien muni d'une base orthonormée, deux vecteurs $u = (x, y)$ et $u' = (x', y')$ sont orthogonaux si et seulement si :

$$xx' + yy' = 0.$$

2) Dans l'espace euclidien ordinaire muni d'une base orthonormée, deux vecteurs $u = (x, y, z)$ et $u' = (x', y', z')$ sont orthogonaux si et seulement si :

$$xx' + yy' + zz' = 0.$$

Démonstration. Dans les deux cas, considérons la figure suivante.



1) Le vecteur u' est orthogonal au vecteur u si et seulement si les vecteurs $u' + u$ et $u' - u$ sont de même norme, ce qui donne dans le plan :

$$\sqrt{(x' + x)^2 + (y' + y)^2} = \sqrt{(x' - x)^2 + (y' - y)^2}.$$

Cette condition est équivalente à l'égalité des carrés :

$$(x' + x)^2 + (y' + y)^2 = (x' - x)^2 + (y' - y)^2$$

qui donne la condition de l'énoncé en développant. Géométriquement, on peut dire qu'un parallélogramme est un rectangle si et seulement si ses diagonales sont égales.

On procède de même pour le 2). □

Définition : produit scalaire de deux vecteurs. Soient $u = (x, y)$ et $u' = (x', y')$ deux vecteurs du plan muni d'une base orthonormée. On appelle produit scalaire de u et de u' le scalaire $\langle u, u' \rangle = xx' + yy'$.

De même, soient $u = (x, y, z)$ et $u' = (x', y', z')$ deux vecteurs de l'espace muni d'une base orthonormée. On appelle produit scalaire de u et de u' le scalaire $\langle u, u' \rangle = xx' + yy' + zz'$.

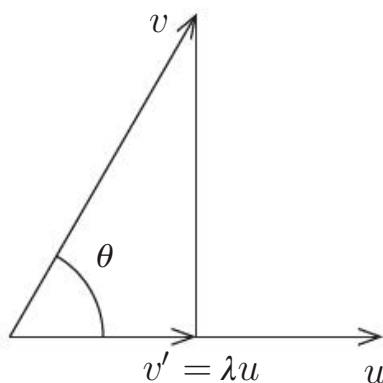
La proposition précédente s'énonce donc de la façon suivante.

Proposition 1'. *Deux vecteurs du plan ou de l'espace sont orthogonaux si et seulement si leur produit scalaire est nul.* $\square$

La norme d'un vecteur vaut $\|u\| = \langle u, u \rangle^{1/2}$, ce qui équivaut à $\|u\|^2 = \langle u, u \rangle$.

Projection orthogonale d'un vecteur sur un axe. Soient u et v deux vecteurs du plan ou de l'espace. On suppose $u \neq 0$. La projection orthogonale de v sur u (ou sur l'axe engendré par u) est de la forme λu . Pour trouver λ , écrivons que $v - \lambda u$ est orthogonal à u . L'égalité $\langle v - \lambda u, u \rangle = 0$ donne $\langle v, u \rangle - \lambda \langle u, u \rangle = 0$. Comme $\langle u, u \rangle \neq 0$, on obtient $\lambda = \frac{\langle v, u \rangle}{\langle u, u \rangle}$, où λ est positif, négatif ou nul suivant les cas.

La projection orthogonale v' de v sur u est donc $v' = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$.



Cosinus d'un angle de vecteurs. Soient u et v deux vecteurs non nuls du plan ou de l'espace. Considérons la projection orthogonale $v' = \lambda u$ de v sur u . L'angle non orienté θ que forment u et v est un angle compris entre 0 et π . Son cosinus est $\frac{\lambda \|u\|}{\|v\|}$ soit :

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

L'ordre de u et v n'importe pas pour cette définition. Le cosinus est nul si et seulement si les deux vecteurs sont orthogonaux.

16.3 PRODUIT SCALAIRE

Le produit scalaire défini en 16.2 est une fonction $\varphi : E \times E \rightarrow \mathbb{R}$ (où E désigne le plan ou l'espace). Cette fonction possède des propriétés à l'origine des définitions suivantes.

Définition 1 : forme bilinéaire, forme bilinéaire symétrique. Soient K un corps, E un K -espace vectoriel, $\varphi : E \times E \rightarrow K$ une application. On dit que φ est bilinéaire si :

pour tous u, u', v, v' de E et tout λ de $\mathbb{R}$:

$$\varphi(u + u', v) = \varphi(u, v) + \varphi(u', v)$$

$$\varphi(u, v + v') = \varphi(u, v) + \varphi(u, v')$$

$$\varphi(\lambda u, v) = \lambda \varphi(u, v) = \varphi(u, \lambda v).$$

C'est le cas $p = 2$ de la définition 1 de 14.4.

On dit que φ est bilinéaire symétrique si :

pour tous u, v de E :

$$\varphi(u, v) = \varphi(v, u)$$

Définition 2 : forme bilinéaire symétrique définie, positive. Soient E un $\mathbb{R}$ -espace vectoriel, $\varphi : E \times E \rightarrow \mathbb{R}$ une forme bilinéaire symétrique. On dit que φ est positive si :

pour tout u de E :

$$\varphi(u, u) \geq 0.$$

On dit que φ est définie si :

$$\varphi(u, u) = 0 \text{ équivaut à } u = 0.$$

Définition 3 : produit scalaire. Soit E un $\mathbb{R}$ -espace vectoriel. On appelle produit scalaire sur E , une forme bilinéaire symétrique définie positive $\varphi : E \times E \rightarrow \mathbb{R}$.

On notera le produit scalaire de deux vecteurs $\langle u, v \rangle$ sauf quand on aura besoin de l'application φ .

Produit scalaire sur un sous-espace. Soit F un sous-espace de E . On définit un produit scalaire sur F en prenant la restriction à $F \times F$ du produit scalaire défini sur $E \times E$.

Définition 4 : espace euclidien et préhilbertien réel. Un $\mathbb{R}$ -espace vectoriel muni d'un produit scalaire est appelé :

- espace euclidien s'il est de dimension finie ;
- espace préhilbertien réel autrement.

Définition 5 : vecteurs orthogonaux. Dans un espace muni d'un produit scalaire, deux vecteurs u et v sont dits orthogonaux si $\langle u, v \rangle = 0$.

Exemples.

- 1) L'exemple de base est celui de $\mathbb{R}^n$ muni de sa base canonique. On définit un produit scalaire en posant, pour $u = (x_1, \dots, x_n)$, $v = (y_1, \dots, y_n)$:

$$\langle u, v \rangle = \sum_{1 \leq i \leq n} x_i y_i.$$

Cet exemple généralise les produits scalaires sur $\mathbb{R}^2$ et $\mathbb{R}^3$ vus ci-dessus. En particulier, on retrouve la condition d'orthogonalité de 16.2.

On peut généraliser et définir une structure euclidienne sur un $\mathbb{R}$ -espace vectoriel E de dimension finie n muni d'une base B en posant $\varphi(u, v) = \sum_{1 \leq i \leq n} x_i y_i$, pour $u = (x_1, \dots, x_n)$ et $v = (y_1, \dots, y_n)$.

- 2) L'exemple précédent conduit à définir un produit scalaire sur l'espace des matrices carrées d'ordre n à coefficients réels : le produit scalaire des matrices $A = (a_{ij})$ et $B = (b_{ij})$ est $\langle A, B \rangle = \sum_{1 \leq i, j \leq n} a_{ij} b_{ij}$. On vérifiera que cette somme est égale à $\text{Tr}(AB)$.
- 3) Soient $I = [a, b]$ un intervalle fermé borné de $\mathbb{R}$ et $E = \text{Cont}(I, \mathbb{R})$ l'espace des fonctions continues de I dans $\mathbb{R}$. On définit un produit scalaire sur E en posant $\langle f, g \rangle = \int_I f(t)g(t) dt$. Ce n'est qu'un des nombreux exemples de produit scalaire sur un espace de fonctions. La dimension de l'espace est ici infinie.

Par exemple, pour $I = [0, 2\pi]$, les fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$ sont orthogonales dans $\text{Cont}([0, 2\pi], \mathbb{R})$ car $\int_0^{2\pi} \sin t \cos t dt = 0$.

C'est Schmidt qui, en 1905, introduit dans sa thèse l'adjectif orthogonal pour qualifier deux fonctions dont le produit scalaire est nul.

16.4 EXPRESSION DU PRODUIT SCALAIRE

Expression du produit scalaire dans une base. Soit E un espace euclidien de dimension n et soit $B = (e_1, \dots, e_n)$ une base de E . Soient $u = \sum_{1 \leq i \leq n} x_i e_i$ et $v = \sum_{1 \leq i \leq n} y_i e_i$ deux vecteurs de E . Le produit scalaire de u et v s'exprime en fonction des produits scalaires $\langle e_i, e_j \rangle$ des vecteurs de base grâce à la bilinéarité :

$$\langle u, v \rangle = \left\langle \sum_{1 \leq i \leq n} x_i e_i, \sum_{1 \leq j \leq n} y_j e_j \right\rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle.$$

En particulier, on a :

$$\langle u, u \rangle = \sum_{1 \leq i, j \leq n} x_i x_j \langle e_i, e_j \rangle.$$

Pour tenir compte de la symétrie du produit scalaire qui donne $\langle e_i, e_j \rangle = \langle e_j, e_i \rangle$ pour $1 \leq i, j \leq n$, on peut encore écrire :

$$\begin{aligned}\langle u, v \rangle &= \sum_{1 \leq i \leq n} x_i y_i \langle e_i, e_i \rangle + \sum_{1 \leq i < j \leq n} (x_i y_j + x_j y_i) \langle e_i, e_j \rangle \\ \langle u, u \rangle &= \sum_{1 \leq i \leq n} x_i^2 \langle e_i, e_i \rangle + 2 \sum_{1 \leq i < j \leq n} x_i x_j \langle e_i, e_j \rangle.\end{aligned}$$

En posant $a_{ij} = \langle e_i, e_j \rangle$ pour $1 \leq i, j \leq n$, on a donc :

$$\begin{aligned}\langle u, v \rangle &= \sum_{1 \leq i \leq n} a_{ii} x_i y_i + \sum_{1 \leq i < j \leq n} a_{ij} (x_i y_j + x_j y_i) \\ \langle u, u \rangle &= \sum_{1 \leq i \leq n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j\end{aligned}$$

Représentation matricielle du produit scalaire dans un espace euclidien. Soit E un espace euclidien de dimension n et soit $B = (e_1, \dots, e_n)$ une base de E . Soient $u = \sum_{1 \leq i \leq n} x_i e_i$ et $v = \sum_{1 \leq i \leq n} y_i e_i$ deux vecteurs de E et U et V les matrices colonnes associées. On appelle matrice associée au produit scalaire φ de E dans la base B la matrice $\Phi = (a_{ij})$ avec $a_{ij} = \langle e_i, e_j \rangle$.

Proposition. La matrice Φ est symétrique, inversible et :

$$\langle u, v \rangle = {}^t U \Phi V = {}^t V \Phi U$$

Démonstration. On vérifie que ${}^t U \Phi V = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle = \langle u, v \rangle$ et que ${}^t ({}^t V \Phi U) = {}^t U \Phi V$. Pour tout $u \neq 0$, on a ${}^t U \Phi U = \langle u, u \rangle$, donc $\Phi U \neq 0$; on en déduit que Φ est inversible. $\square$

Le produit ${}^t U \Phi V$ est une matrice carrée d'ordre 1 dont on n'écrira jamais, comme on l'a dit en 8.2, les parenthèses.

Exemple. Dans $\mathbb{R}^n$ muni de la base canonique, la matrice du produit scalaire défini en 16.3, exemple 1, est I_n .

Changement de base. Soit E un espace euclidien de dimension n et soient $B = (e_1, \dots, e_n)$ et $B' = (e'_1, \dots, e'_n)$ deux bases de E . Soient $u = \sum_{1 \leq i \leq n} x_i e_i$ et $v = \sum_{1 \leq i \leq n} y_i e_i$ deux vecteurs de E , U et V les matrices colonnes associées dans la base B , U' et V' les matrices colonnes associées dans la base B' . Posons $\Phi = (\langle e_i, e_j \rangle)$ et $\Phi' = (\langle e'_i, e'_j \rangle)$. Si on pose $P = M(\text{id}, B', B)$, on a $U = P U'$ et $V = P V'$ donc ${}^t U \Phi V = {}^t U' {}^t P \Phi P V'$. Cette formule étant valable pour tout u, v de E , on a la formule de changement de base :

$$\Phi' = {}^t P \Phi P$$

On notera la différence avec la formule de changement de base pour les applications linéaires.

Reconnaître un produit scalaire. Reprenons les notations précédentes. Nous allons exposer une méthode pour reconnaître si une application $\varphi : E \times E \rightarrow \mathbb{R}$ définie par $\varphi(u, v) = \sum_{1 \leq i \leq n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} a_{ij} (x_i y_j + x_j y_i)$ est un produit scalaire sur E ou non.

Une telle application est bilinéaire et symétrique : on peut facilement s'en assurer.

La difficulté est de savoir si φ est définie positive. Le problème revient donc à l'étude de $\varphi(u, u) = \sum_{1 \leq i \leq n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$. Une telle expression est appelée forme quadratique ; nous approfondirons l'étude de ces formes au chapitre 22.

Exemples. Prenons d'abord quelques exemples de formes φ définies sur $\mathbb{R}^3$. Si $u = (x, y, z)$, on a $\varphi(u, u) = ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz$ avec a, b, c, d, e, f réels.

- Si $\varphi(u, u) = x^2 + 3y^2 + 5z^2$, il est clair que $\varphi(u, u) \geq 0$ pour tout u de $\mathbb{R}^3$ et que $\varphi(u, u) = 0$ si et seulement si $x = y = z = 0$ donc si et seulement si $u = 0$: φ définit un produit scalaire.
- Si $\varphi(u, u) = x^2 - 3y^2 + 5z^2$, la forme φ n'est pas définie positive : pour $u = (0, 1, 0)$, on a $\varphi(u, u) < 0$.
- Si $\varphi(u, u) = x^2 + 3y^2$, la forme φ n'est pas définie positive : pour $u = (0, 0, 1)$, on a $\varphi(u, u) = 0$.
- Si $\varphi(u, u) = 3xy + xz$, la forme φ n'est pas définie positive : pour $u = (1, -1, 0)$, on a $\varphi(u, u) < 0$.

Méthode de Gauss. La méthode de Gauss consiste à examiner d'abord si $a_{ii} > 0$ pour tout i , $1 \leq i \leq n$. Si ce n'est pas le cas, par exemple si $a_{11} \leq 0$, on a $\varphi(u, u) \leq 0$ pour $u = (1, 0, \dots, 0)$ et φ n'est pas définie positive.

Si $a_{ii} > 0$ pour tout i , $1 \leq i \leq n$, on transforme l'expression de $\varphi(u, u)$. On considère d'abord tous les termes où x_1 apparaît : en mettant a_{11} en facteur c'est : $f_1(u) = x_1 + \sum_{1 \leq i \leq n} \frac{a_{1i}}{a_{11}} x_i$. On peut écrire $\varphi(u, u) = a_{11} (f_1(u))^2 + \varphi_1(u, u)$; $\varphi_1(u, u)$ est une forme quadratique où x_1 n'apparaît pas et on applique la même méthode avec les termes en x_2 ; on recommence jusqu'à ce qu'il ne reste aucun terme.

On obtient finalement une décomposition $\varphi(u, u) = \alpha_1 f_1(u)^2 + \alpha_2 f_2(u)^2 + \dots + \alpha_p f_p(u)^2$, où apparaissent des carrés de formes linéaires $f_1(u) = x_1 + \dots$, $f_2(u) = x_2 + \dots$, etc.

- Si $p = n$ et si, pour $1 \leq i \leq n$, on a $\alpha_i > 0$, alors φ est définie positive car $\varphi(u, u) = 0$ équivaut à un système triangulaire de n équations :

$$\begin{cases} f_1(u) = 0 \\ \dots \\ f_n(u) = 0 \end{cases}$$

qui admet $u = 0$ comme seule solution.

- Si $p < n$, le système précédent est de rang inférieur à n et admet des solutions non nulles, donc φ n'est pas un produit scalaire.
- Si $p = n$ et s'il existe i , $1 \leq i \leq n$ tel que $\alpha_i < 0$, il existe un vecteur u solution des équations $f_j(u) = 0$, $j \neq i$ et tel que $f_i(u) \neq 0$. On a $\varphi(u, u) < 0$ et φ n'est pas un produit scalaire.

Exemple. Considérons la forme φ définie par $\varphi(u, u) = 2x^2 + y^2 + 2z^2 - 4xy - 2yz + 4xz$, la méthode de Gauss donne $\varphi(u, u) = 2(x - y + z)^2 - (y - z)^2 + z^2$. Pour $u = (1, 1, 0)$, on a $\varphi(u, u) = 2 \times 0^2 - 1 + 0^2 < 0$, donc φ n'est pas un produit scalaire.

16.5 NORME ET ANGLE

Dans les espaces euclidiens et préhilbertiens réels, on généralise les notions de norme et d'angle.

Définition 1 : norme d'un vecteur. Dans un espace vectoriel muni d'un produit scalaire, on appelle norme du vecteur u le scalaire $\|u\| = \langle u, u \rangle^{1/2}$.

Exemples. Donnons les normes associées aux produits scalaires des premier et troisième exemples de 16.3.

- 1) Dans $\mathbb{R}^n$, si $u = (x_1, \dots, x_n)$, on a :

$$\|u\| = \left(\sum_{1 \leq i \leq n} x_i^2 \right)^{1/2}$$

- 3) Dans $E = \text{Cont}(I, \mathbb{R})$ la norme de $f: I \rightarrow \mathbb{R}$ est :

$$\|f\| = \left(\int_I f(t)^2 dt \right)^{1/2}$$

Proposition 1 : inégalité de Cauchy-Bunyakovski-Schwarz (CBS).

Soit E un $\mathbb{R}$ -espace vectoriel muni d'un produit scalaire

- 1) Pour tous u, v de E , on a :

$$\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2 \quad \text{ou} \quad |\langle u, v \rangle| \leq \|u\| \|v\|$$

2) *Cas d'égalité* : pour tous u, v de E : $\langle u, v \rangle^2 = \|u\|^2 \|v\|^2$ si et seulement si u et v sont colinéaires.

Démonstration. L'inégalité et le cas d'égalité sont évidents pour $u = 0$. Supposons $u \neq 0$.

1) Pour tout réel λ , on a :

$$0 \leq \|v + \lambda u\|^2 = \langle v + \lambda u, v + \lambda u \rangle = \|v\|^2 + 2\lambda \langle v, u \rangle + \lambda^2 \|u\|^2$$

Le trinôme du second degré en λ est positif ou nul pour toutes les valeurs de λ . Son discriminant est donc négatif ou nul, ce qui donne l'inégalité.

2) Notons v' la projection orthogonale de v sur u . On a vu que $v' = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$, donc :

$$\|v - v'\|^2 = \langle v - \frac{\langle v, u \rangle}{\langle u, u \rangle} u, v - \frac{\langle v, u \rangle}{\langle u, u \rangle} u \rangle = \|v\|^2 - \frac{\langle v, u \rangle^2}{\|u\|^2} = 0$$

Par conséquent, $v - v' = 0$ et v est colinéaire à u . □

L'inégalité précédente est souvent appelée inégalité de Cauchy-Schwarz. C'est plus court, mais c'est injuste pour Viktor Bunyakovski (1804-1889) qui a travaillé avec Cauchy en 1825 et a publié en 1859, 25 ans avant Hermann Schwarz (1843-1921), des résultats sur la forme fonctionnelle de l'inégalité.

Exemples. Donnons les inégalités associées aux produits scalaires des premier et troisième exemples donnés en 16.3.

1) Dans $\mathbb{R}^n$, étant donnés $u = (x_1, \dots, x_n)$ et $v = (y_1, \dots, y_n)$, on a :

$$\left(\sum_{1 \leq i \leq n} x_i y_i \right)^2 \leq \left(\sum_{1 \leq i \leq n} x_i^2 \right) \left(\sum_{1 \leq i \leq n} y_i^2 \right)$$

C'est cette inégalité que montre Cauchy dans une note de son cours d'analyse de 1821.

2) Dans $E = \text{Cont}(I, \mathbb{R})$, étant données $f, g : I \rightarrow \mathbb{R}$, on a :

$$\left(\int_I f(t)g(t) dt \right)^2 \leq \left(\int_I f(t)^2 dt \right) \left(\int_I g(t)^2 dt \right)$$

Proposition 2 : propriétés de la norme.

- Pour tout u de E , on a : $\|u\| \geq 0$
- $\|u\| = 0$ équivaut à $u = 0$.
- Pour tout u de E et tout λ de $\mathbb{R}$, on a : $\|\lambda u\| = |\lambda| \|u\|$.

➤ *Inégalité triangulaire*

Pour tous u, v de E , on a :

$$\|u + v\| \leq \|u\| + \|v\|.$$

➤ *Cas d'égalité de l'inégalité triangulaire*

Si $\|u + v\| = \|u\| + \|v\|$, il existe $\lambda > 0$ tel que $u = \lambda v$ ou $v = \lambda u$.

➤ *Relation entre norme et produit scalaire*

$$2 \langle u, v \rangle = \|u + v\|^2 - \|u\|^2 - \|v\|^2.$$

Démonstration. Les premières propriétés sont faciles à vérifier. La dernière relation résulte de $\|u + v\|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + 2 \langle u, v \rangle + \langle v, v \rangle$.

L'inégalité triangulaire résulte de l'inégalité de CBS :

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2 \|u\| \|v\| + \|v\|^2 \end{aligned}$$

L'égalité a lieu si on est dans le cas d'égalité de l'inégalité de CBS et si $\langle u, v \rangle = |\langle u, v \rangle|$, d'où le résultat. $\square$

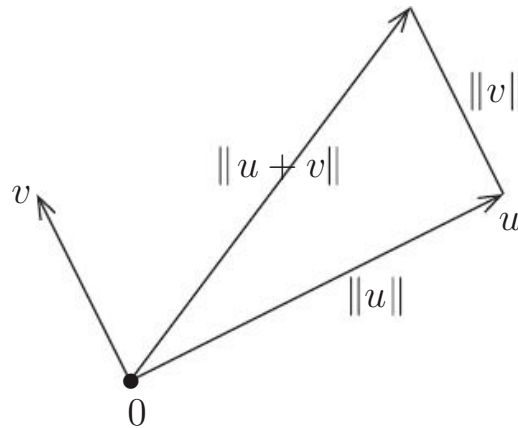
L'inégalité triangulaire généralise l'inégalité bien connue entre la longueur d'un côté d'un triangle et la somme des longueurs des deux autres côtés.

Un autre résultat classique se généralise. Il porte le nom de Pythagore, mais des mathématiciens babyloniens le connaissaient déjà 1 200 ans avant lui, vers – 1750 (voir exercice 12.6). D'ailleurs la renommée de Pythagore en mathématique n'est mentionnée que dans des textes anciens largement postérieurs à lui et il est probable que l'un des seuls mathématiciens connus des élèves du secondaire le soit à tort ; peut-être n'a-t-il été que le chef d'une secte observant des interdictions plus que loufoques (pour nous).

Théorème de Pythagore. Dans un espace muni d'un produit scalaire, deux vecteurs u et v sont orthogonaux si et seulement si :

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Démonstration. Cela résulte de $\|u + v\|^2 = \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2$. $\square$



Définition 2 : angle de deux vecteurs. Dans un espace muni d'un produit scalaire, on définit l'angle (non orienté) de deux vecteurs u et v non nuls comme le réel θ de $[0, \pi]$ tel que :

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

ce qui est possible car ce rapport est dans $[-1, 1]$ d'après l'inégalité de CBS.

Quand deux vecteurs sont orthogonaux, leur produit scalaire, donc leur cosinus, est nul ; leur angle est donc $\frac{\pi}{2}$. La réciproque est immédiate.

Cette expression du cosinus, dans le cas particulier de l'espace $\mathbb{R}^n$, apparaît sous sa forme analytique dans les travaux de Jordan de 1873.

16.6 BASES ORTHOGONALES ET ORTHONORMÉES

Définition 1 : famille orthogonale. Dans un espace E muni d'un produit scalaire, une famille $(e_1, \dots, e_k)$ est dite orthogonale si $\langle e_i, e_j \rangle = 0$ pour tous i, j avec $1 \leq i, j \leq k$ et $i \neq j$.

Proposition 1. Dans un espace E muni d'un produit scalaire, une famille orthogonale $(e_1, \dots, e_k)$ sans vecteur nul est une famille libre.

Démonstration. Supposons qu'on ait $0 = \lambda_1 e_1 + \dots + \lambda_k e_k$; pour tout $i, 1 \leq i \leq k$, on a $0 = \langle \lambda_1 e_1 + \dots + \lambda_k e_k, e_i \rangle = \lambda_i \|e_i\|^2$, donc $\lambda_i = 0$. $\square$

Cette proposition donne une nouvelle méthode pour montrer la liberté d'une famille.

Définition 2 : bases orthogonales et orthonormées. Dans un espace euclidien E de dimension n , une base $B = (e_1, \dots, e_n)$ est dite :

- orthogonale si $\langle e_i, e_j \rangle = 0$ pour tous i, j avec $1 \leq i, j \leq n$ et $i \neq j$;
- orthonormée si elle est orthogonale et si $\|e_i\| = 1$ pour tout i , $1 \leq i \leq n$.

Si $B = (e_1, \dots, e_n)$ est une base orthogonale, la base $\left(\frac{e_1}{\|e_1\|}, \dots, \frac{e_n}{\|e_n\|} \right)$ est orthonormée.

Proposition 2 : calculs dans une base orthonormée. Soit E un espace euclidien de dimension n et soit $B = (e_1, \dots, e_n)$ une base orthonormée de E . Soient $u = \sum_{1 \leq i \leq n} x_i e_i$ et $v = \sum_{1 \leq i \leq n} y_i e_i$ deux vecteurs de E et U et V les matrices colonnes associées.

- 1) $x_k = \langle u, e_k \rangle$ pour $1 \leq k \leq n$.
- 2) $\|u\| = \left(\sum_{1 \leq i \leq n} x_i^2 \right)^{1/2}$.
- 3) $\langle u, v \rangle = \sum_{1 \leq i \leq n} x_i y_i$.
- 4) La matrice du produit scalaire est I_n et $\langle u, v \rangle = {}^t U V = {}^t V U$.

Démonstration. 1) $\langle u, e_k \rangle = \langle \sum_{1 \leq i \leq n} x_i e_i, e_k \rangle = x_k$.

$$2) \|u\|^2 = \langle \sum_{1 \leq i \leq n} x_i e_i, \sum_{1 \leq i \leq n} x_i e_i \rangle = \sum_{1 \leq i \leq n} x_i^2.$$

$$3) \langle u, v \rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle = \sum_{1 \leq i \leq n} x_i y_i.$$

$$4) \text{ La matrice du produit scalaire est } (\langle e_i, e_j \rangle) = I_n.$$

□

Construction de bases orthogonales : procédé d'orthogonalisation de Gram-Schmidt. Schmidt présente cette méthode dans le cadre du produit scalaire de l'exemple 3 donné en 16.3. Ses idées rencontrent celles d'un article de Jorgen Gram (1850-1916) de 1881 qui étudiait la méthode des moindres carrés.

Soit E un espace euclidien de dimension n et soit $B = (e_1, \dots, e_n)$ une base de E . Le procédé d'orthogonalisation de Gram-Schmidt permet de construire à partir de la base B une base orthogonale $C = (\varepsilon_1, \dots, \varepsilon_n)$ telle que

$$\text{Vect}(e_1, \dots, e_i) = \text{Vect}(\varepsilon_1, \dots, \varepsilon_i) \text{ pour } 1 \leq i \leq n.$$

On va construire C par récurrence sur i .

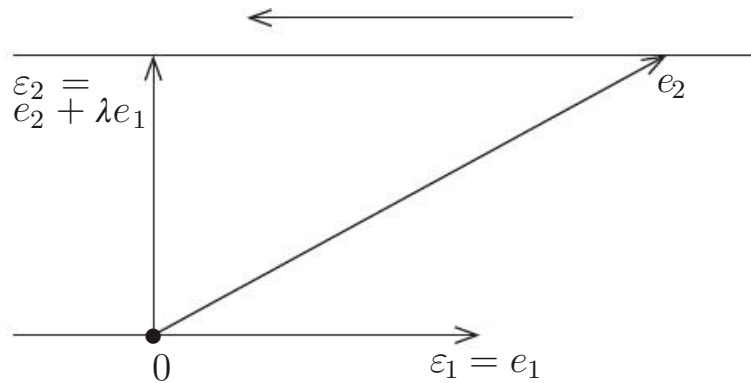
On pose d'abord $\varepsilon_1 = e_1$.

On construit ε_2 dans $\text{Vect}(e_1, e_2)$ en cherchant ε_2 sous la forme $e_2 + \lambda \varepsilon_1$ et en déterminant λ pour que ε_2 soit orthogonal à ε_1 . On doit avoir $\langle \varepsilon_2, \varepsilon_1 \rangle = 0$, d'où

$$\lambda = -\frac{\langle e_2, \varepsilon_1 \rangle}{\|\varepsilon_1\|^2}.$$

Illustrons ce calcul par un dessin : la flèche suggère le déplacement

de l'extrémité de e_2 quand λ varie pour obtenir un vecteur orthogonal à ε_1 .



Supposons la famille $(\varepsilon_1, \dots, \varepsilon_n)$ construite jusqu'au rang i . On cherche ε_{i+1} sous la forme $\varepsilon_{i+1} = e_{i+1} + \sum_{1 \leq k \leq i} \lambda_k \varepsilon_k$. Pour $1 \leq k \leq i$, la condition d'orthogonalité de

ε_{i+1} et de ε_k donne $\lambda_k = \frac{\langle e_{i+1}, \varepsilon_k \rangle}{\|\varepsilon_k\|^2}$. On détermine donc ainsi ε_{i+1} et on vérifie qu'il

satisfait aux conditions indiquées. Le calcul de ε_{i+1} à partir $(\varepsilon_1, \dots, \varepsilon_i)$ et de e_{i+1} peut s'illustrer par un dessin analogue au précédent.

On déduit immédiatement de cette construction qu'une famille orthogonale sans vecteur nul d'un espace euclidien E peut être complétée en une base orthogonale de E .

Polynômes trigonométriques. Prenons $I = [0, 2\pi]$, $E = \text{Cont}(I, \mathbb{R})$ dans l'exemple 3 de 16.3 ; le produit scalaire de deux éléments de E est défini par $\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt$. Posons $s_n : x \mapsto \sin nx$ pour $n > 0$, $c_n : x \mapsto \cos nx$ pour $n \geq 0$ (c_0 est la fonction constante 1).

Des calculs d'intégrales simples (il faut se rappeler les formules de trigonométrie) montrent les relations d'orthogonalité :

- $\langle s_n, s_m \rangle = 0$ pour $m \neq n$;
- $\langle c_n, c_m \rangle = 0$ pour $m \neq n$;
- $\langle s_n, c_m \rangle = 0$ pour tous m, n ;
- $\langle s_n, s_n \rangle = \pi$ pour tout $n > 0$;
- $\langle c_n, c_n \rangle = \pi$ pour tout $n > 0$ et $\langle c_0, c_0 \rangle = 2\pi$.

Avec le résultat de l'exercice 16.2 a), on voit que la famille des s_n et des c_n forme une base orthogonale de l'espace F , appelé espace des polynômes trigonométriques, qu'elle engendre. En divisant ces vecteurs par leurs normes, on obtient une base orthonormée B de F :

$$B = \left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left(\frac{1}{\sqrt{\pi}} c_n \right)_{n>0} \cup \left(\frac{1}{\sqrt{\pi}} s_n \right)_{n>0}.$$

Notons $e_0 = \frac{1}{\sqrt{2\pi}}$, $e_n = \frac{1}{\sqrt{\pi}} c_n$, $e'_n = \frac{1}{\sqrt{\pi}} s_n$, pour $n > 0$, les vecteurs de la base B .

Si $f \in E$, on définit, pour $n \geq 0$, les coefficients de Fourier de f comme les produits scalaires de f avec les vecteurs de la base B :

$$a_0 = \langle e_0, f \rangle = \frac{1}{\sqrt{2\pi}} \langle 1, f \rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t) dt ;$$

$$a_n = \langle e_n, f \rangle = \frac{1}{\sqrt{\pi}} \langle c_n, f \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} \cos nt f(t) dt \text{ pour } n > 0 ;$$

$$a'_n = \langle e'_n, f \rangle = \frac{1}{\sqrt{\pi}} \langle s_n, f \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} \sin nt f(t) dt \text{ pour } n > 0.$$

Si $f = a_0 e_0 + \sum_{n \geq 1} a_n e_n + \sum_{n \geq 1} a'_n e'_n \in F$, on a donc, d'après la proposition 2 de 16.6 :

$$\|f\|^2 = a_0^2 + \sum_{n \geq 1} a_n^2 + \sum_{n \geq 1} a_n'^2.$$

où les différentes sommes n'ont qu'un nombre fini de termes non nuls.

Si $g \in F$ s'écrit $g = b_0 e_0 + \sum_{n \geq 1} b_n e_n + \sum_{n \geq 1} b'_n e'_n$, la proposition 2 de 16.6 donne :

$$\langle f, g \rangle = a_0 b_0 + \sum_{n \geq 1} a_n b_n + \sum_{n \geq 1} a'_n b'_n.$$

Ces formules sont dites formules de Parseval, du nom de Marc-Antoine Parseval des Chênes (1755-1836) qui en publia une première forme en 1799.

Quand f est un élément quelconque de E , on peut toujours calculer les coefficients $a_n, n \geq 0, a'_n, n > 0$ par les formules précédentes. On les appelle encore coefficients de Fourier de f et on définit la série $S(f)$, qu'on appelle série de Fourier de f , par :

$$S(f) = \frac{1}{\sqrt{2\pi}} a_0 + \frac{1}{\sqrt{\pi}} \sum_{n \geq 1} a_n c_n + \frac{1}{\sqrt{\pi}} \sum_{n \geq 1} a'_n s_n.$$

L'étude des conditions de convergence de la série $S(f)$ dans des cas plus généraux que celui des fonctions continues est délicat et difficile et on entre là dans un domaine extrêmement vaste où l'algèbre et l'analyse s'entremêlent.

16.7 ORTHOGONALITÉ DE SOUS-ESPACES

Définition : orthogonal d'un sous-espace. Soit E un espace euclidien et soit F un sous-espace de E . Il est facile de vérifier que l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F est un sous-espace vectoriel de E . On le note $F^\perp$.

On peut définir également le sous-espace orthogonal à une partie A de E comme l'espace $(\text{Vect}(A))^\perp$.

On vérifie facilement que si $F = \text{Vect}(u_1, \dots, u_k)$, on a $u \in F^\perp$ si et seulement si $\langle u, u_i \rangle = 0$ pour $1 \leq i \leq k$.

Proposition : propriétés de l'orthogonal. Soit E un espace euclidien et soit F un sous-espace de E . On a :

- 1) $E = F \oplus F^\perp$.
- 2) $\dim(F) + \dim(F^\perp) = \dim(E)$.
- 3) Soit G un sous-espace vectoriel de E ; si $G \subset F$ alors $F^\perp \subset G^\perp$.
- 4) $(F^\perp)^\perp = F$.

Démonstration. 1) Complétons une base $(e_1, \dots, e_k)$ de F en une base $(e_1, \dots, e_n)$ de E et appliquons le procédé d'orthogonalisation de Gram-Schmidt. On obtient une base orthogonale $B = (\varepsilon_1, \dots, \varepsilon_n)$ de E telle que $(\varepsilon_1, \dots, \varepsilon_k)$ soit une base orthogonale de F . Soit u un vecteur de E ; dans la base B , il s'écrit : $u = \sum_{1 \leq i \leq n} x_i \varepsilon_i$. Il appartient à $F^\perp$ si et seulement si $\langle u, \varepsilon_i \rangle = 0$ pour $1 \leq i \leq k$. Comme $\langle u, \varepsilon_i \rangle = x_i$, cette dernière condition équivaut à $x_1 = \dots = x_k = 0$, c'est-à-dire $u \in \text{Vect}(\varepsilon_{k+1}, \dots, \varepsilon_n)$. On en déduit que $F^\perp = \text{Vect}(\varepsilon_{k+1}, \dots, \varepsilon_n)$, d'où $E = F \oplus F^\perp$.

- 2) C'est une conséquence immédiate du 1).
- 3) Soit u un vecteur orthogonal à tous les vecteurs de F ; il est orthogonal à tous les vecteurs de G , donc $u \in G^\perp$.
- 4) On a $\dim(F) + \dim(F^\perp) = \dim(E) = \dim(F^\perp) + \dim(F^\perp)^\perp$, donc $\dim(F^\perp)^\perp = \dim(F)$. D'autre part, tout vecteur de F est orthogonal à $F^\perp$, donc $F \subset (F^\perp)^\perp$, ce qui permet de conclure. $\square$

Nous allons maintenant montrer des liens entre la relation d'orthogonalité de ce chapitre et celle étudiée au chapitre 10.

Formes linéaires définies par un produit scalaire. Soit E un espace euclidien et soit u un vecteur non nul de E . Notons $\varphi : E \times E \rightarrow \mathbb{R}$ le produit scalaire. Pour tout u de E , l'application $\varphi_u : v \mapsto \varphi(u, v)$ est une application linéaire de E dans $\mathbb{R}$. C'est donc une forme linéaire sur E , autrement dit un élément de ce que nous avons appelé au chapitre 10 l'espace dual de E et que nous avons noté E^* .

Hyperplan orthogonal. Dans ces conditions, l'orthogonal d'un vecteur u non nul de E est l'hyperplan $\ker(\varphi_u)$. On voit alors que l'orthogonal d'un sous-espace est l'intersection des hyperplans orthogonaux aux vecteurs du sous-espace.

Isomorphisme de E et E^* . Soit E un espace euclidien. Notons $\varphi_g : E \rightarrow E^*$ l'application $u \mapsto \varphi_u$ (le g indique que u est à gauche).

L'application φ_g est un isomorphisme d'espaces vectoriels.

Démonstration. On vérifie d'abord que φ_g est une application linéaire. Son noyau est formé des vecteurs u de E tels que φ_u soit nulle. Comme $\varphi_u(u) = \varphi(u, u)$, cela

implique $u = 0$ (on peut aussi remarquer que $\ker \varphi_g = E^\perp = \{0\}$). Donc φ_g est injective ; comme E et E^* ont même dimension, φ_g est un isomorphisme. $\square$

On avait noté en 10.3 que la définition d'un isomorphisme entre un espace E de dimension finie et son dual E^* dépendait du choix d'une base de E . Quand E est euclidien, l'isomorphisme φ_g ne dépend pas d'un tel choix : il est donné par le produit scalaire.

Matrice de φ_g . Soient E un espace euclidien, $B = (e_1, \dots, e_n)$ une base de E , $B' = (e_1^*, \dots, e_n^*)$ la base duale de B et Φ la matrice du produit scalaire dans la base B . On a $\Phi = M(\varphi_g, B, B')$.

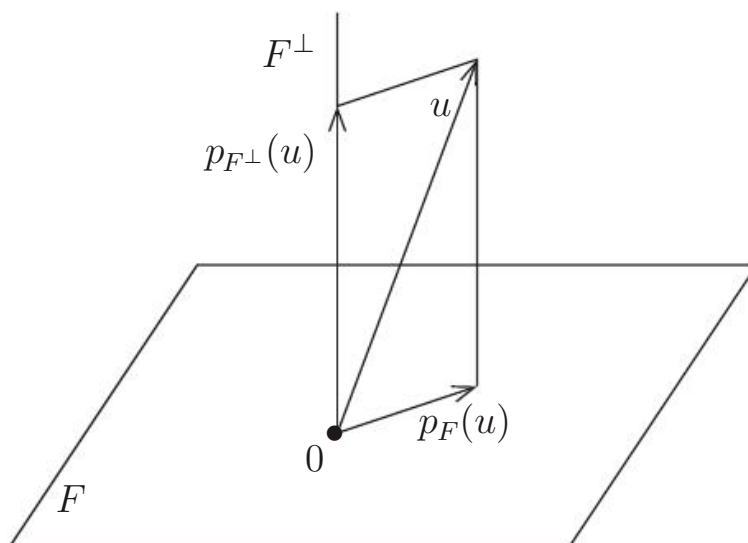
En effet : $\varphi_g(e_i) = \sum_{1 \leq j \leq n} \langle e_i, e_j \rangle e_j^*$.

16.8 PROJECTION ORTHOGONALE

On a défini les projections en 9.5 et on a vu que la donnée d'une projection était équivalente à la donnée d'une décomposition en somme directe.

Définition 1 : projection orthogonale. Soit E un espace euclidien et F un sous espace vectoriel de E . On appelle projection orthogonale (ou projecteur orthogonal) sur F , la projection $p_F : E \rightarrow E$ sur F parallèlement à $F^\perp$. Pour tout vecteur u de E , $p_F(u)$ est appelé projeté orthogonal de u sur F .

Pour tout u de E , $p_F(u)$ est donc un élément de F tel que $u - p_F(u) \in F^\perp$, ce qu'illustre le dessin suivant.



On a vu au chapitre 9 qu'une projection n'allait jamais seule : si $p : E \rightarrow E$ est une projection d'image F , $\text{id}_E - p$ est la projection sur $\ker(p)$ parallèlement à F . Ici, on peut vérifier que $\text{id}_E - p_F$ est la projection orthogonale sur $F^\perp$ et que $u = p_F(u) + (u - p_F(u)) = p_F(u) + p_{F^\perp}(u)$.

Inégalité. Le théorème de Pythagore appliqué aux vecteurs orthogonaux $p_F(u)$ et $u - p_F(u)$ donne : $\|p_F(u)\|^2 + \|u - p_F(u)\|^2 = \|u\|^2$. On en déduit $\|p_F(u)\| \leq \|u\|$, l'égalité ayant lieu si et seulement si $u - p_F(u) = 0$, c'est-à-dire si $u \in F$.

Proposition. Pour tous u et v de E , on a $\langle p_F(u), v \rangle = \langle u, p_F(v) \rangle$.

Démonstration. Décomposons u et v dans la somme directe $E = F \oplus F^\perp$: $u = u_1 + u_2$, $v = v_1 + v_2$ avec $u_1, v_1 \in F$ et $u_2, v_2 \in F^\perp$. On a $p_F(u) = u_1$, $p_F(v) = v_1$ et $\langle p_F(u), v \rangle = \langle u_1, v \rangle = \langle u_1, v_1 \rangle = \langle u, p_F(v) \rangle$. $\square$

Projection sur une droite ou un sous-espace. La formule vue en 16.2 est valable dans un espace euclidien, la projection orthogonale $p_F(v)$ de v sur la droite F engendrée par un vecteur u non nul est : $p_F(v) = \frac{\langle v, u \rangle}{\langle u, u \rangle} u$. Si F est un sous-espace

muni d'une base orthonormée $(e_1, \dots, e_k)$, la projection $p_F(v) = \sum_{1 \leq i \leq k} x_i e_i$ est définie par $x_i = \langle v, e_i \rangle$.

La projection de v sur l'hyperplan $F^\perp$ orthogonal à u est :

$$v - p_F(v) = v - \frac{\langle v, u \rangle}{\langle u, u \rangle} u.$$

Matrice de la projection sur un sous-espace. Soient E un espace euclidien de dimension finie n et F un sous-espace de dimension k de E . Choisissons une base orthogonale $(e_1, \dots, e_k)$ de F et complétons-la en une base orthogonale $B = (e_1, \dots, e_n)$ de E . La matrice $M(p_F, B, B)$ est de la forme :

$$\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

Distance à un sous-espace. Soient E un espace euclidien de dimension finie n et F un sous-espace de dimension k de E . On définit la distance d'un vecteur u de E à F par $d(u, F) = \|u - p_F(u)\|$.

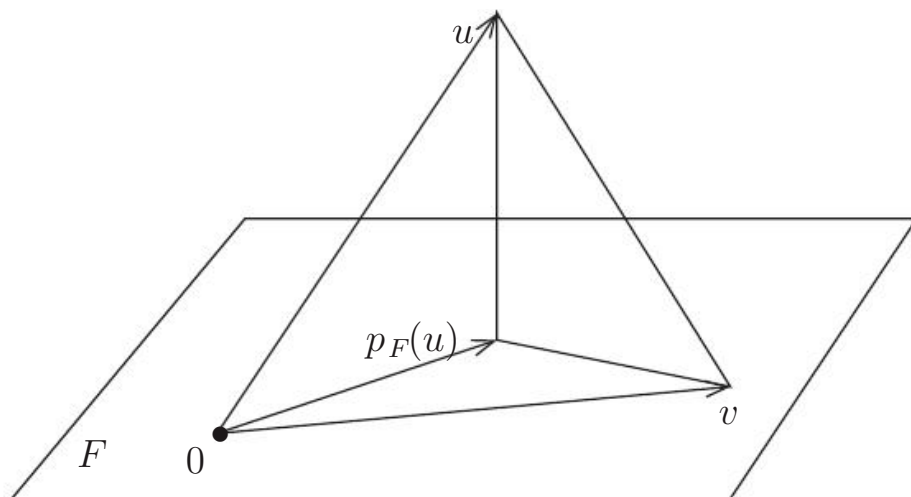
Proposition. Pour tout v de F , on a $\|u - v\| \geq d(u, F)$, l'égalité n'ayant lieu que si $v = p_F(u)$.

Démonstration. On a $u - v = u - p_F(u) + p_F(u) - v$. Comme $u - p_F(u)$ est orthogonal à F , il est orthogonal à $p_F(u) - v$; le théorème de Pythagore donne

Chapitre 16 • Orthogonalité

$\|u - p_F(u)\|^2 + \|p_F(u) - v\|^2 = \|u - v\|^2$ d'où $d(u, F) \leq \|u - v\|$, l'égalité n'ayant lieu que si $\|p_F(u) - v\| = 0$, soit $v = p_F(u)$. $\square$

Géométriquement, cette proposition signifie que quand on s'éloigne dans F du pied de la perpendiculaire abaissée de u sur F , on augmente la distance. On peut considérer cette proposition comme un résultat d'approximation : l'élément de F le plus proche de u est le projeté orthogonal de u sur F .



Exemple. Soit $\text{Cont}([0, 1], \mathbb{R})$ l'espace des fonctions continues sur $[0, 1]$ muni du produit scalaire défini dans l'exemple 3 de 16.3. Cet espace est de dimension infinie, mais on va se limiter au sous-espace E des fonctions polynômes de degré ≤ 2 . Posons $e_k : t \mapsto t^k$ pour $k = 0, 1, 2$ et notons F le sous-espace $\text{Vect}(e_0, e_1)$ de E . Chercher le projeté orthogonal de e_2 sur F , c'est chercher la fonction $ae_1 + be_0$ la plus proche de e_2 dans F , autrement dit, c'est chercher a et b tels que $\|e_2 - (ae_1 + be_0)\|^2 = \int_0^1 (t^2 - at - b)^2 dt$ soit minimum. Pour trouver a et b , on écrit les conditions d'orthogonalité de $e_2 - (ae_1 + be_0)$ à F :

$$\begin{aligned} \int_0^1 (t^2 - at - b) dt &= \langle e_2 - (ae_1 + be_0), e_0 \rangle = 0 \\ \int_0^1 (t^3 - at^2 - bt) dt &= \langle e_2 - (ae_1 + be_0), e_1 \rangle = 0 \end{aligned}$$

Les deux calculs d'intégrales donnent les conditions :

$$\begin{aligned} \frac{1}{3} - \frac{a}{2} - b &= 0 \\ \frac{1}{4} - \frac{a}{3} - \frac{b}{2} &= 0 \end{aligned}$$

d'où $a = 1$ et $b = -\frac{1}{6}$. L'élément de F le plus proche de e_2 est donc $t \mapsto t - \frac{1}{6}$.

On peut alors calculer que $d(e_2, F) = \frac{1}{180}$.

L'analyse des données. L'exemple qui précède illustre l'idée d'approximation par une projection orthogonale dans un espace de fonctions.

On retrouve aussi l'idée d'approximation en analyse des données. Les données recueillies pour une étude peuvent souvent être représentées dans un espace euclidien E de grande dimension, 1 000 par exemple. Dans un tel espace, il est impossible de voir comment les données s'organisent. Des algorithmes ont été mis au point (les premiers à Rennes sous la direction de Jean-Paul Benzécri (1932-2019), en 1963-1964) ; depuis, cette discipline s'est considérablement développée et est utilisée dans de nombreuses enquêtes et recherches de toutes disciplines) pour trouver un sous-espace de petite dimension : 1, 2, 3 le plus souvent, tels que la projection orthogonale de E sur ce sous-espace déforme le moins possible les données. Sur ce sous-espace, on peut *mieux voir* les données : celles qui se regroupent, celles qui s'opposent, etc.

Méthode des moindres carrés. Bien des mesures de phénomènes naturels conduisent à rechercher une relation de la forme $y = ax + b$ entre deux variables. On dispose de couples de valeurs $(x_1, y_1), \dots, (x_n, y_n)$ qui conduisent en général à un système d'équations $ax_i + b = y_i$, $1 \leq i \leq n$ incompatibles. On recherche donc a et b qui satisfassent *au mieux* ces équations. On peut chercher à minimiser la somme des écarts $\sum_{1 \leq i \leq n} |y_i - (ax_i + b)|$, ce qui ne conduit pas à des calculs faciles ; on préfère souvent minimiser la somme des carrés des écarts : $\sum_{1 \leq i \leq n} (y_i - (ax_i + b))^2$, ce qui revient, en posant $u = (x_1, \dots, x_n)$, $v = (y_1, \dots, y_n)$, $\omega = (1, \dots, 1)$, à chercher le vecteur de $\mathbb{R}^n$ de la forme $au + b\omega$ le plus proche de v , donc à déterminer le projeté orthogonal de u sur le plan Vect (u, ω) .

On prendra garde à la signification de ce qu'on fait : la projection de u sur le plan Vect (u, ω) ne donne pas la même loi.

C'est le premier janvier 1801 à Palerme que l'astronome Giuseppe Piazzi (1746-1826) découvre un nouvel objet gravitant autour du soleil. Il l'identifie rapidement comme étant une nouvelle planète d'orbite intermédiaire entre celle de Mars et celle de Jupiter. Ses observations continuent jusqu'au 11 février 1801 où la planète est perdue.

À partir des observations de Piazzi, Gauss détermine les paramètres de l'orbite de la planète. Le 31 décembre 1801, la planète, baptisée Cérès, fut retrouvée dans le ciel à la position indiquée par ses calculs. Le succès de Gauss tient d'une part à son approche théorique : sa méthode de calcul d'orbite peut s'appliquer à tout objet du système solaire et ne présuppose pas une orbite circulaire, et, d'autre part, à une

méthode nouvelle de calcul numérique : la méthode des moindres carrés, qu'il utilise depuis plusieurs années et qu'il publie en 1809.

16.9 TRANSFORMATIONS ORTHOGONALES

Intéressons-nous maintenant aux applications linéaires qui respectent le produit scalaire.

Proposition 1. *Soient E et E' deux espaces euclidiens (on ne distinguera dans les notations ni les produits scalaires ni les normes de ces deux espaces). Soit $f: E \rightarrow E'$ une application linéaire. Les conditions suivantes sont équivalentes.*

- 1) *f conserve les produits scalaires, c'est-à-dire que pour tous u, v de E on a $\langle f(u), f(v) \rangle = \langle u, v \rangle$;*
- 2) *f conserve les normes, c'est-à-dire que pour tout u de E on a $\|f(u)\| = \|u\|$.*

Démonstration. Si la première condition est vérifiée, on a :

$$\|f(u)\|^2 = \langle f(u), f(u) \rangle = \langle u, u \rangle = \|u\|^2.$$

Si la seconde condition est vérifiée, la proposition 2 de 16.5 donne :

$$2 \langle f(u), f(v) \rangle = \|f(u) + f(v)\|^2 - \|f(u)\|^2 - \|f(v)\|^2,$$

ce qui permet de montrer que f respecte le produit scalaire. □

Définition 1 : transformations orthogonales. Soient E et E' deux espaces euclidiens de même dimension et $f: E \rightarrow E'$ une application linéaire. On dit que f est une application (ou une transformation) orthogonale si f vérifie l'une des conditions équivalentes précédentes.

Nous dirons transformation orthogonale plutôt qu'application orthogonale pour suivre les usages. Cette appellation se réfère à la conservation de l'orthogonalité. Mais on parle aussi pour la même notion d'isométrie, ce qui fait référence à la conservation des normes.

Proposition 2 : groupe orthogonal. *Les applications orthogonales d'un espace euclidien E dans lui-même et la composition des applications forment un groupe appelé groupe orthogonal de E et noté $O(E)$.*

Démonstration. La seconde condition de la proposition précédente montre que si f est une transformation orthogonale, $\ker(f) = \{0\}$, donc f est bijective, donc son inverse f^{-1} est linéaire et conserve les normes ; par conséquent, f^{-1} est une transformation orthogonale. De plus, le composé de deux transformations orthogonales et l'identité sont des transformations orthogonales. □

Le groupe orthogonal est un sous-groupe du groupe linéaire. Son introduction est due à Jordan (1870). Elle précède l'article de 1873 sur la géométrie à n dimensions dont nous avons parlé en 3.1.

Proposition 3 : matrices des transformations orthogonales. Soient E un espace euclidien, $B = (e_1, \dots, e_n)$ une base orthonormée de E et $f : E \rightarrow E$ une application linéaire. On pose $A = M(f, B, B)$.

1) Si f est une transformation orthogonale, la matrice A :

- est inversible ;
- vérifie ${}^tAA = A {}^tA = I$, autrement dit : ${}^tA = A^{-1}$;
- est telle que $\det(A) = \pm 1$.

2) Si A vérifie ${}^tAA = A {}^tA = I$, autrement dit : ${}^tA = A^{-1}$, alors f est une transformation orthogonale.

Démonstration.

- 1) A est inversible parce que f est un isomorphisme. Pour tous u, v de E , de matrices U et V dans la base B , on a $\langle f(u), f(v) \rangle = \langle u, v \rangle$ donc ${}^t(AU)AV = {}^tU {}^tAAV = {}^tUV$, d'où ${}^tAA = I$. Enfin, en prenant les déterminants des deux membres, on obtient $(\det(A))^2 = 1$.
- 2) Pour tous u, v de E , avec les mêmes notations qu'au 1), on a $\langle f(u), f(v) \rangle = {}^t(AU)AV = {}^tU {}^tAAV = {}^tUV = \langle u, v \rangle$, donc f est une transformation orthogonale. $\square$

Une matrice carrée A vérifiant l'une des conditions équivalentes du 2) :

${}^tAA = I, A {}^tA = I, {}^tA = A^{-1}$, est appelée matrice orthogonale.

Définition 2 : groupe spécial orthogonal. Soit E un espace euclidien. Les transformations orthogonales de déterminant 1 forment un sous-groupe de $O(E)$ appelé groupe spécial orthogonal et noté $SO(E)$ (ou $O^+(E)$).

Les éléments de ce groupe conservent l'orientation de E ; ce sont des transformations orthogonales directes appelées parfois rotations.

Les transformations orthogonales de déterminant -1 ne conservent pas l'orientation et sont appelées indirectes.

Définition 3 : symétrie orthogonale. Soient E un espace euclidien, F un sous-espace de E . On définit la symétrie orthogonale s par rapport à F comme la symétrie par rapport à F parallèlement à $F^\perp$.

Si p note la projection orthogonale sur F , pour tout u de E , $s(u)$ est une somme de deux vecteurs orthogonaux : $s(u) = p(u) + (p(u) - u)$, l'un dans F , l'autre dans $F^\perp$. Pour tout u de E , le théorème de Pythagore donne :

$$\|s(u)\|^2 = \|p(u)\|^2 + \|p(u) - u\|^2 = \|u\|^2.$$

Donc s est une transformation orthogonale.

En choisissant une base orthonormée de F et une base orthonormée de $F^\perp$, on obtient une base orthonormée de E , par rapport à laquelle la matrice de s est diagonale ; les éléments diagonaux sont 1 ou -1 . Le nombre de -1 est égal à la dimension de $F^\perp$. Par conséquent s est une transformation directe si $\dim(F^\perp)$ est paire et indirecte si $\dim(F^\perp)$ est impaire.

Proposition 4 : transformations orthogonales et bases orthonormées. Soient E un espace euclidien, $B = (e_1, \dots, e_n)$ une base orthonormée de E et $f : E \rightarrow E$ une application linéaire.

- 1) Si f est une transformation orthogonale, $(f(e_1), \dots, f(e_n))$ est une base orthonormée de E .
- 2) Si $(f(e_1), \dots, f(e_n))$ est une base orthonormée de E , f est une transformation orthogonale.

Démonstration. 1) On a $\|f(e_i)\| = \|e_i\| = 1$ pour $1 \leq i \leq n$ et $\langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = 0$ pour $1 \leq i, j \leq n$ et $i \neq j$.

2) Soient $u = \sum_{1 \leq i \leq n} x_i e_i$ et $v = \sum_{1 \leq i \leq n} y_i e_i$ deux vecteurs de E . On a :

$$\begin{aligned} \langle f(u), f(v) \rangle &= \langle f\left(\sum_{1 \leq i \leq n} x_i e_i\right), f\left(\sum_{1 \leq i \leq n} y_i e_i\right) \rangle \\ &= \sum_{1 \leq i, j \leq n} x_i y_j \langle f(e_i), f(e_j) \rangle = \sum_{1 \leq i} x_i y_i = \langle u, v \rangle. \end{aligned} \quad \square$$

Corollaire. Soient E un espace euclidien, $B = (e_1, \dots, e_n)$ et $B' = (e'_1, \dots, e'_n)$ deux bases orthonormées de E . La matrice de passage $P = M(\text{id}, B', B)$ est une matrice orthogonale.

Démonstration. Notons f la transformation définie par $f(e_i) = e'_i$ pour $1 \leq i \leq n$. Comme f est orthogonale d'après la proposition précédente et que $P = M(f, B, B)$, on voit que P est une matrice orthogonale. $\square$

Proposition 5 : modèle d'espace euclidien de dimension n . Un espace euclidien E de dimension n est isométrique à l'espace $\mathbb{R}^n$ muni du produit scalaire usuel (voir 16.3, exemple 1).

Démonstration. Notons $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$ et choisissons dans E une base $(\varepsilon_1, \dots, \varepsilon_n)$ orthonormée. L'isomorphisme d'espaces vectoriels $f : \mathbb{R}^n \rightarrow E$ défini par $f(e_i) = \varepsilon_i$ pour $1 \leq i \leq n$ est une isométrie. $\square$

Valeurs propres d'une transformation orthogonale. Soient E un espace euclidien et $f : E \rightarrow E$ une transformation orthogonale. Les valeurs propres réelles de f sont ± 1 . En effet, si λ est une valeur propre réelle de f et v un vecteur propre associé, on a $f(v) = \lambda v$ et $\|f(v)\| = \|v\|$ d'où $|\lambda| = 1$ et $\lambda = \pm 1$.

Notons A la matrice de f ; on a ${}^tAA = I$. Si λ est une valeur propre non réelle de A , notons v un vecteur propre associé dans $\mathbb{C}^n$ et V la matrice de v dans la base canonique de $\mathbb{C}^n$. On a $AV = \lambda V$, donc $A\bar{V} = \bar{\lambda}\bar{V}$ puisque les coefficients de A sont réels. Par conséquent, ${}^t\bar{V}V = {}^t\bar{V}{}^tAAV = {}^t\bar{A}VAV = |\lambda|^2 {}^t\bar{V}V$ et on voit que $|\lambda| = 1$.

Les valeurs propres, réelles ou non, d'une transformation orthogonale sont donc de module 1. Nous reviendrons sur ce passage à un espace vectoriel sur $\mathbb{C}$ en 22.1.

16.10 GROUPE ORTHOGONAL DE $\mathbb{R}^2$

Description d'un élément de $O(\mathbb{R}^2)$. Soit $B = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$. Elle définit un produit scalaire et une orientation de $\mathbb{R}^2$. Soit f une transformation orthogonale de $\mathbb{R}^2$. La matrice $A = M(f, B, B) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est orthogonale et vérifie ${}^tAA = I$, ce qui donne les relations ;

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

Les deux premières relations montrent qu'il existe des réels θ et θ' tels que :

$$\begin{cases} a = \cos \theta & c = \sin \theta \\ b = \cos \theta' & d = \sin \theta' \end{cases}$$

La troisième relation impose $\cos(\theta - \theta') = 0$, soit $\theta' = \theta + \frac{\pi}{2} \bmod 2\pi$ ou $\theta' = \theta - \frac{\pi}{2} \bmod 2\pi$. On trouve donc deux types de matrices orthogonales :

$$\text{Type 1: } \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \text{Type 2: } \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$$

Dans le premier cas, f est la rotation $\rho(\theta)$ d'angle θ . Son déterminant est 1 ; c'est une transformation orthogonale directe. Les valeurs propres de f sont $e^{\pm i\theta}$; elles sont de module 1. Comme cas particuliers, on trouve les matrices $\pm I$.

Dans le second cas, le déterminant de f est -1 et le polynôme caractéristique est $\lambda^2 - 1$ donc f est diagonalisable, de valeurs propres 1 et -1 . Si ε_1 et ε_2 notent des vecteurs propres associés aux valeurs propres 1 et -1 , on a $f(\varepsilon_1) = \varepsilon_1$ et $f(\varepsilon_2) = -\varepsilon_2$.

Donc f est une symétrie orthogonale par rapport à Vect (ε_1) et sa matrice dans la base $(\varepsilon_1, \varepsilon_2)$ s'écrit :

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

Pour déterminer ε_1 , on peut chercher la symétrie orthogonale qui transforme e_1 en $f(e_1) = (\cos \theta, \sin \theta)$, ce qui donne $\varepsilon_1 = \left(\cos \frac{\theta}{2}, \sin \frac{\theta}{2} \right)$ où $\frac{\theta}{2}$ est défini modulo π et ε_1 au signe près.

Proposition 1 : conjugué d'une rotation par une symétrie orthogonale. Soient s une symétrie orthogonale par rapport à un axe défini par e_1 et $\rho(\theta)$ une rotation de $O(\mathbb{R}^2)$. On a :

$$s \circ \rho(\theta) \circ s = \rho(-\theta)$$

Démonstration. Choisissons e_2 tel que (e_1, e_2) soit une base orthonormée de $\mathbb{R}^2$. Le composé du membre de gauche a pour déterminant 1 ; c'est donc une rotation d'après l'étude précédente. Pour déterminer son angle, on calcule $(s \circ \rho(\theta) \circ s)(e_1) = (\cos \theta, -\sin \theta) = \rho(-\theta)(e_1)$, ce qui donne le résultat. $\square$

Proposition 2 : rotations et symétries orthogonales.

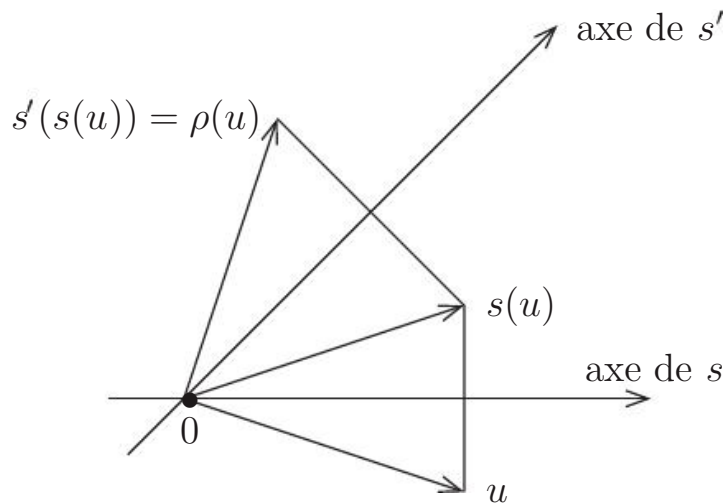
- 1) Soient (e_1, e_2) une base orthonormée de $\mathbb{R}^2$, $\rho(\theta)$ une rotation et s la symétrie orthogonale par rapport à e_1 . Alors $\rho(\theta) \circ s$ est la symétrie orthogonale par rapport à l'axe engendré par $\left(\cos \frac{\theta}{2}, \sin \frac{\theta}{2} \right)$ et $s \circ \rho(\theta)$ est la symétrie orthogonale par rapport à l'axe engendré par $\left(\cos \frac{\theta}{2}, -\sin \frac{\theta}{2} \right)$.
- 2) Toute rotation s'écrit comme un produit de deux symétries orthogonales par rapport à des axes, l'une d'entre elles pouvant être choisie arbitrairement.
- 3) Un produit de deux symétries orthogonales est une rotation.

Démonstration. 1) Les composés $\rho(\theta) \circ s$ et $s \circ \rho(\theta)$ sont des transformations orthogonales de déterminant -1 de $O(\mathbb{R}^2)$. Ce sont donc des symétries. Pour connaître leurs axes, il suffit de déterminer l'image de e_1 .

Comme $\rho(\theta)(s(e_1)) = (\cos \theta, \sin \theta)$, $\rho(\theta) \circ s$ est la symétrie par rapport à l'axe indiqué. Comme $s(\rho(\theta)(e_1)) = (\cos \theta, -\sin \theta)$, $s \circ \rho(\theta)$ est la symétrie par rapport à l'axe indiqué.

- 2) Soient $\rho(\theta)$ une rotation et s une symétrie orthogonale par rapport à un axe quelconque ; on a montré au 1) que $\rho(\theta) \circ s = s'$ et que $s \circ \rho(\theta) = s''$ sont des symétries orthogonales, d'où le résultat.
- 3) Le composé de deux symétries orthogonales s et s' a pour déterminant 1 ; c'est donc une rotation.

On peut préciser l'angle de cette rotation. Soient e_1, e'_1 des vecteurs orthonormés tels que $s(e_1) = e_1$, $s'(e'_1) = e'_1$ et soit θ l'angle tel que $\rho(\theta)(e_1) = e'_1$. On a $s'(s(e_1)) = \rho(2\theta)(e_1)$ donc $s' \circ s = \rho(2\theta)$. $\square$



Propriétés des groupes $O(\mathbb{R}^2)$ et $SO(\mathbb{R}^2)$. Le groupe des rotations $SO(\mathbb{R}^2)$ est un groupe commutatif. La formule $s \circ \rho(\theta) \circ s = \rho(-\theta)$ montre que le groupe $O(\mathbb{R}^2)$ n'est pas commutatif. Le centre de $O(\mathbb{R}^2)$ contient par définition les transformations qui commutent avec tous les éléments de $SO(\mathbb{R}^2)$. Aucune symétrie ne peut donc y appartenir. Le centre ne contient que les rotations telles que $\rho(\theta) = \rho(-\theta)$, c'est-à-dire $\rho(0)$ et $\rho(\pi)$; il est réduit à $\{ \text{id}, -\text{id} \}$.

Le 2) de la proposition 2 montre aussi que l'ensemble des symétries orthogonales engendre $O(\mathbb{R}^2)$.

16.11 GROUPE ORTHOGONAL DE $\mathbb{R}^3$

Description d'un élément de $SO(\mathbb{R}^3)$. Soit $B = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$. Cette base définit un produit scalaire et une orientation de $\mathbb{R}^3$.

Soit f une transformation orthogonale directe de $\mathbb{R}^3$. On a $\det(f) = 1$. Montrons que 1 est valeur propre de f .

Comme le polynôme caractéristique de f est de degré impair, il admet au moins une racine réelle λ . On a vu que $\lambda = \pm 1$. Si f n'admet pas la valeur propre 1, $\lambda = -1$ est valeur propre de f . Les autres valeurs propres μ et μ' de f ne peuvent

être réelles car elles seraient égales à -1 et on aurait $\det(f) = \lambda\mu\mu' = -1$. Si elles n'étaient pas réelles, alors on aurait $\mu' = \bar{\mu}$ et $\det(f) = \lambda\mu\mu' = -1$. Donc $\lambda = 1$ est valeur propre de f .

Soit ε_1 un vecteur propre associé à la valeur propre 1 et posons $F = \text{Vect}(\varepsilon_1)$. L'espace $F^\perp$ est stable par f car si $u \in F^\perp$, on a :

$$\langle f(u), \varepsilon_1 \rangle = \langle f(u), f(\varepsilon_1) \rangle = \langle u, \varepsilon_1 \rangle = 0.$$

La restriction f' de f à $F^\perp$ est donc une transformation orthogonale de déterminant 1 . Choisissons une base orthonormée $B' = (\varepsilon_2, \varepsilon_3)$ de $F^\perp$ de façon que $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ soit une base directe de $\mathbb{R}^3$.

Comme $\det(f') = 1$, f' est une rotation et la matrice de f par rapport à la base $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est de la forme :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

Détermination de l'axe et de l'angle de rotation. Si $f \neq \text{id}$, 1 est valeur propre simple de f et l'axe de la rotation est déterminé par un vecteur propre ε_1 associé à la valeur propre 1 .

Pour déterminer l'angle θ , on remarque que la trace de f , qui ne dépend pas de la base dans laquelle elle est calculée, donne $\text{Tr}(f) = 1 + 2 \cos \theta$, ce qui donne $\cos \theta$.

Pour déterminer complètement θ , il reste à déterminer le signe de $\sin \theta$. Choisissons un vecteur u quelconque non colinéaire à ε_1 (l'un des trois vecteurs de la base canonique convient sûrement). Le procédé d'orthogonalisation de Gram-Schmidt permet de trouver, à partir de la base formée par ε_1, u et un troisième vecteur, une base orthonormée directe $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ de $\mathbb{R}^3$ telle que $u = \alpha \varepsilon_1 + \beta \varepsilon_2$. On a $f(u) = \alpha f(\varepsilon_1) + \beta f(\varepsilon_2) = \alpha \varepsilon_1 + \beta \cos \theta \varepsilon_2 + \beta \sin \theta \varepsilon_3$, d'où :

$$\det(\varepsilon_1, u, f(u)) = \begin{vmatrix} 1 & \alpha & \alpha \\ 0 & \beta & \beta \cos \theta \\ 0 & 0 & \beta \sin \theta \end{vmatrix} = \beta^2 \sin \theta.$$

Par conséquent, le signe de ce déterminant, calculé dans la base initiale, donne le signe cherché.

Exemple. Considérons l'application linéaire f définie par la matrice :

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

dans la base canonique. On vérifie que f est une transformation orthogonale en vérifiant que ${}^tAA = I$. Comme $\det(A) = 1$, f est une rotation.

L'axe de rotation est défini par un vecteur propre associé à la valeur propre 1. On vérifiera qu'on peut prendre $\varepsilon_1 = (1, 1, 1)$.

Comme $\text{Tr}(f) = 0$, on trouve $\cos\theta = -\frac{1}{2}$.

On prend alors $u = e_1$; on a $f(u) = e_2$ et $\det(\varepsilon_1, u, f(u)) = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1$. On

a donc $\sin\theta > 0$, donc $\theta = \frac{2\pi}{3} \bmod 2\pi$.

Le choix de $\varepsilon_1 = (-1, -1, -1)$ conduit à $\sin\theta < 0$, donc $\theta = -\frac{2\pi}{3} \bmod 2\pi$.

Soulignons que l'angle d'une rotation dans $\mathbb{R}^3$ dépend de l'orientation de son axe. Pour s'en convaincre, on peut s'imaginer avec un tournevis à la main ou dans la peau du bonhomme d'Ampère.

Demi-tour. On appelle demi-tour une rotation d'angle $\theta = \pi$. Il existe donc une base orthonormée dans laquelle la matrice d'un demi-tour est :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Description d'un élément de déterminant -1 . Soit $f \in O(\mathbb{R}^3)$ de déterminant -1 . Supposons $f \neq \text{id}$ et posons $g = -f$. Puisque $\det(g) = 1$, on peut appliquer ce qui précède à g . Comme g a un vecteur propre ε_1 associé à la valeur propre 1, f a -1 pour valeur propre et ε_1 vérifie $f(\varepsilon_1) = -\varepsilon_1$. Notons s la symétrie orthogonale par rapport au plan orthogonal à ε_1 . On a $\det(f \circ s) = 1$, donc $f \circ s$ est une rotation ρ ; son axe est défini par ε_1 et son angle comme ci-dessus. On a $f = \rho \circ s = s \circ \rho$.

On peut dire que toute isométrie indirecte est produit d'une rotation par une symétrie par rapport au plan orthogonal à son axe. La matrice de f par rapport à une base orthonormée de premier vecteur ε_1 est de la forme :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Proposition. Toute rotation de $\mathbb{R}^3$ est produit de deux symétries orthogonales par rapport à des plans vectoriels.

Démonstration. Soit f une rotation de $\mathbb{R}^3$ d'axe ε_1 et d'angle θ . Notons $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ une base orthonormée directe de $\mathbb{R}^3$ et s la symétrie orthogonale par rapport au plan Vect $(\varepsilon_1, \varepsilon_2)$. On vérifie que $f \circ s$ est une symétrie orthogonale par rapport au plan défini par ε_1 et $\cos \frac{\theta}{2} \varepsilon_2 + \sin \frac{\theta}{2} \varepsilon_3$. $\square$

Autres propriétés des groupes $O(\mathbb{R}^3)$ et $SO(\mathbb{R}^3)$. Le groupe des rotations $SO(\mathbb{R}^3)$ n'est pas un groupe commutatif : par exemple, le produit des rotations d'angle $\frac{\pi}{2}$ autour des axes définis par e_1 et e_2 ne commute pas.

Le centre de $O(\mathbb{R}^3)$ est égal à $\{\text{id}, -\text{id}\}$, celui de $SO(\mathbb{R}^3)$ est égal à $\{\text{id}\}$.

L'ensemble des symétries orthogonales par rapport à des plans engendre $O(\mathbb{R}^3)$.

L'étude des sous-groupes finis de $SO(\mathbb{R}^3)$ est liée à l'étude des polyèdres réguliers de l'espace. Nous ne pouvons l'entreprendre ici.

Prolongements. L'étude du groupe orthogonal peut être continuée et les problèmes précédents étudiés dans le cadre des espaces euclidiens $\mathbb{R}^n$ (voir 23.1).

16.12 ENDOMORPHISME ADJOINT ET AUTOADJOINT

16.12.1 Endomorphisme adjoint

Proposition et définition. Soient E un espace euclidien et f un endomorphisme de E .

1) Il existe un unique endomorphisme noté f^* , appelé adjoint de f , tel que, pour tous u, v de E , on ait :

$$\langle f(u), v \rangle = \langle u, f^*(v) \rangle$$

2) Si B est une base orthonormée de E et si A est la matrice de f dans la base B , alors la matrice de f^* dans la base B est ${}^tA$.

Démonstration. Soit $B = (e_1, \dots, e_n)$ une base orthonormée de E . En notant A, A^* respectivement les matrices de f, f^* dans la base B , on doit avoir, pour tous u, v de E de matrices U, V par rapport à la base B :

$$\langle f(u), v \rangle = {}^t(AU)V = {}^tU {}^tAV$$

et

$$\langle u, f^*(v) \rangle = {}^tU A^* V.$$

On en déduit que $A^* = {}^tA$ est la seule possibilité, puis le 2). $\square$

Propriétés de l'adjonction. Soient E un espace euclidien et f, g deux endomorphismes de E .

- 1) $(f^*)^* = f$;
- 2) $(f + g)^* = f^* + g^*$;
- 3) $(\lambda f)^* = \lambda f^*$;
- 4) $(g \circ f)^* = f^* \circ g^*$.

Adjoint dans une base quelconque. Soient E un espace euclidien et f un endomorphisme de E . Si B est une base quelconque de E , si Φ est la matrice du produit scalaire dans la base B et si A est la matrice de f dans la base B , alors la matrice A^* de f^* dans la base B est $\Phi^{-1t} A \Phi$.

En effet, pour tous u, v de E , on a $\langle f(u), v \rangle = {}^tU {}^tA \Phi V$ et $\langle u, f^*(v) \rangle = {}^tU \Phi A^* V$, donc ${}^tA \Phi = \Phi A^*$ et le résultat.

16.12.2 Endomorphisme autoadjoint

Définition. Soit E un espace euclidien. Un endomorphisme de E est dit autoadjoint ou symétrique si, pour tous u, v de E , on a :

$$\langle f(u), v \rangle = \langle u, f(v) \rangle,$$

autrement dit si $f^* = f$.

Matrice d'un endomorphisme symétrique. Soient E un espace euclidien, f un endomorphisme de E , $B = (e_1, \dots, e_n)$ une base orthonormée de E et $A = M(f, B, B)$ la matrice de f dans la base B .

Comme $A^* = {}^tA$, l'endomorphisme f est symétrique si et seulement si $A = {}^tA$, autrement dit si et seulement si la matrice A est symétrique.

Exemples. Nous avons vu qu'une projection orthogonale, une symétrie orthogonale, une homothétie sont des endomorphismes symétriques. Par contre, une rotation n'est pas un endomorphisme symétrique.

Le résultat important est celui-ci.

Théorème : diagonalisation des endomorphismes symétriques.

- 1) Soit f un endomorphisme symétrique de l'espace euclidien $\mathbb{R}^n$ avec son produit scalaire usuel.
 - a) Les valeurs propres de f sont réelles.
 - b) Les sous-espaces propres de f sont deux à deux orthogonaux.
 - c) f est diagonalisable.
- 2) Soit A une matrice symétrique à coefficients réels. A est diagonalisable et on peut trouver une matrice orthogonale P telle que $P^{-1}AP = {}^tPAP$ soit diagonale.

Démonstration. 1) Nous allons utiliser les nombres complexes pour cette démonstration. Il existe des démonstrations de ce théorème qui évitent cela. Notons $\psi : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ l'application définie par $\psi(u, v) = \sum_{1 \leq i \leq n} x_i \overline{y_i}$ pour $u = (x_1, \dots, x_n)$

et $v = (y_1, \dots, y_n)$. L'application ψ ressemble à une forme bilinéaire sur $\mathbb{C}^n$. Elle possède les propriétés suivantes, pour tous u, u', v, v' de $\mathbb{C}^n$ et tout λ de $\mathbb{C}$:

- $\psi(u + u', v) = \psi(u, v) + \psi(u', v)$;
- $\psi(u, v + v') = \psi(u, v) + \psi(u, v')$;
- $\psi(\lambda u, v) = \lambda \psi(u, v)$;
- $\psi(u, \lambda v) = \bar{\lambda} \psi(u, v)$;
- $\psi(u, u) = \sum_{1 \leq i \leq n} |x_i|^2 \geq 0$ et $\psi(u, u) = 0$ équivaut à $u = 0$.

a) Notons A la matrice de f dans la base canonique de $\mathbb{R}^n$. Les coefficients de A sont réels mais nous allons considérer A comme la matrice à coefficients complexes d'une application linéaire $g : \mathbb{C}^n \rightarrow \mathbb{C}^n$. Le polynôme caractéristique $\chi(A)$ de A est à coefficients réels donc si a non réel est racine de $\chi(A)$, $\bar{a}$ est aussi racine de $\chi(A)$.

Pour tous u, v de $\mathbb{C}^n$, on a $\psi(g(u), v) = \psi(u, g(v))$ car les coefficients de g sont réels.

Soit λ une valeur propre de A et $u = (x_1, \dots, x_n)$ un vecteur propre associé dans $\mathbb{C}^n$. On a :

$\lambda \psi(u, u) = \psi(\lambda u, u) = \psi(g(u), u) = \psi(u, g(u)) = \psi(u, \lambda u) = \bar{\lambda} \psi(u, u)$. Comme $\psi(u, u) > 0$, on a donc $\lambda = \bar{\lambda}$, ce qui prouve que λ est réel.

b) Le rôle de $\mathbb{C}$ est maintenant terminé. Soient λ et μ deux valeurs propres réelles distinctes de A et u et v des vecteurs propres associés dans $\mathbb{R}^n$. En utilisant le produit scalaire usuel de $\mathbb{R}^n$, on a $\lambda \langle u, v \rangle = \mu \langle u, v \rangle$ par un calcul analogue au précédent. On en déduit que u et v sont orthogonaux.

c) Si la somme F des espaces propres associés aux différentes valeurs propres de A n'est pas $\mathbb{R}^n$, $F^\perp$ n'est pas réduit à $\{0\}$. L'espace F est stable par f , puisque tout vecteur de F est une combinaison linéaire de vecteurs propres de f . Soit $u \in F^\perp$; pour tout v de F , on a $\langle f(u), v \rangle = \langle u, f(v) \rangle = 0$ car $f(v) \in F$. Donc $F^\perp$ est stable par f . La restriction f' de f à $F^\perp$ est un endomorphisme symétrique. On vient de voir que les valeurs propres d'un tel endomorphisme sont réelles, donc il existe dans $F^\perp$ un vecteur propre de f' donc de f , ce qui contredit la définition de F .

2) Choisissons une base de vecteurs propres orthonormée. Les matrices de passage P et P^{-1} sont alors orthogonales. □

La matrice P joue donc ici un rôle double : elle sert à exprimer à la fois le changement de base pour le produit scalaire et pour l'application linéaire.

16.13 POLYNÔMES ORTHOGONAUX : EXEMPLE DES POLYNÔMES DE LEGENDRE

Introduction. Dans ce paragraphe, nous présentons un système de polynômes orthogonaux : celui des polynômes de Legendre. C'est en 1783 que Adrien-Marie Legendre (1752-1833) introduit ces polynômes pour calculer l'attraction newtonienne sur un point matériel d'un corps homogène de révolution et symétrique par rapport à l'équateur. Le but des calculs de Legendre est de déterminer la forme des planètes soumises à une attraction. Dans ces calculs apparaissent des inverses de distances de deux points du plan. Si ces points sont $(1, 0)$ et $(r \cos \theta, r \sin \theta)$, l'inverse de leur distance est $\frac{1}{\sqrt{1 - 2r \cos \theta + r^2}}$ et Legendre pense à développer cette

expression sous la forme $\sum_{n \geq 0} r^n L_n(\cos \theta)$ où les L_n sont des polynômes. L'année suivante, il dégage les principales propriétés des polynômes L_n , en particulier les relations d'orthogonalité, la réalité des racines et leur localisation entre -1 et 1 , la majoration $|L_n(x)| \leq 1$ entre -1 et 1 . Laplace complète très vite les résultats de Legendre : la compétition entre les deux savants est alors très vive.

Au cours du XIX^e siècle, bien d'autres systèmes de polynômes orthogonaux ont été inventés pour résoudre des problèmes d'analyse issus de la physique.

Les relations d'orthogonalité découvertes par Legendre vont jouer un rôle essentiel dans ce paragraphe.

16.13.1 Systèmes de polynômes orthogonaux associés au produit scalaire $\langle f, g \rangle = \int_I f(t)g(t) dt$

Notations. On pose $I = [-1, 1]$ et on note E l'espace $\text{Cont}(I, \mathbb{R})$ des fonctions continues sur I . On munit E du produit scalaire défini dans l'exemple 3 de 16.3 :

$$\varphi(f, g) = \langle f, g \rangle = \int_I f(t)g(t) dt$$

Dans la suite de ce sous-paragraphe, l'orthogonalité fera toujours référence à ce produit scalaire.

Définition : systèmes de polynômes orthogonaux. On appelle système de polynômes orthogonaux de E toute suite $(P_n)_{n \geq 0}$ de polynômes de $\mathbb{R}[X]$ vérifiant les deux propriétés :

- a) $\deg(P_n) = n$;
- b) $\langle P_m, P_n \rangle = 0$ pour tous entiers $m, n \geq 0$ distincts.

Premières propriétés.

- 1) Pour tout $n \geq 0$, les polynômes $P_0, \dots, P_n$ forment une base de $\mathbb{R}_n[X]$ puisqu'ils sont de degrés $0, 1, \dots, n$ tous distincts.
- 2) Une suite $(P_n)_{n \geq 0}$ de polynômes de $\mathbb{R}[X]$ est un système de polynômes orthogonaux si et seulement si :
 - a) $\deg(P_n) = n$;
 - b) pour tout entier $n \geq 0$, P_n est orthogonal à l'espace $\mathbb{R}_{n-1}[X]$
- 3) Si (P_n) et (Q_n) sont des systèmes de polynômes orthogonaux, il existe une suite (α_n) de réels non nuls tels que, pour tout $n \geq 0$, on ait $Q_n = \alpha_n P_n$. En effet, pour tout $n \geq 0$, les polynômes P_n et Q_n sont des polynômes de $\mathbb{R}_n[X]$ orthogonaux à $\mathbb{R}_{n-1}[X]$. Comme l'orthogonal de ce sous-espace dans $\mathbb{R}_n[X]$ est de dimension 1, on en déduit que P_n et Q_n sont colinéaires.

Calcul de P_0, P_1, P_2, P_3 . Comme ces polynômes sont définis à un coefficient multiplicatif près, on va chercher des polynômes unitaires.

On prend $P_0 = 1$.

P_1 est un polynôme de la forme $X + a$; la condition $\langle P_1, P_0 \rangle = 0$ donne $\int_I (t+a) dt = 0$ d'où $a = 0$ et $P_1 = X$.

P_2 est un polynôme de la forme $X^2 + aX + b$; les conditions $\langle P_2, P_0 \rangle = 0$ et $\langle P_2, P_1 \rangle = 0$ donnent $\int_I (t^2 + at + b) dt = 0$ et $\int_I (t^3 + at^2 + bt) dt = 0$ d'où $a = 0$, $b = -\frac{1}{3}$ et $P_2 = X^2 - \frac{1}{3}$.

P_3 est un polynôme de la forme $X^3 + aX^2 + bX + c$; les conditions $\langle P_3, P_0 \rangle = 0$, $\langle P_3, P_1 \rangle = 0$ et $\langle P_3, P_2 \rangle = 0$ donnent $P_3 = X^3 - \frac{3}{5}X$.

Existence de systèmes de polynômes orthogonaux. Considérons la suite des polynômes $(X^n)_{n \geq 0}$. L'orthogonalisation de Gram-Schmidt permet de construire une suite de polynômes $(P_n)_{n \geq 0}$ telle que P_n soit de la forme $X^n + \sum_{0 \leq k < n} \alpha_k X^k$ et orthogonal à $\mathbb{R}_{n-1}[X]$. C'est donc un système de polynômes orthogonaux.

On va maintenant montrer des résultats permettant de travailler plus commodément avec les systèmes de polynômes orthogonaux.

Proposition 1 : relation de récurrence. Soit (P_n) un système de polynômes orthogonaux. Pour tout $n \geq 0$, notons λ_n le coefficient du terme de degré n de P_n . Il existe des suites (a_n) , (b_n) , (c_n) de réels telles que, pour tout entier $n \geq 0$, la relation (R_n) soit satisfaite :

$$X P_{n+1} = a_{n+1} P_{n+2} + b_{n+1} P_{n+1} + c_{n+1} P_n \quad (R_n)$$

Démonstration. 1) Soit $n \geq 0$. Comme $X P_{n+1}$ est un polynôme de degré $n+2$, il s'écrit sous la forme $X P_{n+1} + \sum_{0 \leq k < n+2} \alpha_k P_k$.

2) Calcul de α_{n+2}

La comparaison des coefficients des termes de plus haut degré de la relation précédente donne $\lambda_{n+1} = \alpha_{n+2} \lambda_{n+2}$, d'où $\alpha_{n+2} = \frac{\lambda_{n+1}}{\lambda_{n+2}}$.

3) On a $\langle X P, Q \rangle = \langle P, X Q \rangle$ pour tous polynômes P, Q : il suffit de considérer les intégrales définies par les deux membres.

4) Calcul de α_k pour $k < n$

Pour $k < n$ on a : $\langle X P_{n+1}, P_k \rangle = \langle P_{n+1}, X P_k \rangle = 0$ car $X P_k$ est un polynôme de degré $< n+1$. D'autre part,

$$\langle X P_{n+1}, P_k \rangle = \langle \sum_{0 \leq i \leq n+2} \alpha_i P_i, P_k \rangle = \alpha_k \|P_k\|^2 ;$$

d'où $\alpha_k = 0$. On en déduit que $X P_{n+1}$ est bien de la forme $a_{n+1} P_{n+2} + b_{n+1} P_{n+1} + c_{n+1} P_n$.

5) Calcul de b_{n+1}

$$\langle X P_{n+1}, P_{n+1} \rangle = b_{n+1} \|P_{n+1}\|^2 \text{ donne } b_{n+1} = \frac{\langle X P_{n+1}, P_{n+1} \rangle}{\|P_{n+1}\|^2}.$$

6) Calcul de c_{n+1}

D'une part, $\langle X P_{n+1}, P_n \rangle = c_{n+1} \|P_n\|^2$. D'autre part,

$$\langle X P_{n+1}, P_n \rangle = \langle P_{n+1}, X P_n \rangle = \langle P_{n+1}, \frac{\lambda_n}{\lambda_{n+1}} P_{n+1} \rangle ;$$

$$\text{d'où } c_{n+1} = \frac{\lambda_n \|P_{n+1}\|^2}{\lambda_{n+1} \|P_n\|^2}.$$

□

Autre forme de la relation de récurrence. Soit (P_n) un système de polynômes orthogonaux tel que les coefficients λ_n des termes de plus haut degré des P_n soient strictement positifs. Il existe des suites (u_n) , (v_n) , (ω_n) de réels telles que, pour tout entier $n \geq 0$, on ait $u_n > 0$, $\omega_n < 0$ et

$$P_{n+2} = (u_n X + v_n) P_{n+1} + \omega_n P_n \quad (S_n)$$

Démonstration. Il suffit de diviser la relation (R_n) ci-dessus par a_{n+1} . □

Proposition 3 : parité des P_n . Soit (P_n) un système de polynômes orthogonaux. Pour tout entier n , le polynôme P_n est de la parité de n .

Démonstration. Utilisons l'orthogonalité plutôt que de faire un raisonnement par récurrence.

Définissons la suite de polynômes (Q_n) par $Q_n(X) = P_n(-X)$. On a $\deg(Q_n) = n$ et, pour tous entiers m, n :

$$\langle Q_m, Q_n \rangle = \int_I P_m(-t)P_n(-t)dt = \int_I P_m(u)P_n(u)du = \langle P_m, P_n \rangle$$

avec le changement de variable $u = -t$. On en déduit que la suite (Q_n) est un système de polynômes orthogonaux. Par conséquent, d'après la propriété 4, il existe une suite (α_n) telle que, pour tout entier $n \geq 0$, on ait $Q_n = \alpha_n P_n$. Comme $Q_n(X) = P_n(-X)$, on a $\alpha_n = (-1)^n$, en considérant les termes de plus haut degré des deux membres, ce qui permet de conclure. $\square$

Nouvelle forme de R_n et de S_n . L'intégrale d'une fonction impaire sur I étant nulle, on a : $\langle X P_{n+1}, P_{n+1} \rangle = \int_I t(P_{n+1}(t))^2 dt = 0$ d'où $b_{n+1} = v_n = 0$ pour tout $n \geq 0$; les relations R_n et S_n s'écrivent donc :

$$X P_{n+1} = a_{n+1} P_{n+2} + c_{n+1} P_n \quad (R_n)$$

$$P_{n+2} = u_n X P_{n+1} + \omega_n P_n \quad (S_n)$$

Autres produits scalaires définissant des systèmes de polynômes orthogonaux.

Beaucoup d'autres produits scalaires sur des espaces de fonctions permettent de définir des systèmes de polynômes orthogonaux. On se donne un intervalle I dans $\mathbb{R}$ et une fonction $p : I \rightarrow \mathbb{R}$ (appelée fonction poids) ayant les propriétés suivantes :

- a) p est une fonction définie au moins à l'intérieur de I ;
- b) p est une fonction continue et à valeurs strictement positives à l'intérieur de I ;
- c) p est une fonction intégrable sur I ;
- d) si I n'est pas bornée, p est à décroissance rapide à l'infini au sens suivant : $\lim_{x \rightarrow \infty} x^n p(x) = 0$ pour tout entier n , ce qui assure l'existence de $\int_I P(t)p(t)dt$ pour tout polynôme P .

On note E l'espace des fonctions continues sur I telles que $\int_I f(t)^2 p(t) dt$ existe et on munit E du produit scalaire défini par :

$$\langle f, g \rangle = \int_I f(t)g(t)p(t) dt.$$

Pour donner une idée de la diversité des situations, donnons quelques exemples :

- a) polynômes de Jacobi : $I =]-1, 1[$ ou $I = [-1, 1]$ et $p(x) = (1-x)^\alpha(1+x)^\beta$ avec $\alpha, \beta > -1$ avec les deux cas particuliers suivants ;
- b) polynômes de Legendre : $I = [-1, 1]$ et $p(x) = 1$;
- c) polynômes de Tchebycheff¹ : $I =]-1, 1[$ et $p(x) = \frac{1}{\sqrt{1-x^2}}$;

1. Pafnuty Tchebycheff (1821-1894)

- d) polynômes de Laguerre : $I = [0, \infty]$ et $p(x) = x^\alpha e^{-x}$, avec $\alpha > -1$;
e) polynômes de Hermite¹ : $I = \mathbb{R}$ et $p(x) = e^{-x^2}$.

Ce qui est remarquable est la similitude entre ces différents cas. On pourrait adapter ce paragraphe pour étudier toutes les situations précédentes, ce que propose l'exercice 16.14 pour les polynômes de Tchebycheff.

16.13.2 Polynômes de Legendre

On appelle polynômes de Legendre les polynômes définis pour tout entier $n \geq 0$ par la formule :

$$L_n(X) = \frac{1}{2^n n!} ((X^2 - 1)^n)^{(n)}$$

Nous avons dit au début de ce paragraphe comment Legendre avait introduit ses polynômes à partir de la formule donnant l'inverse de la distance de deux points, mais la définition précédente, due à Olinde Rodrigues (1794-1851) en 1815, est plus agréable.

Proposition 4. *La suite (L_n) est un système de polynômes orthogonaux.*

Démonstration. 1) On vérifie d'abord que $\deg(L_n) = n$: en effet, L_n est la dérivée n -ième du polynôme $(X^2 - 1)^n$ de degré $2n$.

2) Rappelons la formule de Leibniz donnant la dérivée r -ième du produit de deux fonctions u et v :

$$(uv)^{(r)} = \sum_{0 \leq k \leq r} C_r^k u^{(k)} v^{(r-k)}$$

3) Comme $(X^2 - 1)^n = (X - 1)^n (X + 1)^n$, la dérivée r -ième de $(X^2 - 1)^n$ est $\sum_{0 \leq k \leq r} C_r^k ((X - 1)^n)^{(k)} ((X + 1)^n)^{(r-k)}$.

4) Pour $r < n$, les différents termes de cette somme sont nuls en 1 et en -1 car $((X - 1)^n)^{(k)}$ est proportionnel à $(X - 1)^{n-k}$ et $((X + 1)^n)^{(r-k)}$ est proportionnel à $(X + 1)^{n-(r-k)}$.

5) Calculons $\langle L_n, X^k \rangle$ pour $k < n$. À l'aide d'intégrations par parties, on a :

2. Charles Hermite (1822-1901)

$$\begin{aligned}
 \langle 2^n n! L_n, X^k \rangle &= \int_I t^k ((t^2 - 1)^n)^{(n)} dt \\
 &= [t^k ((t^2 - 1)^n)^{(n-1)}]_1^1 - k \int_I t^{k-1} ((t^2 - 1)^n)^{(n-1)} dt \\
 &= -k \int_I t^{k-1} ((t^2 - 1)^n)^{(n-1)} dt \\
 &= \dots \\
 &= (-1)^k k! \int_I ((t^2 - 1)^n)^{(n-k)} dt \\
 &= 0
 \end{aligned}$$

6) Les calculs ci-dessus montrent que L_n est orthogonal à $\mathbb{R}_{n-1}[X]$, d'où le résultat. $\square$

Calcul de $L_n(1)$ et de $L_n(-1)$. Le 3) ci-dessus a montré comment calculer les dérivées de $(X^2 - 1)^n$; $L_n(1)$ est la valeur en 1 de $\frac{1}{2^n n!} ((X-1)^n)^{(n)} (X+1)^n$; on trouve $L_n(1) = 1$. De même, $L_n(-1)$ est la valeur en -1 de $\frac{1}{2^n n!} ((X+1)^n)^{(n)} (X-1)^n$; on trouve $L_n(-1) = (-1)^n$.

Norme de L_n . On a $\|L_n\|^2 = \frac{2}{2n+1}$. Ce résultat s'obtient avec une succession d'intégrations par parties, nous l'admettrons.

Relation de récurrence des polynômes de Legendre. Avec les résultats précédents, la relation (R_n) s'écrit :

$$(2n+1)X L_n = (n+1)L_{n+1} + nL_{n-1}$$

Cette relation donne un calcul facile des premiers polynômes de Legendre. À partir de $L_0 = 1$ et de $L_1 = X$, on trouve $L_2 = \frac{3X^2 - 1}{8}$ et $L_3 = \frac{15X^3 - 9X}{2}$. Ces polynômes sont bien proportionnels à ceux que nous avons calculé au début.

Proposition 5 : racines des polynômes de Legendre. Pour tout $n > 0$, L_n est un polynôme scindé à racines simples et ces racines appartiennent toutes à I .

Démonstration. Pour $n > 0$, notons $a_1, \dots, a_r$ les racines de L_n appartenant à I et d'ordre impair. Supposons que $r < n$ et posons $R = \prod_{1 \leq i \leq r} (X - a_i)$; on a $\deg(R) = r < n$. Toutes les racines du polynôme $L_n R$ dans I sont d'ordre pair, donc $L_n R$ garde un signe constant sur I et n'est pas nul, donc $\langle L_n, R \rangle = \int_I L_n(t) R(t) dt \neq 0$ ce qui contredit l'orthogonalité de L_n et de R . $\square$

16.13.3 Calcul approché d'intégrales

C'est le 16 septembre 1814 que Gauss présente à Göttingen son procédé de calcul numérique des intégrales.

Nous allons décrire cette méthode pour le calcul numérique d'intégrales de la forme $\int_I f(t) dt$ où $I = [-1, 1]$ (on peut toujours se ramener à intégrer sur cet intervalle par un changement linéaire de variable).

L'idée de la méthode est de rechercher n points $a_1, \dots, a_n$ de I et n rels $A_1, \dots, A_n$ tels que $A_1 f(a_1) + \dots + A_n f(a_n)$ soit une bonne approximation de $\int_I f(t) dt$.

Pour obtenir cette qualité d'approximation, on cherche d'abord à avoir une formule exacte pour tout polynôme P de $\mathbb{R}_{2n-1}[X]$, c'est-à-dire :

$$\int_I P(t) dt = A_1 P(a_1) + \dots + A_n P(a_n) \quad (R_n).$$

Posons $T = \prod_{1 \leq i \leq n} (X - a_i)$ et $T_k = \frac{T}{X - a_k}$ pour $1 \leq k \leq n$.

Comme $\deg(T T_k) = 2n - 1$, on doit avoir, pour $1 \leq k \leq n$:

$$\langle T, T_k \rangle = \int_I T(t) T_k(t) dt = \sum_{1 \leq i \leq n} A_i T(a_i) T_k(a_i) = 0$$

Par conséquent, T est orthogonal aux T_k pour $1 \leq k \leq n$. Comme ces polynômes forment une base de $\mathbb{R}_{n-1}[X]$, T est colinéaire au polynôme de Legendre L_n , ce qui impose que les a_k , $1 \leq k \leq n$, soient les racines de L_n . Ainsi les polynômes de Legendre s'introduisent naturellement dans le calcul approché d'intégrales.

La valeur des A_k pour $1 \leq k \leq n$ est donnée par $\int_I T_k(t) dt = A_k T_k(a_k)$. Avec ces A_k , on aura donc des valeurs exactes pour les intégrales des T_k ; comme les T_k forment une base de $\mathbb{R}_{n-1}[X]$, la condition (R_n) sera remplie pour les polynômes de cet espace. Il faut maintenant vérifier que les A_k déterminés permettent de remplir la condition (R_n) pour un polynôme P de $\mathbb{R}_{2n-1}[X]$. La division euclidienne de P par T donne $P = TQ + R$, avec $\deg(Q) < n$ et $\deg(R) < n$ et on a :

$$\begin{aligned} \int_I P(t) dt &= \int_I (T(t)Q(t) + R(t)) dt = \int_I R(t) dt \\ &= \sum_{1 \leq i \leq n} A_i R(a_i) = \sum_{1 \leq i \leq n} A_i P(a_i) \end{aligned}$$

Il resterait à montrer la qualité de cette approximation pour calculer des intégrales de fonctions. On peut montrer qu'elle est meilleure que la précision donnée par la méthode de Simson. Seulement, elle demande de calculer avec les racines des polynômes de Legendre, ce qui suppose de calculer ces racines avec une bonne

approximation, puis d'évaluer la fonction en ces points. On avait construit des tables donnant ces racines avec beaucoup de décimales. Le nombre de décimales rendait la méthode difficile à mettre en œuvre car, jusque vers 1960, les calculs devaient être faits à la main ; la méthode de Gauss était peu utilisée. Une autre façon de calculer des intégrales était alors de construire des machines, de plus en plus compliquées, pour intégrer mécaniquement. Depuis, les ordinateurs, qui manient les nombres avec beaucoup de décimales sans états d'âme, ont permis à la méthode de Gauss de prendre place parmi les méthodes de calcul approché d'intégrales. D'autres problèmes limitent l'utilisation de la méthode : on ne peut obtenir une bonne précision si la fonction est peu dérivable et la simple méthode des trapèzes est souvent préférée.

Exemple. Soit $f: I = [-1, 1] \rightarrow \mathbb{R}$ la fonction définie par $f(t) = \frac{1}{t^2 + 2t + 5}$. On peut calculer directement $\int_I f(t) dt$ puisqu'une primitive de f est $t \mapsto \frac{1}{2} \arctan \frac{t+1}{2}$. On trouve $\frac{\pi}{8} = 0,392\,699$.

La méthode de Gauss avec les racines de $L_3 = \frac{15X^3 - 9X}{2}$ donne $\int_I f(t) dt = A_1 f\left(-\frac{\sqrt{3}}{\sqrt{5}}\right) + A_2 f(0) + A_3 f\left(\frac{\sqrt{3}}{\sqrt{5}}\right)$.

La condition R_3 pour les fonctions $t \mapsto 1$, $t \mapsto t$, $t \mapsto t^2$, donne trois équations permettant d'obtenir $A_1 = A_3 = \frac{5}{9}$ et $A_2 = \frac{8}{9}$. Dans la pratique, ces nombres ont été calculés une fois pour toutes.

Finalement, on calcule $\frac{5}{9} f\left(-\frac{\sqrt{3}}{\sqrt{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\frac{\sqrt{3}}{\sqrt{5}}\right)$. On trouve 0,392 263.

Pour comparaison, la formule des trois niveaux (voir exercice 10.5) donne : 0,391 666.

➤ Vers le chapitre 17

Cette deuxième partie, correspondant à la deuxième année de licence, se termine. Il nous reste à rendre hommage à Karl Friedrich Gauss, le prince des mathématiciens, dont nous avons exposé les résultats à plusieurs reprises.

Exercices

16.1

1) Les applications suivantes déterminent-elles des produits scalaires sur $E = \mathbb{R}[X]$?

a) $\varphi(P, Q) = P(0)Q(0) + P(1)Q(1)$;

b) $\varphi(P, Q) = \int_0^1 (P(t)Q(t) + P'(t)Q'(t))dt$.

2) Les applications définies par les formules suivantes déterminent-elles des produits scalaires sur $\mathbb{R}^3$?

a) $\varphi(u, u') = xx' + 2yy' + 4zz'$;

b) $\varphi(u, u') = xx' + 2xy' + 2x'y + yy' + 4zz'$;

(où $u = (x, y, z)$ et $u' = (x', y', z')$ sont des vecteurs de $\mathbb{R}^3$) ?

3) a) En utilisant la méthode de Gauss, déterminer pour quelles valeurs de a la forme $\varphi(u, u') = xx' + 6yy' + azz' - 2x'y - 2xy' - 3xz' - 3x'z$ définit un produit scalaire.

b) Donner la matrice Φ de φ pour $a = 28$ dans la base canonique de $\mathbb{R}^3$; déterminer une base dans laquelle la matrice Φ' de φ soit diagonale et donner une relation entre les deux matrices de φ .

16.2 Liberté et orthogonalité

Utiliser des arguments d'orthogonalité pour résoudre les questions suivantes.

a) Montrer que, pour tout entier n , les fonctions $c_k : x \mapsto \cos kx$, $0 \leq k \leq n$ et $s_k : x \mapsto \sin kx$, $1 \leq k \leq n$ forment une famille libre de $\text{Cont}([0, 2\pi], \mathbb{R})$.

b) Montrer que, pour tout entier n , les fonctions $e_k : x \mapsto e^{ikx}$, $-n \leq k \leq n$ forment une famille libre de $\text{Cont}([0, 2\pi], \mathbb{C})$.

16.3

Soit E un espace euclidien et soient F_1 et F_2 deux sous-espaces de E .

Montrer que $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$ et que $F_1^\perp + F_2^\perp = (F_1 \cap F_2)^\perp$.

16.4 Inégalité

Montrer que, pour tout entier $n \geq 1$, on a :

$$\sum_{1 \leq k \leq n} k\sqrt{k} \leq n(n+1) \frac{\sqrt{2n+1}}{2\sqrt{3}}.$$

16.5 Symétries, rotations et demi-tours

1) a) Soit $u = (a, b, c)$ un vecteur non nul de $\mathbb{R}^3$. Déterminer la matrice de la symétrie orthogonale s par rapport au plan H orthogonal à u , puis la matrice de la symétrie orthogonale s_1 par rapport à u .

b) Soit $f \in O(\mathbb{R}^3)$. Décrire les applications $f \circ s \circ f^{-1}$ et $f \circ s_1 \circ f^{-1}$.

2) Soient $f \in O(\mathbb{R}^3)$ et $r \in SO(\mathbb{R}^3)$ une rotation d'axe u et d'angle θ . Décrire l'application $f \circ r \circ f^{-1}$.

3) Montrer que le groupe $O(\mathbb{R}^3)$ est engendré par les demi-tours.

16.6 Calcul de distance

On note E l'espace $\mathbb{R}_n[X]$ muni du produit scalaire $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$. On pose $F = \text{Vect}(X, \dots, X^n)$ et on cherche à déterminer la distance $d = d(1, F)$ (ici 1 représente le polynôme 1 de $\mathbb{R}_n[X]$).

On pose $S(X) = \frac{(X-1) \dots (X-n)}{(X+1) \dots (X+n+1)}$. La décomposition en éléments simples de S

s'écrit (voir 20.5) : $S(X) = \frac{a_0}{X+1} + \dots + \frac{a_n}{X+n+1}$.

a) Calculer a_0 .

b) Montrer que le polynôme $T(X) = a_0 + \dots + a_n X^n$ est orthogonal à F . En déduire $F^\perp$.

c) Montrer que $d^2 = \frac{1}{a_0^2} \|T\|^2$.

d) En déduire d .

16.7 Espaces de fonctions

1) Pour tout réel $x > 0$, on note E_x l'espace vectoriel des fonctions de classe C^1 sur $[0, x]$ nulles en 0.

a) Montrer que $\langle f, g \rangle_x = \int_{[0,x]} f'(t)g'(t) dt$ définit un produit scalaire sur E_x .

b) Soit $f \in E_1$. Montrer que $\sup_{x \in [0,1]} |f(x)| \leq \left(\int_{[0,1]} (f'(t))^2 dt \right)^{1/2}$.

2) On pose $I = [0, 1]$. Soit E l'espace des fonctions continues $f: I \rightarrow \mathbb{R}$ telles que $\int_I \frac{f(t)^2}{t} dt$ converge.

a) Montrer que E est un espace vectoriel.

b) Montrer que $\langle f, g \rangle = \int_I \frac{f(t)g(t)}{t} dt$ définit un produit scalaire sur E .

16.8 Espace l^2

On note l^2 l'ensemble des suites $(u_n)_{n \geq 0}$ de nombres réels telles que $\sum_{n \geq 0} u_n^2$ converge. Montrer que l^2 est un espace préhilbertien réel pour un produit scalaire qu'on définira.

C'est Hilbert qui a commencé, en 1906, l'étude de cet espace de dimension infinie. L'année suivante, Schmidt a complété les travaux de Hilbert.

16.9 Minimisation

On pose $I = [-1, 1]$. Trouver le polynôme unitaire de degré 4 qui minimise $\int_I (P(t))^2 dt$.

16.10 Déterminant et trace

1) Soient $n \geq 3$ un entier et A une matrice symétrique réelle telle $A^n = I$. Montrer que $A^2 = I$.

2) Soient n un entier et A une matrice carrée symétrique réelle de $M_n(\mathbb{R})$ associée à un produit scalaire.

a) Montrer que les valeurs propres de A sont toutes strictement positives.

b) Montrer que $\det(A) \leq \left(\frac{\text{Tr}(A)}{n} \right)^n$.

16.11 Opérateur autoadjoint

On note E l'espace euclidien $\mathbb{R}_n[X]$ muni du produit scalaire $\langle P_1, P_2 \rangle = \int_0^1 P_1(t)P_2(t) dt$. On définit $L : E \rightarrow E$ par

$$L(P)(x) = \int_0^1 (x+t)^n P(t) dt.$$

Montrer que L est autoadjoint.

16.12 Calculs de sommes

Soient f et g deux endomorphismes de l'espace euclidien $\mathbb{R}^n$. On note $B = (e_1, \dots, e_n)$, $B' = (e'_1, \dots, e'_n)$ deux bases orthonormées de $\mathbb{R}^n$ et f^* l'adjoint de f .

a) Montrer que $\sum_{1 \leq i \leq n} \langle f(e_i), g(e'_j) \rangle^2 = \|(f^* \circ g)(e'_j)\|^2$ pour tout j , $1 \leq j \leq n$.

b) Montrer que $\sum_{1 \leq i, j \leq n} \langle f(e_i), g(e'_j) \rangle^2$ ne dépend pas des bases choisies.

16.13 Conservation des normes et linéarité

Soient E un espace euclidien et $f: E \rightarrow E$ une application. Montrer que, si f conserve les normes, alors f est linéaire (montrer que $\|f(\lambda u + \mu v) - \lambda f(u) - \mu f(v)\|^2 = 0$).

16.14 Déterminant de Gram

Soit E un espace euclidien et soit $S = (u_1, \dots, u_p)$ une famille de vecteurs de E . On appelle déterminant de Gram de la famille S le déterminant $G(S)$ de la matrice carrée $(\langle u_i, u_j \rangle)_{1 \leq i, j \leq p}$.

a) On suppose que S est une famille orthogonale. Calculer $G(S)$ en fonction des $\|u_i\|$, $1 \leq i \leq p$.

b) Montrer que $G(S)$ ne change pas si on ajoute à l'un des vecteurs de S une combinaison linéaire des autres vecteurs de S .

c) On construit à partir de S une famille orthogonale $S' = (u'_1, \dots, u'_p)$ par orthogonalisation de Gram-Schmidt. Montrer que $G(S') = G(S)$.

d) Soit u un vecteur de E et $T = (u_1, \dots, u_p, u)$. Montrer que

$$d(u, \text{Vect}(S))^2 = \frac{G(T)}{G(S)}$$

e) On se place dans l'espace euclidien $\mathbb{R}^2$ muni de la base canonique. Utiliser le déterminant de Gram pour calculer l'aire du triangle de sommets $0 = (0, 0)$, $u = (3, 2)$, $v = (1, 4)$.

16.15 Polynômes de Tchebycheff

On pose $I =]-1, 1[$ et on définit $p: I \rightarrow \mathbb{R}$ par $p(t) = \frac{1}{\sqrt{1-t^2}}$. On note E l'espace

des fonctions continues $f: I \rightarrow \mathbb{R}$ telles que $\int_I f(t)^2 p(t) dt$ converge.

1) Montrer qu'on définit un produit scalaire sur E en posant, pour f et g dans E : $\langle f, g \rangle = \int_I f(t)g(t)p(t) dt$. On considérera dans la suite E muni de ce produit scalaire.

2) Vérifier qu'on peut reprendre entièrement le sous-paragraphe 16.13.1 pour ce produit scalaire.

Pour $n \geq 0$, on définit T_n par $T_n(x) = \cos(n \arccos x)$ pour $x \in I$.

3) Donner une relation de récurrence entre T_{n+1} , T_n et T_{n-1} pour $n \geq 1$.

4) Montrer que les T_n forment un système de polynômes orthogonaux. Quel est le coefficient du terme de plus haut degré de T_n ?

5) a) Montrer que $\sup_{x \in I} |T_n(x)| = 1$.

b) On note F_n l'ensemble des polynômes de degré n ayant leur coefficient du terme de degré n égal à celui de T_n . Montrer que, pour tout P de F_n , $\sup_{x \in I} |P(x)| \geq \sup_{x \in I} |T_n(x)|$.

Cette dernière propriété est une propriété célèbre des polynômes de Tchebycheff qui se démontre sans utiliser le produit scalaire.

Solutions

16.1 1) a) Non. Il s'agit d'une forme bilinéaire symétrique positive, mais $\varphi(P, P)$ peut être nul pour des polynômes non nuls comme $X^2 - X$ et leurs multiples non nuls.

b) Oui.

2) a) Oui.

b) Non. On a $\varphi(u, u) = x^2 + 4xy + y^2 + 4z^2 = (x + 2y)^2 - 3y^2 + 4z^2$. On trouve facilement des vecteurs u tels que $\varphi(u, u) < 0$: par exemple, $u = (-2, 1, 0)$.

3) a) On trouve

$\varphi(u, u) = x^2 + 6y^2 + az^2 - 4xy - 6xz = (x - 2y - 3z)^2 + 2(y - 3z)^2 + (a - 27)z^2$. Donc φ est un produit scalaire pour $a > 27$.

$$\text{b) On trouve } \Phi = \begin{pmatrix} 1 & -2 & -3 \\ -2 & 6 & 0 \\ -3 & 0 & 28 \end{pmatrix} \quad \Phi' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{En posant } P = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & -3 & 1 \end{pmatrix}, \text{ on a } \Phi = P \Phi'^t P.$$

16.2 a) Les formules $2 \cos a \cos b = \cos(a + b) + \cos(a - b)$,

$2 \sin a \sin b = \cos(a - b) - \cos(a + b)$ et $2 \sin a \cos b = \sin(a + b) + \sin(a - b)$

permettent de calculer $\langle c_k, c_l \rangle = \int_0^{2\pi} \cos kx \cos lx \, dx$, $\langle s_k, s_l \rangle = \int_0^{2\pi} \sin kx \sin lx \, dx$,

égaux à 0 si $k \neq l$, à π si $k = l$ et $\langle s_k, c_l \rangle = \int_0^{2\pi} \sin kx \cos lx \, dx = 0$.

Si une combinaison linéaire f de c_k et de s_l est la fonction nulle : $f = \sum_{0 \leq k \leq n} \lambda_k c_k +$

$\sum_{1 \leq l \leq n} \mu_l s_l = 0$, on obtient $\lambda_k = 0$ en calculant $\langle f, c_k \rangle = \int_0^{2\pi} f(x) \cos kx \, dx$ et $\mu_l = 0$

en calculant $\langle f, s_l \rangle = \int_0^{2\pi} f(x) \sin lx \, dx$.

b) Si $f = \sum_{-n \leq k \leq n} \lambda_k e_k = 0$, on obtient $\lambda_k = 0$ en calculant $\langle f, e_{-k} \rangle = \int_0^{2\pi} f(x) e^{-ikx} \, dx$.

16.3 Si u est orthogonal à $F_1 + F_2$, il est orthogonal à F_1 et à F_2 donc $(F_1 + F_2)^\perp \subset F_1^\perp \cap F_2^\perp$. Réciproquement, soit $u \in F_1^\perp \cap F_2^\perp$. Tout vecteur v de $F_1 + F_2$ s'écrit $v_1 + v_2$ avec $v_1 \in F_1$ et $v_2 \in F_2$; on a $\langle u, v_1 + v_2 \rangle = \langle u, v_1 \rangle + \langle u, v_2 \rangle = 0$ donc $u \in (F_1 + F_2)^\perp$. D'où l'égalité demandée.

Comme pour l'exercice 10.1.c), on obtient la seconde égalité en remplaçant F_1 et F_2 par leurs orthogonaux dans la première égalité.

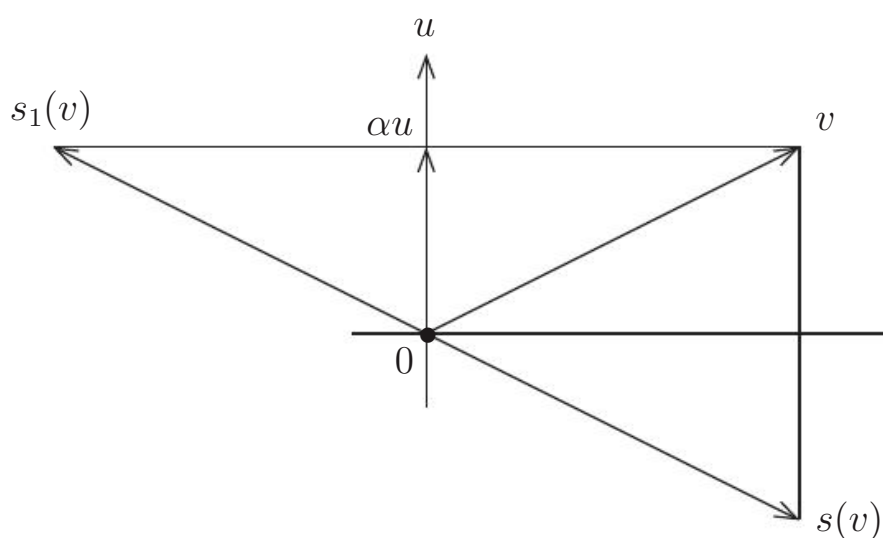
16.4 Appliquer l'inégalité de CBS dans $\mathbb{R}^n$ muni du produit scalaire usuel aux vecteurs $u = (1, 2, \dots, n)$ et $v = (1, \sqrt{2}, \dots, \sqrt{n})$; puis utiliser les égalités

$$\sum_{1 \leq k \leq n} k = \frac{n(n+1)}{2} \text{ et } \sum_{1 \leq k \leq n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

16.5 1) a) Soit $u = (x, y, z)$ un vecteur de $\mathbb{R}^3$. Le projeté orthogonal de v sur l'axe engendré par u est αu avec $\alpha = \frac{\langle u, v \rangle}{\langle u, u \rangle} = \frac{ax + by + cz}{a^2 + b^2 + c^2}$ (voir 16.2). On a $s(v) = v - 2\alpha u$, ce qui donne la matrice :

$$-\frac{1}{a^2 + b^2 + c^2} \begin{pmatrix} a^2 - b^2 - c^2 & 2ab & 2ac \\ 2ab & b^2 - a^2 - c^2 & 2bc \\ 2ac & 2bc & c^2 - a^2 - b^2 \end{pmatrix}$$

Comme $s_1 = -s$, la matrice de s_1 est la matrice ci-dessus changée de signe. Voici un dessin illustrant la situation dans le plan de u et v .



b) Posons $s' = f \circ s \circ f^{-1}$, $s'_1 = f \circ s_1 \circ f^{-1}$, $H' = f(H)$, $u' = f(u)$. Comme $f \in O(\mathbb{R}^3)$, H' et u' sont orthogonaux. Un vecteur de $\mathbb{R}^3$ s'écrit $au' + v'$ avec a réel et $u' \in H'$. On

a $s'(au' + v') = f(s(au + v))$ avec $v = f^{-1}(v') \in H$, d'où $s'(au' + v') = -au' + v'$. De même, $s'_1(au' + v') = au' - v'$. On voit que s' est la symétrie orthogonale par rapport à H' et que s'_1 est la symétrie orthogonale par rapport à u' .

2) Si B est une base orthogonale de $\mathbb{R}^3$, $f(B)$ en est aussi une. Si on choisit B de façon que la matrice de r soit $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$, alors la matrice de

$f \circ r \circ f^{-1}$ dans la base $f(B)$ est aussi M (calculer les images des vecteurs de la base $f(B)$), donc $f \circ r \circ f^{-1}$ est une rotation d'axe $f(u)$ et d'angle θ .

3) Un élément $f \in O(\mathbb{R}^3)$ est une rotation. Soit ε son axe. Dans le plan H orthogonal à ε , f induit une rotation plane qu'on sait pouvoir décomposer (voir 16.10) en un produit $s' \circ s$ de symétries planes d'axes u, u' dans H . Notons σ, σ' les demi-tours définis par u et u' respectivement. Les restrictions de σ et σ' à H sont égales à s et s' respectivement et $\sigma'(\sigma(\varepsilon)) = \varepsilon$; par conséquent, $f = \sigma' \circ \sigma$; on voit que f est un produit de demi-tours, d'où le résultat.

16.6 a) En évaluant la fraction rationnelle

$$(X+1)S = \frac{(X-1) \dots (X-n)}{(X+2) \dots (X+n+1)}$$

$$= a_0 + (X+1) \left(\frac{a_1}{X+2} + \dots + \frac{a_n}{X+n+1} \right) \text{ en } -1,$$

$$\text{on obtient } a_0 = \frac{(-1)^n (n+1)!}{n!} = (-1)^n (n+1).$$

b) Pour $k = 1, \dots, n$, on a

$$\langle T, X^k \rangle = \int_0^1 (a_0 t^k + \dots + a_n t^{n+k}) dt$$

$$= \frac{a_0}{k+1} + \dots + \frac{a_n}{k+n+1} = S(k) = 0.$$

Comme $\dim(F^\perp) = 1$, on a $F^\perp = \text{Vect}(T)$.

c) Comme $E = F \oplus F^\perp$, on sait que 1 s'écrit $P + \alpha T$ où $P \in F$. La comparaison des termes constants donne $1 = \alpha a_0$ d'où $\alpha = \frac{1}{a_0}$. D'autre part, $d^2 = \alpha^2 \|T\|^2$. Donc $d^2 = \frac{1}{a_0^2} \|T\|^2$.

d) On a $\|T\|^2 = \langle T, T \rangle = \langle T, T - a_0 \rangle + \langle T, a_0 \rangle = \langle T, a_0 \rangle$ car $T \in F^\perp$ et $T - a_0 \in F$.

Donc

$$\|T\|^2 = a_0 \int_0^1 (a_0 + \dots + a_n t^n) dt = a_0 S(0) = a_0 \frac{(-1)^n n!}{(n+1)!} = 1.$$

On en déduit $d = \frac{1}{n+1}$.

16.7 1) a) Si $\langle f, f \rangle_x = 0$, on a $\int_{[0, x]} f'(t)^2 dt = 0$ donc f' est nulle. Comme $f(0) = 0$, f est aussi nulle. Les autres vérifications sont faciles.

b) Pour tout réel x de $[0, 1]$, l'inégalité de CBS donne $(\langle f, \text{id} \rangle_x)^2 \leq \langle f, f \rangle_x \langle \text{id}, \text{id} \rangle_x$, soit $\left(\int_0^x f'(t) dt\right)^2 \leq x \int_0^x (f'(t))^2 dt$. On en déduit que, pour tout x de $[0, 1]$, on a $(f(x))^2 \leq x \int_{[0, 1]} (f'(t))^2 dt$ d'où le résultat.

2) a) La seule difficulté est de montrer que E est stable par somme. Soient f et g dans E ; les intégrales $\int_I \frac{(f(t))^2}{t} dt$ et $\int_I \frac{(g(t))^2}{t} dt$ sont donc convergentes. Pour montrer que $f + g \in E$, on est ramené à montrer que $\int_I \frac{(f(t)g(t))}{t} dt$ converge. Pour tout $\varepsilon > 0$ de I on a :

$$\left(\int_{[\varepsilon, 1]} \frac{|f(t)g(t)|}{t} dt \right)^2 \leq \int_{[\varepsilon, 1]} \frac{f(t)^2}{t} dt \int_{[\varepsilon, 1]} \frac{g(t)^2}{t} dt$$

en appliquant l'inégalité de CBS aux fonctions $\frac{f(t)}{\sqrt{t}}$ et $\frac{g(t)}{\sqrt{t}}$ sur l'intervalle $[\varepsilon, 1]$, ce qui permet de conclure. On aurait pu utiliser simplement ici l'inégalité $2|ab| \leq a^2 + b^2$.

b) Si $\langle f, f \rangle = \int_I \frac{(f(t))^2}{t} dt = 0$, on a $\int_{[\varepsilon, 1]} \frac{f(t)^2}{t} dt = 0$. Donc f est nulle sur $[\varepsilon, 1]$ et par continuité $f(0) = 0$.

16.8 On sait que les suites (u_n) de nombres réels forment un espace vectoriel. Il faut montrer que l^2 en est un sous-espace. La difficulté est de montrer que la somme de deux suites de l^2 est dans l^2 .

Soient $u = (u_n)$ et $v = (v_n)$ deux suites de l^2 .

Posons $U = \sum_{n \geq 0} u_n^2$ et $V = \sum_{n \geq 0} v_n^2$.

Pour tout entier N , u et v définissent des éléments $u_N = (|u_0|, \dots, |u_N|)$ et $v_N = (|v_0|, \dots, |v_N|)$ de $\mathbb{R}^N$. L'inégalité de CBS donne :

$$\left(\sum_{0 \leq k \leq N} |u_k v_k| \right)^2 \leq \left(\sum_{0 \leq k \leq N} u_k^2 \right) \left(\sum_{0 \leq k \leq N} v_k^2 \right) \leq UV$$

On en déduit que la série $\sum_{0 \leq k} u_k v_k$ est absolument convergente donc la série $\sum_{0 \leq k} (u_k + v_k)^2 = \sum_{0 \leq k} (u_k^2 + 2u_k v_k + v_k^2)$ converge, ce qui montre que $u + v \in l^2$. Pour la majoration, on peut ici aussi utiliser l'inégalité $2|u_k v_k| \leq (u_k^2 + v_k^2)$. Le produit scalaire sur l^2 est défini par $\langle u, v \rangle = \sum_{0 \leq k} u_k v_k$. Les dernières vérifications ne posent aucun problème.

16.9 On va suivre la démarche de l'exemple de 16.8. On considère le produit scalaire sur $\mathbb{R}_4[X]$ défini par $\langle P, Q \rangle = \int_I P(t)Q(t) dt$. Un polynôme unitaire de degré 4 s'écrit $X^4 - S$ avec $\deg(S) \leq 3$. On doit chercher le projeté orthogonal S de X^4 sur $\mathbb{R}_3[X]$. On l'obtient en écrivant que $X^4 - S$ est orthogonal à 1, X , X^2 , X^3 . En posant $S = aX^3 + bX^2 + cX + d$, on est conduit à la résolution d'un système linéaire. On trouve $a = c = 0$, $b = \frac{6}{7}$ et $d = -\frac{3}{35}$. Le minimum est obtenu pour $P = X^4 - \frac{6}{7}X^2 + \frac{3}{35}$. Le minimum est $\frac{128}{11\,025}$.

16.10 1) On sait que A est semblable à une matrice diagonale D . Chaque valeur propre λ de A est réelle et vérifie $\lambda^n = 1$. On a donc $\lambda = \pm 1$, donc $\lambda^2 = 1$. On en déduit $D^2 = I$ puis $A^2 = I$.

2) a) Dans une base où A est une matrice diagonale D de termes diagonaux $\lambda_1, \dots, \lambda_n$ l'expression du produit scalaire est $\sum_{1 \leq i \leq n} \lambda_i x_i^2$. Les λ_i sont strictement positifs.

b) On a $\det(A) = \prod_{1 \leq i \leq n} \lambda_i$ et $\frac{\text{Tr}(A)}{n} = \frac{\sum_{1 \leq i \leq n} \lambda_i}{n}$. L'inégalité entre la moyenne géométrique et la moyenne arithmétique des λ_i donne le résultat.

16.11 On veut montrer que, pour tous P_1 et P_2 de E , on a : $\langle L(P_1), P_2 \rangle = \langle P_1, L(P_2) \rangle$.

Le premier membre est égal à :

$$\int_0^1 \left(\int_0^1 (x+t)^n P_1(t) dt \right) P_2(x) dx = \sum_{0 \leq k \leq n} C_n^k \int_0^1 t^{n-k} P_1(t) dt \int_0^1 x^k P_2(x) dx.$$

Le second membre est égal à :

$$\int_0^1 P_1(x) \left(\int_0^1 (x+t)^n P_2(t) dt \right) dx = \sum_{0 \leq k \leq n} C_n^k \int_0^1 x^{n-k} P_1(x) dx \int_0^1 t^k P_2(t) dt.$$

Ils sont donc égaux.

16.12

$$\text{a)} \sum_{1 \leq i \leq n} \langle f(e_i), g(e'_j) \rangle^2 = \sum_{1 \leq i \leq n} \langle e_i, (f^* \circ g)(e'_j) \rangle^2 = \|(f^* \circ g)(e'_j)\|^2$$

d'après la définition de f^* et le théorème de Pythagore.

$$\begin{aligned} \text{b)} \quad \sum_{1 \leq i, j \leq n} \langle f(e_i), g(e'_j) \rangle^2 &= \sum_{1 \leq j \leq n} \langle (f^* \circ g)(e'_j), (f^* \circ g)(e'_j) \rangle \\ &= \sum_{1 \leq j \leq n} \langle e'_j, (f^* \circ g)^*((f^* \circ g)(e'_j)) \rangle \\ &= \text{Tr}((f^* \circ g)^* \circ (f^* \circ g)) \end{aligned}$$

le résultat final est donc : $\text{Tr}(g^* \circ f \circ f^* \circ g)$.

16.13 Si a, b, c sont trois vecteurs de E , la bilinéarité du produit scalaire montre que

$$\begin{aligned} \|a - b - c\|^2 &= \langle a - b - c, a - b - c \rangle \\ &= \|a\|^2 + \|b\|^2 + \|c\|^2 - 2\langle a, b \rangle - 2\langle a, c \rangle + 2\langle b, c \rangle. \end{aligned}$$

Pour tous u, v de E et tous λ, μ réels, on a donc :

$$\begin{aligned} &\|f(\lambda u + \mu v) - \lambda f(u) - \mu f(v)\|^2 = \\ &\|f(\lambda u + \mu v)\|^2 + \lambda^2 \|f(u)\|^2 + \mu^2 \|f(v)\|^2 - \\ &2\lambda \langle f(\lambda u + \mu v), f(u) \rangle - 2\mu \langle f(\lambda u + \mu v), f(v) \rangle + 2\lambda\mu \langle f(u), f(v) \rangle. \end{aligned}$$

Comme f conserve les normes, donc les produits scalaires, on a :

$$\begin{aligned} \|f(\lambda u + \mu v) - \lambda f(u) - \mu f(v)\|^2 &= \|\lambda u + \mu v\|^2 + \lambda^2 \|u\|^2 + \mu^2 \|v\|^2 - \\ &2\lambda \langle \lambda u + \mu v, u \rangle - 2\mu \langle \lambda u + \mu v, v \rangle + 2\lambda\mu \langle u, v \rangle. \end{aligned}$$

D'après l'égalité initiale, cette dernière expression est égale à $\|\lambda u + \mu v - \lambda u - \mu v\|^2$, donc est nulle. On en déduit que $\|f(\lambda u + \mu v) - \lambda f(u) - \mu f(v)\|^2 = 0$ puis que f est linéaire.

16.14 a) La matrice $(\langle u_i, u_j \rangle)_{1 \leq i, j \leq p}$ est une matrice diagonale et $\langle u_i, u_i \rangle = \|u_i\|^2, 1 \leq i \leq p$. Donc $G(S) = \prod_{1 \leq i \leq p} \|u_i\|^2$.

Le déterminant de Gram est donc le carré de l'aire du rectangle défini par S en dimension 2, le carré du parallélepède défini par S en dimension 3.

b) On peut se ramener à montrer que $G(S) = G(S')$ dans le cas où $S' = (u'_1, u_2, \dots, u_p)$ et $u'_1 = u_1 + \sum_{2 \leq i \leq p} \alpha_i u_i$. On a alors :

$$G(S') = \begin{vmatrix} \langle u'_1, u'_1 \rangle & \langle u'_1, u_2 \rangle & \cdots & \langle u'_1, u_p \rangle \\ \langle u_2, u'_1 \rangle & \langle u_2, u_2 \rangle & \cdots & \langle u_2, u_p \rangle \\ \vdots & \cdots & \cdots & \cdots \\ \langle u_p, u'_1 \rangle & \cdots & \cdots & \langle u_p, u_p \rangle \end{vmatrix}.$$

En retranchant de la première ligne une combinaison linéaire convenable des autres lignes puis de la première colonne une combinaison linéaire convenable des autres colonnes, on retrouve $G(S)$.

c) Le b) s'applique à chaque étape de l'orthogonalisation.

d) En orthogonalisant la famille T , on obtient la famille T' dont les p premiers vecteurs forment S' . Notons u' le dernier vecteur de T' . D'après le a), on a :

$$\frac{G(T)}{G(S)} = \|u'\|^2 = d(u, \text{Vect}(S))^2.$$

e) Notons A l'aire du triangle. Alors $2A$ est l'aire du parallélogramme de sommets définis par 0 , u , v et $u + v$ et $(2A)^2$ est le déterminant de Gram

$$\begin{vmatrix} \langle u, u \rangle & \langle u, v \rangle \\ \langle u, v \rangle & \langle v, v \rangle \end{vmatrix} = \begin{vmatrix} 13 & 11 \\ 11 & 17 \end{vmatrix} = 100. \text{ D'où } A = 5. \text{ Bien sûr, le calcul du détermi-}$$

nant $\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix}$ donnait immédiatement le résultat.

16.15 1,2) La seule difficulté est la longueur des vérifications.

3) La formule $\cos(n+1)a + \cos(n-1)a = 2 \cos na \cos a$ donne :

$$T_{n+1} = 2X T_n - T_{n-1}$$

On en déduit que les T_n définissent des polynômes de $\mathbb{R}[X]$.

4) Le 3) permet de montrer par récurrence que $\deg(T_n) = n$ et que le coefficient du terme de plus haut degré de T_n est 2^{n-1} . L'orthogonalité se montre à l'aide du changement de variable $u = \arccos x$.

5) a) On a $|T_n(x)| \leq 1$ pour tout x de I d'après la définition de T_n . On a $|T_n(x)| = 1$ dans I pour $n \arccos x = k\pi$ donc pour $x = \cos \frac{k\pi}{n}$, $0 \leq k \leq n$.

b) Supposons qu'il existe un polynôme $P \in F_n$ tel que $\sup_{x \in I} |P(x)| < 1$. Le polynôme $P - T_n$ change de signe entre deux valeurs successives de la suite des $n+1$ nombres $\cos \frac{k\pi}{n}$, $0 \leq k \leq n$. Il s'annule donc n fois, ce qui n'est pas possible car $\deg(P - T_n) < n$.

CARL FRIEDRICH GAUSS (1777-1855)

17

Brunswick

Les Allemands disent Braunschweig depuis le xvi^e siècle. C'est à la fois une ville et un duché. La ville est bien située, à 150 kilomètres au sud de Brême et Hambourg, et se développe à partir des xi-xii^e siècles. Elle fait partie de la Hanse et ses liens commerciaux s'étendent aux Flandres, en Angleterre, dans les pays scandinaves et jusqu'en Russie.

À l'époque moderne, c'est aussi un centre culturel important : les premiers musées allemands y sont construits en 1754 et la première du Faust (première partie) de Goethe y est donnée en 1829.

La ville est en grande partie détruite par les bombardements du 15 octobre 1944.

Le duc

Après la mort de Frédéric II, le duc de Brunswick, Charles Guillaume Ferdinand de Brunswick (1735-1806) est reconnu comme le plus grand militaire de son temps. Il est aussi réputé comme diplomate, philosophe et a des sympathies pour la Révolution française de 1789. En 1792, il se voit proposer à la fois le commandement des troupes françaises par les Girondins (on aurait même pensé à lui pour remplacer Louis XVI !) et le commandement des troupes coalisées contre les Français.

Il accepte finalement de diriger ces dernières mais, indécis, il est vaincu à Valmy par les manœuvres audacieuses de Dumouriez et bat en retraite. Les émigrés français ne sont pas du tout contents de son action ; cependant, un bon nombre sont accueillis dans son duché et, parmi eux, le futur Louis XVIII.

Les parents de Gauss

Le père de Gauss, Gebhard Gauss (1744-1808), est d'une famille de fermiers pauvres qui est venue s'établir à Brunswick vers 1740 en espérant une vie meilleure. Mais ces espoirs ne se réalisent pas et son père vivra de petits métiers : jardinier, boucher, employé aux travaux du canal... En secondes noces, en 1776, il épouse Dorothea Benze, plus âgée que lui d'un an et fille d'un tailleur de pierre mort à 30 ans de tuberculose.

Johann Karl Friedrich naît un an plus tard, le 30 avril 1777.

Le père de Gauss savait lire, écrire et compter. Mais Gauss le trouvait un peu fruste, grossier et violent. Il pensait que le mariage de ses parents n'avait pas été heureux. Ses capacités intellectuelles venaient, selon lui, du côté de sa mère.

Premières années d'études

Ce que nous savons des premières années de Gauss nous vient de Gauss lui-même, des récits qu'il aimait volontiers raconter à ses proches, de lettres qu'il a écrites. Gauss avait une excellente mémoire de ses jeunes années ; il faut sans doute le croire.

Tout petit, il tombe dans le grand canal qui passe devant la maison de ses parents sur le Wendengraben, le fossé des Slaves ; il est repêché à temps.

À 3 ans, Gauss sait compter et commence à faire des opérations élémentaires. Il est capable de corriger son père dans un calcul de salaire. Il apprend à lire et à écrire seul, questionnant les adultes autour de lui.

En 1784, il entre à l'école Sainte-Catherine de Brunswick. Il y avait sans doute beaucoup plus de 50 enfants dans la classe de Gauss et son maître, J. G. Büttner, enseignait avec un fouet à la main. Un jour, il demande à ses élèves le calcul de la somme des nombres de 1 à 100 ; le jeune Gauss écrit immédiatement 5050 sur son ardoise et va la déposer sur le bureau du maître. Quand celui-ci prend connaissance des résultats de ses élèves, il doit constater que Gauss a juste. Gauss explique calmement qu'il suffisait de grouper les nombres en 50 paquets de somme 101 : $100 + 1$, $99 + 2$, $98 + 3$, etc. Büttner se procure alors un traité d'arithmétique pour son génial élève et son père le dispense de filer du lin tous les soirs.

Gauss a la chance de rencontrer Johann Martin Bartels (1769-1836), alors très jeune, qui vient assister son instituteur. Ils se lient d'amitié et commencent à étudier ensemble les mathématiques, la formule du binôme, les séries infinies, etc. Bartels deviendra professeur de mathématiques (1808-1820) à l'Université de Kazan : cette université est créée en 1804 ; il y a quatre professeurs de mathématiques, trois sont allemands et enseignent en latin. Nikolai Lobatchevski (1792-1856) y entre en 1807 et suit des cours de Bartels. Comme Bartels a maintenu des liens avec Gauss, certains historiens se demandent si des idées de Gauss sur les géométries non euclidiennes ne sont pas arrivées à Lobatchevski par Bartels !

Gymnasium Catharineum

Peut-être convaincu par Bartels ou Büttner, son père n'empêche pas Gauss de continuer ses études malgré ses faibles moyens. Gauss entre au Gymnasium Catharineum en 1788 ; c'est une école préparatoire au collège. Ses talents en mathématiques sont

reconnus et il fait de rapides progrès dans les disciplines classiques comme le latin, considéré alors comme la base des études. Gauss apprend aussi l'allemand officiel. Jusque-là, il ne parlait que son dialecte. Chez lui, Gauss étudie la nuit dans son lit dans les combles, à la lueur d'une lanterne faite d'un navet évidé, un peu de gras et une mèche de coton.

Bartels le recommande à Eberhard Zimmermann (1743-1815), professeur au Collegium Carolinum et conseiller privé du duc de Brunswick. Zimmermann lui offre son amitié et ses encouragements. En juin 1791, il introduit Gauss à la cour et le présente au duc. Le duc, impressionné, offre à Gauss une bourse de 10 thalers par an et un de ses ministres lui donne une table de logarithmes. Ainsi, Gauss devient indépendant de son père et peut poursuivre ses études. Durant les années qui suivent, Zimmermann continue à soutenir Gauss, qui lui en sera extrêmement reconnaissant.

Collegium Carolinum

Cette institution a été fondée par le duc en 1745. Elle forme en particulier le personnel administratif et militaire du duché. D'autres établissements de la même qualité existent ailleurs en Allemagne. Gauss entre au Collegium le 18 février 1792. Il a quelques condisciples remarquables avec lesquels il entretiendra de longues relations amicales. La bibliothèque du collège est riche et Gauss lit probablement beaucoup de mathématiques : des textes de Newton, de Jean Bernoulli, peut-être aussi d'Euler ou de Lagrange.

Université de Göttingen

C'est le 11 octobre 1795 que Gauss quitte sa ville natale pour entrer à l'Université Georgia Augusta de Göttingen, fondée par le roi d'Angleterre George II en 1737, à 100 kilomètres au sud. Il semble encore hésiter entre les mathématiques et la philologie. Le duc de Brunswick l'aide toujours financièrement.

À l'époque, les études à l'Université ne comportaient aucune contrainte, aucune obligation d'assister à des cours, aucun examen régulier.

Parmi les 200 livres empruntés par Gauss à la bibliothèque, on trouve des livres de mathématiques : Lagrange, Lambert, Euler, Fermat, de nombreux volumes des académies de Paris, Berlin, etc., une grammaire de suédois et des livres dans cette langue, un certain nombre de romans et de textes classiques : Clarisse de Richardson, Gil Blas de Lesage, Jean-Jacques Rousseau, des auteurs latins, etc.

Le 31 octobre 1796, Gauss rencontre Farkas Wolfgang von Bolyai (1775-1856) qui vient de sa Transylvanie étudier la philosophie. Ils s'apprécient mutuellement. Ils se quittent le 25 mai 1799, à Clausthal, un petit village entre Brunswick et Göttingen, et ne se reverront plus. Mais leur correspondance durera toute leur vie. Elle nous permet de mieux connaître l'un et l'autre. À la mort de Gauss, Bolyai fera l'éloge de son ami pour l'Académie des sciences de Hongrie :

Il était très modeste et discret... Nous avons une passion commune pour les mathématiques et nous partageons les mêmes convictions. Souvent, nous nous

promenions ensemble, plongés dans nos pensées, sans échanger un mot pendant des heures.

Le professeur de mathématiques Abraham Kästner (1719-1800) est apprécié pour ses cours et ses livres. Il s'intéresse à la philosophie, à la logique, à l'optique, à l'astronomie et à la poésie. Durant les dernières années de sa vie, il écrit une *Histoire des mathématiques* en 4 volumes. Est-ce la vieillesse : il ne décèle pas le génie de son nouvel étudiant. De son côté, Gauss ironisera jusqu'à la fin de sa vie sur la médiocrité de son professeur : le meilleur poète parmi les mathématiciens, le meilleur mathématicien parmi les poètes, disait-il.

Gauss est plus intéressé par le professeur de physique, Georg Christoph Lichtenberg (1742-1799), par ailleurs célèbre auteur d'aphorismes, à qui l'on doit les dénominations de charges électriques positive et négative qu'il note + et –, et surtout par le philologue et humaniste réputé Christian Gottlob Heyne (1729-1812) avec lequel il correspondra jusqu'à sa mort.

À la fin de mars 1796, Gauss revient pour quelques jours à Brunswick.

Le 30 mars 1796

Ce matin-là, Gauss est encore au lit. Il réfléchit à la résolution algébrique de l'équation $z^n = 1$. Il réussit à analyser le problème avec une telle clarté, raconte-t-il, qu'il peut immédiatement calculer les racines pour $n = 17$ et qu'il découvre que le polygone régulier à 17 côtés peut être construit avec une règle et un compas.

En résulte, le 1^{er} juin, la première note mathématique de Gauss, très courte, dans le journal de l'Université :

Chaque débutant en géométrie sait qu'il est possible de construire différents polygones réguliers : triangles, pentagones, polygones à 15 côtés et les polygones réguliers en résultant, en doublant le nombre de côtés. Cela était déjà connu du temps d'Euclide et il semble qu'on pense depuis que la géométrie s'arrête là ; je ne connais aucune tentative d'étendre ces résultats. Par conséquent, il me semble intéressant qu'en plus de ces polygones, il y en ait d'autres qui soient constructibles géométriquement, comme celui à 17 côtés. Cette découverte est un corollaire d'une théorie qui n'est pas encore complète et que je publierai quand elle le sera. C. F. Gauss, Brunswick, étudiant en mathématiques à Göttingen.

Le résultat peut paraître anecdotique et on peut s'interroger sur son intérêt. Mais il est admirable, ouvrant des perspectives nouvelles en théorie des équations, théorie des groupes... (voir mon livre *Théorie de Galois*). Et cette réussite décide Gauss à se consacrer aux mathématiques.

Il faut croire Lobatchevski qui affirmait, quelques années plus tard, que toutes les branches des mathématiques, même les plus abstraites, finiraient par trouver une application dans l'étude du monde réel.

Retour à Brunswick et thèse

Le 28 septembre 1798, Gauss revient à Brunswick. Il y restera jusqu'en 1807. Il n'habite plus chez ses parents. Le duc continue à l'aider : il reçoit 158 thalers par an, mais doit soutenir une thèse.

Il y a alors un bon mathématicien à l'Université de Helmstedt : Johann Friedrich Pfaff (1765-1825). Gauss va le voir, noue une relation amicale avec lui et profite de la bibliothèque de l'université. En juin 1799, Pfaff lui fait soutenir une thèse dont le duc de Brunswick finance la publication. Gauss est reçu Docteur en philosophie de l'Université de Helmstedt le 16 juillet 1799, en son absence.

Le sujet de la thèse est la démonstration du théorème fondamental de l'algèbre que nous appelons en France théorème de d'Alembert. Ce théorème affirme (voir chapitre 13) qu'un polynôme de degré n à coefficients complexes a exactement n racines. La démonstration de Gauss a été trouvée en octobre 1797 d'après son journal (voir ci-dessous) ; elle comporte, en réalité, encore quelques difficultés. Par la suite, Gauss donnera trois autres démonstrations de ce résultat (1815, 1816, 1849).

La démonstration remplit le dernier tiers de la publication. Dans la première partie, Gauss, comme il aimera souvent à le faire, donne l'histoire et la critique des travaux de ses prédécesseurs sur le sujet : d'Alembert, Bougainville, Euler, de Foncenex, Lagrange, etc. dont les raisonnements sont imprécis ou incomplets. Toute sa vie, Gauss va souligner la nécessité de démonstrations rigoureuses et irréprochables, et son influence sur les mathématiques dans ce domaine aussi est considérable.

Disquisitiones arithmeticae

Gauss retourne travailler chez Pfaff dont il apprécie le caractère « sans aucune de ces émotions violentes qui déshonorent un homme ». Il travaille énormément et ses amis ne le voient sortir de sa chambre que quelques moments le soir.

Les efforts de Gauss aboutissent. Il peut enfin publier, le 29 septembre 1801, ses *Recherches arithmétiques*. C'est un volume de près de 500 pages dans la traduction française de A. C. M. Pouillet-Delisle parue dès 1807. Le texte original est en latin, que Gauss écrit parfaitement ; Cicéron n'y aurait quasiment rien trouvé à reprendre, d'après Moritz Cantor.

Gauss dédie son travail au duc de Brunswick qu'il remercie profondément pour l'avoir si constamment aidé ; il loue sa « libéralité à l'égard de ceux qui semblent portés vers l'étude des sciences et sa protection à celles qui paraissent les plus abstraites et d'une application moins directe aux usages ordinaires de la vie, parce que, dans la profondeur de votre sagesse... vous avez senti la liaison intime et nécessaire qui unit entre elles toutes les sciences. »

Dans sa préface, Gauss explique comment il redécouvrit des résultats déjà connus de Fermat, Euler, Lagrange ou Legendre mais dont il n'avait pas connaissance quand il commença ses recherches. Il précise que sa première redécouverte, en 1795, était que -1 est un carré modulo un nombre premier p si et seulement si p est de la forme $4k + 1$. C'est la partie la plus simple de la loi de réciprocité quadratique. Il ajoute :

ÉPITRE DÉDICATOIRE DE L'AUTEUR.

A SON ALTESSE SÉRÉNISSIME,
MONSEIGNEUR CHARLES-GUILLAUME-FERDINAND,
DUC DE BRUNSWICK ET DE LUNEBOURG.

PRINCE SÉRÉNISSIME,

LORSQUE la reconnaissance m'impose le devoir sacré de vous offrir cet Ouvrage, vous mettez le comble à ma félicité, en me permettant de placer à la tête votre nom illustre et respectable. En effet, PRINCE SÉRÉNISSIME, eussé-je pu me dévouer tout entier aux sciences mathématiques, vers lesquelles une ardeur irrésistible m'a toujours emporté, si votre faveur ne m'en eût ouvert l'entrée, si vos bienfaits continuels n'eussent incessamment soutenu mes travaux. Par vos seules bontés, libre des soins étrangers, et maître de consacrer mon temps à l'étude, j'ai pu entreprendre les recherches dont cet Ouvrage renferme une partie, et m'y livrer pendant plusieurs années. Lorsque j'ai désiré de le mettre au jour, votre munificence a écarté tous les obstacles qui en retardaient la publication. Il m'est plus facile de conserver au fond de mon cœur et d'admirer en silence,

b

Carl Friedrich GAUSS,
Recherches arithmétiques.
Reprint par les Éditions Jacques Gabay, Paris, 1989.

« Je me sentis tellement entraîné par l'attrait de ces questions, qu'il me fut impossible de les abandonner, et comme une vérité me conduisait à une autre, la plus grande partie des quatre premières sections était déjà terminée avant que j'eusse rien vu des travaux des autres géomètres sur ce sujet. »

Gauss avait promis un second volume ; la rédaction d'une partie sera retrouvée dans ses papiers, après sa mort.

La passion de l'astronomie

« Au cours des années 1802, 1803, 1804, les travaux astronomiques prirent la plus grande partie de mon temps et j'ai entrepris avant tout la théorie du calcul des planètes », écrit Gauss.

Gauss venait de publier son livre d'arithmétique. Il avait du temps. Zimmermann lui donne un numéro de la *Monatliche Correspondenz* de Zach où étaient décrites les premières observations de la nouvelle planète Cérès par Piazzi. Gauss applique sa méthode des moindres carrés pour déterminer l'orbite à partir d'un petit nombre d'observations (voir 16.8) et cela le décide à se lancer dans l'astronomie pour le reste de sa vie.

L'annonce de la redécouverte de la planète par Franck Xaver von Zach (1754-1832) le 31 décembre 1801 et par Heinrich Wilhelm Olbers (1758-1840) le lendemain, à l'endroit prévu par les calculs de Gauss, rend celui-ci célèbre dans toute l'Europe. Les Russes lui proposent la direction de l'observatoire de Saint-Pétersbourg ; Gauss attendra un an avant de refuser.

Olbers devient un grand ami de Gauss et leur correspondance va être importante. Pour retenir Gauss, le gouvernement de Brunswick pense à construire un observatoire et achète un télescope (qui posera des problèmes) ; Gauss reçoit maintenant 400 thalers par an du duc. De son côté, Olbers, qui découvre un second astéroïde, Pallas, le 28 mars 1802, sollicite Göttingen pour penser à Gauss afin de diriger leur futur observatoire : « c'est le premier mathématicien de son temps », écrit-il, vantant les connaissances et les capacités intellectuelles exceptionnelles de son ami. À la demande d'Olbers, Gauss calcule l'orbite de Pallas. Il lui faut trois heures, le 18 avril.

Par l'intermédiaire d'Olbers, Gauss entre en contact avec Friedrich Bessel (1784-1846), astronome et mathématicien. Leur correspondance, poursuivie pendant plus de 40 ans, est très intéressante.

La correspondance avec Sophie Germain

Sophie Germain est née en 1776 dans une famille bourgeoise. À 13 ans, elle se passionne pour les mathématiques, lisant tout ce qu'elle trouve, travaillant la nuit, son encre gelant dans l'encrier. C'est la première femme française à faire ainsi des mathématiques à une époque où le domaine scientifique est réservé aux hommes. À 19 ans, elle utilise un pseudonyme masculin : Leblanc, pour prendre contact avec Lagrange qu'elle finit par rencontrer. Elle se passionne pour *La théorie des nombres* publiée par Legendre en 1789 et les *Disquisitiones arithmeticae*. Elle écrit à Gauss ses propres découvertes en novembre 1804, reprenant le pseudonyme de Leblanc. La réponse de Gauss est très chaleureuse : « Monsieur, je me félicite que

l'Arithmétique acquiert en Vous un ami aussi habile... votre nouvelle démonstration... m'a extrêmement plu... » D'autres lettres suivent.

En 1806, Sophie Germain s'inquiète pour Gauss des menaces des troupes françaises et le recommande au général Pernety. Celui-ci prend contact avec Gauss, en novembre 1806, en se recommandant de Sophie Germain. Gauss ne comprend cette intervention que lorsque Sophie Germain lui révèle sa véritable identité, quelques temps plus tard. Il lui répond le 30 avril 1807 : « Lorsqu'une personne de ce sexe qui, par nos mœurs et nos préjugés, doit rencontrer infiniment plus d'obstacles et de difficultés à se familiariser avec ces recherches..., il faut sans doute qu'elle ait le plus noble courage, des talents tout à fait extraordinaires, le génie supérieur... » La correspondance entre Gauss et Sophie Germain cesse après 1808. Sophie Germain n'aura jamais le moindre poste universitaire.

Le journal intime de Gauss

En 1898, 43 ans après la mort de Gauss, on découvrit un petit cahier de 19 pages gardé jusque-là dans les papiers de famille. Gauss y avait noté, sans démonstration, 146 résultats dont il souhaitait probablement se souvenir. La dernière note est du 9 juillet 1814, la première du 30 mars 1796 :

Principe de la division du cercle et comment le diviser géométriquement en 17 parties, etc.

Les premières notes montrent l'incroyable fécondité de Gauss durant les années 1796-1800. Le plaisir de la découverte y éclate. Semaine après semaine, on trouve mention de la loi de réciprocité quadratique, du mouvement des planètes, des fondations de la géométrie, de la date de Pâques, etc. et aussi de résultats bien connus que Gauss devait découvrir pour la première fois.

Le journal prouve que ceux qui avaient pu douter de Gauss quand il affirmait avoir déjà résolu tel ou tel problème dans sa jeunesse avaient tort. S'il n'avait rien publié, c'est parce que ce n'était pas encore dans sa tête sous une forme qui lui plut.

Le journal est publié par Klein en 1903 dans les *Mathematische Annalen*.

Johanna Osthoff

Johanna Osthoff est née le 5 août 1780. La mère de Gauss connaissait sa famille, modeste. On n'a conservé aucun portrait de Johanna, mais tous les témoignages s'accordent sur son charme et sa gaieté.

Gauss la rencontre le 27 juillet 1803. Il est blond, a les yeux très bleus, un peu myopes, mesure 1 mètre 60 et a l'aspect un peu trapu. Il est séduit, elle aussi probablement : c'est elle qui se souvient de la date de leur première rencontre. Mais il doit voyager et part pour six mois.

Le 28 juin 1804, il écrit à Bolyai. Il parle de mathématiques, de vins hongrois, des nouvelles amitiés qu'il a noué au cours des cinq mois précédents à Brême et à Gotha. « La plus belle, cependant, est l'amitié d'une fille merveilleuse, exactement celle avec qui j'ai toujours désiré partager ma vie. Elle a d'abord un merveilleux visage de madonne... des yeux tendres avec quelque chose de romanesque, une tour-

nure irréprochable... elle a encore une âme d'ange, calme, gaie, modeste, chaste, qui ne pourrait faire de mal à aucun être et c'est le meilleur. »

Gauss hésite à se déclarer, voulant être sûr de pouvoir la rendre heureuse comme elle le mérite. Le 12 juillet, il lui écrit une déclaration superbe, grave et profonde. Il attend sa décision. Il fait trois petits voyages pour voir Olbers qui compteront parmi ses plus beaux souvenirs de jeunesse.

Johanna reste longtemps sans répondre. Finalement, elle accepte. Et le 25 novembre de la même année, Gauss peut écrire à Bolyai, après lui avoir tout de même d'abord parlé d'astronomie : « Depuis trois jours cet ange, presque trop céleste pour notre terre, est ma fiancée. Je suis extraordinairement heureux... » et depuis qu'il l'a serré dans ses bras, « La vie m'apparaît comme un printemps éternel aux couleurs étincelantes ». Il n'avait jamais espéré un tel bonheur et il ajoute, citant en français *La nouvelle Héloïse* de Jean-Jacques Rousseau : « Puissances du Ciel : j'avais une âme pour la douleur, donnez m'en une pour la félicité. »

Ils se marient le 9 octobre 1805.

La mort du duc

Après la victoire d'Austerlitz, le 2 décembre 1805, où il écrase les Autrichiens, Napoléon rêve de restaurer l'Empire d'occident. Une nouvelle coalition se met en place contre lui sous l'impulsion de l'Angleterre avec la Russie et la Prusse qui organisent trois armées. Le duc de Brunswick doit prendre, à 71 ans, le commandement de l'une d'elles. Elle est mal équipée, mal entraînée, en infériorité numérique. Elle est battue à Auerstedt par Davout, un peu à l'est de Brunswick, le 14 octobre 1806. Le duc est terriblement blessé aux yeux et doit fuir. Gauss voit passer sa voiture sous ses fenêtres le 25 octobre au matin. Le duc est emmené à Altona (où se déroule une pièce de Sartre ; c'est une ville indépendante de Hambourg jusqu'en 1937) où il meurt le 10 novembre. Ainsi Gauss perd-il celui qui l'avait constamment soutenu depuis 15 ans et auquel il était si fidèlement attaché. Il l'avait revu une dernière fois en mai, pour le remercier de la dernière augmentation de sa subvention.

Ce 14 octobre est aussi le jour de la défaite de l'autre armée prussienne par Napoléon à Iéna. Après ces défaites, la Prusse s'effondre. Napoléon continue sa marche sur Berlin, mais est arrêté par les Russes à Eylau, le 8 février 1807. Quelques semaines plus tard, il reprend l'offensive et la victoire de Friedland contraint les Russes au traité de Tilsit (8 juillet 1807).

La Prusse est alors réduite aux provinces à l'est de l'Elbe et le pays conquis forme le royaume de Westphalie, confié à Jérôme Bonaparte et soumis à une énorme contribution de 160 millions de francs.

La nomination à Göttingen

Depuis qu'il est revenu à Brunswick, Gauss n'a pas réussi à occuper une position stable mais sa notoriété est établie : il est nommé correspondant de l'Institut de France le 30 janvier 1804 et membre de la Royal Society de Londres le 12 avril 1804. Alexandre de Humboldt le recommande à l'empereur Frederick Wilhem :

« Le seul homme qui puisse donner une nouvelle splendeur à l'Académie de Berlin est Gauss » (il y sera nommé à l'unanimité quelques années plus tard, le 18 avril 1810). Après la mort du duc, rester à Brunswick ne se justifie plus. Quand l'Université de Göttingen lui propose un poste, le 9 juillet 1807, Gauss en délibère calmement avec Olbers et accepte. Il est nommé professeur d'astronomie et directeur de l'observatoire de Göttingen le 25 juillet 1807. Son assistant est Ludwig Harding (1765-1834), qui a découvert l'astéroïde Junon (1^{er} septembre 1804).

Gauss déménage à Göttingen le 21 novembre 1807 ; il devait y rester toute sa vie.

Pendant les 50 années qui suivent, Gauss sera souvent très pris par les tâches administratives et devra s'organiser pour continuer ses activités de recherche.

Le bonheur du couple

Des enfants naissent : Joseph, le 21 août 1806, et Wilhelmina surnommée Minna, le 29 février 1808 (Gauss s'amusait de ce qu'elle n'ait qu'un anniversaire tous les quatre ans). Leurs prénoms viennent de ceux des découvreurs des deux premiers astéroïdes, Piazzi et Olbers.

Le grand bonheur simple continue.

En visite chez Olbers en juillet 1807, Gauss écrit : « Ma chère Jeannette (Hannchen), quand j'ai un peu de temps libre, je ne connais pas de plus grand plaisir que de t'écrire pour bavarder un peu avec toi, même si je n'ai rien d'important à te dire... » Il ajoute aussi, commentant l'armistice entre Français et Russes qui allait aboutir au traité de Tilsit : *Es sind tolle Zeiten* (*Ce sont des temps insensés*).

Le 2 septembre 1808, une lettre à Bolyai nous laisse la même impression : « Heureux s'écoulent les jours (Glücklich fließen die Tage) dans la marche uniforme de notre vie domestique. Quand la fille a une nouvelle dent ou quand le garçon apprend de nouveaux mots, c'est presque aussi important que la découverte d'une nouvelle étoile ou d'une nouvelle vérité... »

Les lettres de Johanna donnent la même impression ; elle profite de la vie sociale de Göttingen, s'occupe des enfants, se fait des amies ; mais elle regrette un peu sa famille et Brunswick.

La mort de Johanna

Un troisième enfant naît le 10 septembre 1809, Louis ; le prénom vient de celui de Harding. Ce troisième accouchement est encore plus difficile que le second. Johanna meurt le 11 octobre à 20 heures. Gauss lui ferme les yeux pensant à son bonheur des dernières années. Des textes de Gauss retrouvés plus tard montrent l'émotion qui a été la sienne. Gauss est recueilli quelques jours par Olbers. Il revient à Göttingen le 29 octobre. Louis, *le pauvre Louis*, disait son père, meurt subitement le 1^{er} mars 1810.

L'astronomie

Pendant toute cette période, Gauss continue ses calculs et ses observations astronomiques. L'observatoire de Göttingen est situé dans une vieille tour d'époque médiévale. Gauss y fait ses observations, attendant la construction du nouvel observatoire

que le gouvernement de Westphalie finance à partir de 1810. Il s'occupe en particulier des orbites des deux petites planètes découvertes par Olbers, Pallas et Vesta (le nom est choisi par lui). Il publie, en juin 1809, son second livre, lui aussi considéré comme un chef-d'œuvre, sur la théorie du mouvement des corps célestes, en latin à la demande de son éditeur : *Theoria motus corporum cœlestium...* Y sont exposées des méthodes pour déterminer l'orbite de corps célestes à partir de quelques observations. C'est l'occasion pour Gauss d'exposer sa méthode des moindres carrés (voir 16.8) et d'introduire la fonction définissant la fameuse courbe en cloche que tout le monde connaît en probabilités. Il indique le problème de priorité revendiqué par Legendre, affirme son antériorité, mais ne polémique pas, ce qui serait contraire à son caractère.

L'admiration des grands mathématiciens français et le goût de Napoléon pour les mathématiques auraient permis à Gauss d'occuper une position importante. Mais Gauss refuse toutes les compromissions avec le nouvel Empire. Comme il n'accepte pas l'argent d'un prix décerné par l'Institut de France, Sophie Germain lui fait envoyer une pendule qu'il utilisera toute sa vie.

La vie reprend

Quelques mois après la mort de Johanna, Gauss se tourne vers une amie de son épouse : Friederica Wilhelmine Waldeck, dite Minna (comme sa fille). Il lui écrit en mars et en avril 1810, de longues lettres très sérieuses, lui raconte son passé, son présent, son désir d'un nouveau bonheur, partagé avec elle. Il parle de ses enfants, il lui offre un cœur divisé où le souvenir de sa première femme ne pourra être effacé.

Minna Waldeck est née le 15 avril 1788 ; elle est issue d'un milieu aisé ; son père est professeur de droit et conseiller privé. Ils se marient le 4 août 1810. Des enfants naissent : Eugen, le 29 juillet 1811, Wilhelm, le 23 octobre 1813 et Therese, le 9 juin 1816.

Mais Gauss ne retrouvera jamais l'exubérance des années passées avec Johanna.

Les années 1810-1818

Parmi les nombreuses publications de Gauss datant de cette période, citons deux articles importants : un, de 1812, sur la série hypergéométrique, où le problème de convergence de la série est soigneusement étudié, ce qui est une nouveauté, et un autre, sur le calcul approché d'intégrales (voir 16.13.3).

Les observations astronomiques de Gauss continuent tous les soirs où le temps le permet et sa fille Minna surveille le ciel pour voir si son père va pouvoir rester avec elle ou non. La construction du nouvel observatoire est achevée en octobre 1816 et Gauss y emménage avec sa famille. Les instruments sont installés les uns après les autres de 1814 à 1821.

La carte du Hanovre

Heinrich Christian Schumacher (1780-1850) a été un des premiers étudiants de Gauss à Göttingen. Il s'occupe de la carte du Danemark et sera professeur d'astronomie à Hambourg.

Il demande à Gauss (8 juin 1816) de travailler à la cartographie du Hanovre, commencée et laissée inachevée à l'époque napoléonienne. Le Hanovre dépend maintenant du royaume d'Angleterre. L'idée intéresse Gauss qui va passer plusieurs étés de sa vie à parcourir le pays, de village en village, dormant chez l'habitant, s'irritant des difficultés de toutes sortes et ne retournant que rarement chez lui. Les subventions anglaises lui ont permis de recruter des militaires, mieux respectés des paysans, écrit-il ; il se fait également aider par son fils aîné, Joseph, qui prend goût à ce travail.

Une difficulté est la nature du pays : de grandes forêts sans relief empêchent les visées ; il faut parfois abattre des arbres jusqu'à ce qu'on puisse apercevoir les repères. Gauss utilise l'*héliotrope*, instrument qu'il invente en 1821, ayant remarqué trois ans auparavant la gêne que créait la lumière réfléchie par une fenêtre lointaine ; l'héliotrope, conçu avec un miroir mobile de 10 à 20 centimètres, permet d'effectuer des triangulations sur des distances supérieures à celles de ses prédécesseurs ; il figurait sur le billet de 10 marks consacré à Gauss ; il permet également d'envoyer des signaux et sert de télégraphe optique. Gauss envisage même des signaux vers la lune avec une centaine d'appareils réunis. Après 1825, il continue à superviser les opérations de triangulation, lesquelles dureront jusqu'en 1847.

Les deux rapports publiés par Gauss en 1828 et 1843-1846 auront une grande influence sur les méthodes de géodésie de son époque. Elles rendent compte d'un travail énorme. Les méthodes de Gauss seront utilisées jusqu'à l'apparition de la photographie aérienne.

La courbure des surfaces

Sans cesser ses autres travaux et obligations, Gauss se plonge, à l'automne 1825, dans des recherches théoriques sur les surfaces et leur courbure. Il ne dort pratiquement plus et c'est l'une des périodes les plus épuisantes de sa vie.

L'article sur la courbure des surfaces qui paraît en 1827, *Disquisitiones generales circa superficies curvas*, aura une influence considérable sur les travaux de géométrie différentielle ultérieurs. Einstein dira que si Gauss n'avait pas créé sa géométrie des surfaces, base des travaux de Riemann et de la théorie de la relativité, il était difficile d'imaginer que quelqu'un d'autre l'aurait fait à sa place.

La santé de Minna

En épousant Minna, Gauss est entré dans un monde nouveau pour lui. Il s'y adapte vite, devient homme du monde, gère son argent, fait des placements...

Au début de leur mariage, la vie de famille s'écoule calmement pour Gauss et Minna qui prennent le café ensemble tous les jours. Une grande partie des lettres et des souvenirs de cette période seront envoyés, en 1848, aux enfants partis aux États-Unis ; le bateau sombrera et ils seront perdus. Après leur emménagement dans le nouvel observatoire, la mère de Gauss vient habiter chez eux.

La santé de Minna est fragile. Le mal de cette époque était la tuberculose et il est probable qu'elle minait Minna. Elle va régulièrement aux bains dans les stations thermales, mais son état se détériore peu à peu. Elle est constamment malade en

1823 et 1824 et, quand elle attrappe la rougeole, elle reste quelques temps entre la vie et la mort.

Entre 1821 et 1826, Gauss est fortement sollicité pour venir à l'Université de Berlin. Minna aurait été heureuse de ce changement, la situation de Gauss aurait pu être avantageuse scientifiquement et financièrement et son influence en Prusse énorme. Les négociations, où Alexandre de Humboldt joue un grand rôle, sont longues. Gauss communique la proposition finale de Berlin aux autorités de Hanovre ; celles-ci augmentent son salaire et Gauss décide alors de rester à Göttingen où il a ses amis et sa tranquillité. Beaucoup de ses proches ne le comprennent pas.

L'été 1825, Gauss et Minna passent deux mois dans la station thermale de Baden Baden. La santé de Minna n'en est pas améliorée et Gauss lui-même souffre fortement de la chaleur, comme tous les étés.

Les années suivantes, la santé de Minna se dégrade de plus en plus.

Le dernier voyage de Gauss est une visite à Alexandre de Humboldt l'été 1828, à l'occasion d'un congrès à Berlin. Gauss reste trois semaines et fait la connaissance du physicien Wilhelm Weber (1804-1891).

La fin de Minna

La santé de Minna est devenue critique et les relations de Gauss avec ses fils ne sont pas excellentes. En 1830, Eugen, qui a alors 19 ans, quitte la maison familiale pour Brême. Gauss va le voir, tente de discuter avec lui, sans résultat. Minna est restée à Göttingen. Elle écrit à Gauss une longue lettre désespérée sur son état, son impossibilité à bien tenir la maison, sur les problèmes avec les enfants, regrettant l'intransigeance de Gauss, le suppliant de changer d'attitude : *Carl, Carl, ne rejette pas mes*



CARL FRIEDRICH GAUSS, à environ 55 ans

Croquis de J.B. LISTING, Bibliothèque de l'université de Göttingen.

prières. Le départ d'Eugen, le 13 octobre 1830, aux États-Unis est très douloureux pour elle et Gauss le reprochera longtemps à son fils. Elle meurt le 12 septembre 1831.

Quatre jours plus tard, Gauss écrit à Olbers. Il raconte les souffrances de Minna, sa douleur, le soutien de ses deux filles (elles ont alors 15 et 23 ans).

Pendant 24 ans, la plus jeune, Therese va s'occuper des tâches matérielles dans la maison de son père.

L'ainée, Minna, était la favorite de Gauss. Il disait qu'elle ressemblait à sa mère. Elle se marie en 1830 avec le professeur de langues orientales et de théologie de Göttingen, Heinrich Ewald (1803-1875), le premier à enseigner le sanscrit à Göttingen. C'est un mariage heureux. Mais elle commence déjà, elle aussi, à souffrir de tuberculose.

Travaux en physique

Weber est nommé professeur de physique à Göttingen en 1831, sur la recommandation de Gauss. Il arrive trois jours après la mort de Minna. Une forte amitié lie les deux hommes qui se complètent parfaitement : Gauss est plus théoricien et Weber plus pratique. Ils sont sans cesse ensemble, dînant l'un chez l'autre, causant à l'infini. En 1832, Gauss écrit un texte fondateur pour les systèmes d'unités, proposant de prendre le millimètre, le milligramme et la seconde comme unités fondamentales. Ce texte influencera le congrès international de Paris de 1881 qui choisira le système CGS, basé sur le centimètre, le gramme et la seconde. Avec Weber, Gauss construit le télégraphe électromagnétique en 1833, en reliant l'observatoire de Göttingen et un autre bâtiment de l'université ; même si son télégraphe n'a qu'une portée de quelques mètres, Gauss est le premier dans ce domaine qui allait rapidement devenir essentiel pour les communications dans le monde¹. Il fonde avec Weber en 1837 la *Magnetische Verein* qui va recueillir de nombreuses mesures du champ magnétique terrestre. Puis il étudie avec son énergie habituelle le magnétisme terrestre pendant plusieurs années ; il en résulte la publication d'un article théorique en 1839 et, avec Weber, d'un atlas, en 1840. Dans ces recherches, Gauss découvre plusieurs lois fondamentales du magnétisme et de l'électricité qui seront redécouvertes par d'autres dans les années suivantes. L'unité de mesure d'induction magnétique est aujourd'hui appelée le Gauss.

De la même époque, il ne faut pas oublier les travaux de Gauss en optique. Ils rentrent plus directement dans la définition de sa chaire ; dans ce domaine aussi, les recherches de Gauss, ses méthodes géométriques, son exposition élégante, forment une étape importante.

Gauss et la politique

Gauss avait été reconnaissant au duc de Brunswick de l'avoir aidé. Il a toujours gardé des convictions politiques conservatrices et une attitude de non intervention qu'un épisode illustre nettement.

1. Le télégraphe de Samuel Morse (1791-1872) est expérimenté en 1837 ; le développement du télégraphe dans les années 1850-1900 est très rapide ; il y a plus de 350 000 km de câbles en 1900.

En 1837, le roi Guillaume IV d'Angleterre meurt. C'est la reine Victoria qui monte sur le trône, mais en Angleterre seulement, car le Hanovre applique la loi salique : les femmes sont écartées de la succession. Le Hanovre se détache alors de la couronne d'Angleterre.

Le nouveau roi Ernst-August annule la constitution libérale de 1833, probablement parce qu'elle interdisait le trône à toute personne ayant des déficiences physiques ou morales et que le fils du nouveau roi était aveugle.

La contestation s'installe dans son royaume et sept professeurs de l'Université écrivent au roi. Ils sont immédiatement exclus et le roi déclare qu'il est aussi facile de recruter des professeurs que des danseuses... Gauss cherche à agir en faveur de Weber, demande à Humboldt d'intervenir, en vain. Il ne lui semble pas possible d'intervenir en faveur de son gendre, bien que cela implique le départ de sa fille Minna. Weber et Ewald devront trouver des postes ailleurs. Ils reviendront dix ans plus tard.

Les géométries non euclidiennes

Jusqu'au XIX^e siècle, la seule géométrie est celle d'Euclide avec son fameux cinquième postulat dont un énoncé est : *Par un point extérieur à une droite, on peut mener une et une seule parallèle à cette droite*. Depuis longtemps, on en a cherché une démonstration à partir des autres axiomes, en vain. Au XVIII^e siècle, Giovanni Saccheri (1667-1733) puis Lambert (1728-1777), en en cherchant des démonstrations par l'absurde, ont commencé à montrer des résultats d'une géométrie où ce postulat ne serait pas vrai.

C'est Lobatchevski qui publie le premier exposé approfondi de géométrie non euclidienne, qu'il appelle imaginaire, dans le *Courrier de Kazan*, en 1829. Nous avons cité l'hypothèse d'une influence lointaine de Gauss via Bartels.

En 1831, Janos Bolyai (1802-1860) écrit à son père Wolfgang qu'il vient de faire des découvertes extraordinaires. Il s'agit des résultats de Lobatchevski, obtenus indépendamment. Le père écrit alors à Gauss. La réponse de Gauss est décevante pour les Bolyai : « Louer ce travail serait me louer moi-même ». Gauss ajoute que le travail de Janos est en accord avec ce qu'il a trouvé depuis 30 à 35 ans, qu'il ne pensait rien publier sur ce sujet de son vivant, mais qu'il pensait rédiger quelque chose pour que ce ne soit pas perdu après lui, ce qui lui sera évité par le travail remarquable de Janos. Cette réponse plongera Janos dans l'amertume.

Fin 1837, Gauss commence l'étude du russe qu'il réussira à parler et écrire parfaitement. Il peut alors lire Lobatchevski dans le texte et, en 1842, le faire nommer correspondant de la société scientifique de Göttingen.

Sur ce sujet difficile, Gauss ne publiera pratiquement rien, mais il lui donnera le nom de géométrie non euclidienne. S'il ne publie rien, ce n'est pas faute d'y penser : depuis 30 ans, affirme-t-il à Taurinus quand celui-ci publie quelques travaux sur le sujet. Il précise à Schumacher en 1846 : « Vous savez que je partage les mêmes convictions depuis 1792, sans parler de certains développements qu'ont reçu depuis mes idées sur le sujet. Je n'ai donc trouvé dans l'ouvrage de Lobatchevski aucun fait nouveau pour moi ; mais l'exposition est toute différente de celle que j'avais projeté

et l'auteur a traité la matière de main de maître et avec le véritable esprit géométrique. »

La soixantaine

Gauss vieillit, mais sans ennuis de santé majeurs ; il est toujours indisposé par la chaleur de l'été. Il ne quitte plus son domicile une seule nuit, sauf en 1854 pour prendre le nouveau chemin de fer.

À la fin des années 1830, des êtres chers disparaissent. Sa mère meurt aveugle le 19 avril 1839. Sa fille Minna, qui s'était beaucoup occupée de sa grand-mère, a dû partir pour Tübingen en 1838 avec son mari, comme nous l'avons déjà dit. Sa tuberculose l'affaiblit de plus en plus. Elle meurt à Tübingen le 12 août 1840, à 33 ans ; Gauss est terriblement touché et l'écrit à son mari. Ses amis meurent : Olbers en 1840, Bessel en 1846, Schumacher et Bernhard Lindenau en 1850. Gauss se sent seul. Weber doit aussi partir entre 1843 et 1849 ; il est devenu comme un fils pour Gauss.

Gauss continue ses lectures de toutes sortes. Son auteur allemand préféré est Jean-Paul (Richter) (1763-1825) ; Goethe ne semble pas l'avoir intéressé et Schiller lui déplaisait. Il adore les romans d'aventure de Walter Scott ou de Fenimore Cooper. Il a lu tous les grands classiques français et anglais. Il se débrouille moins bien en suédois, italien et espagnol. L'étude du russe lui permet d'aimer Pouchkine. Il aborde un moment l'étude du sanscrit.

Les événements de 1848 lui donnent l'occasion d'exprimer des opinions très conservatrices. Il souhaite un monarque fort et intelligent, le calme et la paix et n'a qu'une médiocre opinion de l'intelligence du peuple.

Dans les années 1840, les honneurs et les hommages se succèdent. On fête son jubilé le 16 juillet 1849. Il y a des discours, un grand banquet, des fleurs à profusion, des gens venus de partout le fêter et lui témoigner admiration et affection. Jacobi est assis à sa droite ; il veut aborder des sujets mathématiques ; Gauss élude et Jacobi le regrette.

Le dernier cours

Le dernier cours de Gauss, durant l'hiver 1850, est suivi par Richard Dedekind (1831-1916), un des grands mathématiciens de la seconde moitié du siècle, qui collaborera plus tard à l'édition des œuvres de Gauss. Les cours de Gauss depuis 1808 portaient souvent sur des sujets d'astronomie. Les dernières années, Gauss reprend toujours le même thème : la méthode des moindres carrés. Sur les neuf auditeurs du cours de 1851 (ce n'était pas un mauvais chiffre, certains cours de Gauss n'ont eu que trois auditeurs), sept paieront 25 thalers (des frais de scolarité, en quelque sorte) à Gauss. Dedekind avait conservé un merveilleux souvenir de ce cours. Il a raconté, en 1901, que Gauss prenait alors plaisir à enseigner, trois heures par semaine, assis les mains sur les genoux pour donner des explications lumineuses ou debout très droit devant le tableau pour écrire d'une très belle écriture les formules ; il notait les exemples numériques sur de petites feuilles de papier.

Les dernières années

En 1851, Gauss publie son dernier livre, un traité sur les fonds de pension. Le problème initial était celui du fond de pension pour les veuves des professeurs de Göttingen, mais, comme d'habitude, Gauss traite un problème beaucoup plus général, montrant sa connaissance approfondie des mécanismes économiques et des données statistiques.

Sa dernière observation astronomique est celle d'une éclipse de soleil, le 28 juillet 1851.

L'été 1853, le cœur de Gauss donne des signes de fatigue.

Le 7 décembre 1853, Gauss félicite Humboldt d'avoir atteint l'âge de Newton : 30766 jours.

Il écrit à Bolyai qu'on peut lui envier tout ce qu'il a fait mais que, pour lui, cela le rend conscient du vide de la vie et lui offre la meilleure garantie qu'une plus belle métamorphose doit suivre la mort.

La leçon d'habilitation de Riemann, le 10 juin 1854, est mémorable. L'usage était alors de présenter trois sujets et d'exposer le premier. Mais Gauss demande à Riemann, à la surprise de celui-ci, d'exposer le second sujet. Ce texte prophétique de Riemann fonde la géométrie dite riemanienne. Gauss en sort enthousiasmé.

Fin 1854, Rudolf Wagner (1805-1864), professeur à Göttingen, vient le voir et prend des notes de ses conversations qu'on retrouvera beaucoup plus tard. Il est spécialiste de physiologie. Gauss est convaincu que dans ce domaine comme dans d'autres, une partie peut être exprimée mathématiquement. Il dit à Wagner que la vie sans l'immortalité lui apparaîtrait comme une absurdité et qu'il envie ceux qui peuvent croire simplement du fond de leur cœur.

Gauss perd des forces. Ses dernières lettres datent de janvier 1855. Son cœur faiblit encore et, après plusieurs alertes, il meurt paisiblement dans son sommeil le 23 février 1855 à 1 heure 5 du matin.

Son cerveau est prélevé. Ses circonvolutions et son poids semblent remarquables. Il est conservé à Göttingen.

À ses obsèques, il y a beaucoup de monde ; son gendre Ewald et son ami Sartorius (1809-1876) font son éloge avec émotion.

Que sont-ils devenus ?

Les enfants de Gauss ont connu des destins très divers. Nous avons dit que les relations de Gauss avec ses fils ont été parfois difficiles.

Joseph, après avoir aidé son père pour la cartographie du Hanovre, devient militaire. En 1836, il voyage aux États-Unis pour étudier les techniques de construction de chemins de fer. Il se marie en 1840. Il quitte la vie militaire, où il n'a pas d'avenir, en 1846, pour les chemins de fer du Hanovre dont il devient ingénieur en chef. En 1849, il donnera à Gauss le seul petit-fils qu'il ait connu.

Nous avons déjà parlé de la fille aînée de Gauss, Minna.

Eugen semble avoir hérité de son père une partie de ses qualités. Il était très doué pour les langues, parlait parfaitement le français, mais Gauss s'opposa à la carrière

scientifique souhaitée par son fils, le poussant vers le droit. L'incompréhension du père ne put que pousser le fils à faire des bêtises : dettes de jeux,... Nous avons déjà raconté son départ aux États-Unis. Gauss écrit l'année suivante à ce *bon à rien* une lettre de morale. Puis Eugen évolue. Il s'engage dans l'armée, devient presbytérien, entre dans une compagnie qui fait le commerce des fourrures dans le Missouri, apprend le sioux et s'occupe de la transcription de la Bible dans cette langue. En 1844, il se marie et Gauss lui envoie ses souhaits de bonheur. La progression d'Eugen est alors remarquable. Il fonde la First National Bank dont il est le premier directeur et devient très aisé. Il a sept enfants. Les dernières années de sa vie, il était encore capable de performances de calcul mental étonnantes. Il a brûlé beaucoup de lettres de son père. Il meurt en 1896.

Au début des années 1830, Wilhelm a désiré sans beaucoup de succès s'établir comme fermier. Il épouse en 1837 une nièce de Bessel avec laquelle il émigre aux États-Unis ; Gauss apprend que leur bateau a dû affronter une tempête ; il reste plusieurs mois sans nouvelles, très inquiet. Il est enfin heureux d'apprendre la naissance de son premier petit-fils, Charles Frederick, son prénom, en 1838. Wilhelm échoue une nouvelle fois à devenir agriculteur, mais il connaît la réussite en créant une entreprise de vente en gros de chaussures à Saint-Louis. Il meurt millionnaire en 1879. Il a eu huit enfants.

Therese, si proche de son père durant tant d'années, est terriblement touchée de sa disparition. Elle ne peut d'abord quitter la maison où tout lui rappelle son père, puis part quelques semaines en Suisse. Ses frères l'invitent aux États-Unis. Elle se marie enfin en septembre 1856 avec un homme de théâtre ; elle meurt sans enfant le 11 février 1864.

Les descendants de Gauss aux États-Unis sont maintenant extrêmement nombreux.

Vers l'immortalité

L'immortalisation de Gauss commence par la publication d'une biographie rédigée par Sartorius en 1856, qui va servir de base aux études sur Gauss pour un siècle.

Gauss a entretenu une correspondance suivie avec de nombreux scientifiques de son époque. On connaît environ 7000 lettres de lui. En dehors des nouvelles sur les états de santé des uns et des autres, elles sont intéressantes par les discussions scientifiques qu'elles permettent de suivre, les points de vue de Gauss sur un grand nombre de recherches de son époque et les informations sur ses propres travaux. La correspondance avec Schumacher est publiée entre 1860 et 1865, 5 volumes pour 596 lettres de Gauss et 808 de Schumacher. Par la suite, on publiera les correspondances de Gauss avec Bessel en 1880 (76 lettres de Gauss et 119 de Bessel), Olbers en 1900 (329 lettres de Gauss et 383 de Olbers), avec sa famille en 1927...

Les œuvres de Gauss sont éditées lentement : volumes 1 et 2 (1863), 3 (1866), 4 (1873), 5 (1867), 6 (1874), 7 (1871), 8 (1900), 9 (1903), 10 (1917), 11 (1927), 12 (1929). Mais les principaux textes sont réédités, traduits, à de nombreuses reprises.

La biographie la plus complète de Gauss est écrite par Waldo Dunnington en 1955 pour le centenaire de sa mort. Elle m'a servi de base pour ce chapitre avec, entre autres, des livres d'Uta Merzbach, Walter Bühler et Pierre Gabriel.

Troisième année

Il existe un certain consensus pour les sujets d'algèbre enseignés durant les deux premières années de licence dans les Universités et dans les classes préparatoires. Pour la troisième année de licence, d'une université à une autre, les choix ne sont pas uniformes. Nous ne pouvons couvrir tous les sujets traités en France et avons dû faire des choix, ouvrant des fenêtres sur certaines notions, en ignorant d'autres, cherchant à donner un aperçu d'utilisations actuelles tout en racontant un peu d'histoire de temps à autre.

OUVERTURES SUR LES GROUPES

18

Au chapitre 11, nous avons introduit les notions de groupe, de sous-groupe, d'homomorphisme et d'isomorphisme de groupes. Nous avons étudié le groupe des permutations d'un ensemble fini, le théorème de Lagrange.

Cela donne une petite idée de la théorie des groupes pour les deux premières années d'université, mais c'est bien peu pour une théorie immense. Ce chapitre a pour but d'étudier des notions de base que nous n'avons pas encore abordées (sous-groupes distingués et groupes quotients, groupes cycliques, produits de groupes, action d'un groupe sur un ensemble) et de présenter quelques exemples importants. La structure des groupes abéliens de type fini, la notion de produit semi-direct, les théorèmes de Sylow et de Jordan-Hölder ne sont pas traités.

Structure de groupe. Deux groupes sont dits avoir la même structure s'ils sont isomorphes. Deux groupes n'ont pas la même structure quand une propriété de la théorie des groupes conservée par isomorphisme est vraie dans l'un des groupes et pas dans l'autre, par exemple la commutativité ou, pour les groupes finis, les propriétés d'avoir le même nombre d'éléments, le même nombre d'éléments de tel ordre, le même nombre de sous-groupes de tel ordre, etc.

Pour les petites valeurs de n , la liste complète des structures de groupe à n éléments a été établie. Si G est un groupe d'ordre n , il n'est pas toujours facile de montrer l'existence d'un isomorphisme du groupe G avec un des groupes, disons Γ , de la liste ; une méthode indirecte est de montrer que le groupe G n'est isomorphe à aucun groupe de la liste autre que Γ .

On dit, par abus de langage, des phrases comme : *il existe deux groupes à 6 éléments*, ce qui signifie qu'il existe deux groupes à 6 éléments non isomorphes définissant donc deux structures différentes.

18.1 RELATION D'ÉQUIVALENCE SUR UN ENSEMBLE

Pour expliquer la notion de quotient, nous allons rappeler et reprendre des notions de base de la théorie des ensembles : partition d'un ensemble, relation d'équivalence, ensemble quotient.

Partition d'un ensemble. Une partition d'un ensemble E est une famille $(E_i)_{i \in I}$ de sous-ensembles (on dit aussi de parties) de E , non vides et deux à deux disjoints, dont l'union est E , autrement dit :

- pour tout $i \in I$, $E_i \neq \emptyset$;
- pour tout $i, j \in I$, $E_i \cap E_j = \emptyset$ si $i \neq j$;
- $\cup_{i \in I} E_i = E$.

On définit une application surjective $\pi : E \rightarrow I$ en posant $\pi(x) = i$ si $x \in E_i$; inversement, une surjection $\pi : E \rightarrow I$ définit la partition $(E_i)_{i \in I}$ avec $E_i = \pi^{-1}(i)$.

L'application π fait correspondre à chaque x de E l'indice de la partie à laquelle il appartient. La démarche inverse consiste à choisir (attention, il y a une difficulté, voir ci-dessous), pour tout $i \in I$, un élément $x \in E_i$ (E_i est supposé non vide). On définit ainsi une application $s : I \rightarrow E$ telle que pour tout $i \in I$, $s(i) \in E_i$, autrement dit, on a $\pi \circ s = \text{id}_I$. L'application s est appelée section de π .

Système de représentants. On appelle système de représentants d'une partition une famille $(x_i)_{i \in I}$ d'éléments de E , un dans chacun des E_i . Se donner un système de représentants d'une partition $(E_i)_{i \in I}$ d'un ensemble E équivaut à se donner une application $s : I \rightarrow E$ telle que $s(i) \in E_i$ pour tout $i \in I$.

Axiome de choix. L'affirmation précédente, sur l'existence de l'application s section de π , autrement dit sur la possibilité de définir simultanément pour chaque i de I un élément $x \in E_i$, paraît naturelle. Elle a posé cependant une difficulté à ceux qui ont cherché à construire une base axiomatique à la théorie des ensembles dans les années 1900. Si I est fini ou infini dénombrable, elle se démontre par récurrence, mais, pour des ensembles I quelconques, elle ne résulte pas des autres axiomes de la théorie des ensembles et doit faire l'objet d'un axiome particulier, l'axiome de choix (on dit souvent axiome du choix).

La fondation axiomatique de la théorie des ensembles a été un domaine de recherches intenses autour de 1900. Ernst Zermelo (1871-1953) fut conduit à poser l'axiome de choix en 1904 pour montrer une conjecture de Georg Cantor (1845-1918). L'axiome posé par Zermelo fut l'objet, à l'époque, de discussions passionnées.

Il permet de justifier un certain nombre de raisonnements, par exemple, que le produit d'une famille d'ensembles non vides est non vide, puisqu'on peut en construire un élément. Nous rencontrerons quelques applications de l'axiome de choix dans la suite de ce livre. Signalons, sans que nous puissions entrer dans les détails, que ses relations avec les autres axiomes ont été étudiées.

Relation d'équivalence sur un ensemble et ensemble quotient. Nous avons déjà rappelé en 12.3 que se donner une relation d'équivalence $\mathcal{R}$ sur un ensemble E équivaut à se donner une partition de cet ensemble : les éléments de la partition sont les classes d'équivalence modulo $\mathcal{R}$ et l'ensemble de ces éléments est appelé ensemble quotient de E par $\mathcal{R}$ et noté $E/\mathcal{R}$. On définit ce qu'on appelle la projection canonique $\pi : E \rightarrow E/\mathcal{R}$ en associant à un élément x de E sa classe d'équivalence modulo $\mathcal{R}$.

Pour parler d'un élément de $E/\mathcal{R}$, on en prend souvent un représentant : si $\pi(x) = \alpha$, on dit que x est un représentant de la classe α . On écrira $\alpha = x \bmod \mathcal{R}$ quand cela ne créera pas de confusion.

Dans un lycée, si on quotientte l'ensemble des élèves par la relation *être dans la même classe*, l'ensemble quotient sera l'ensemble des classes du lycée et on parlera aussi bien de la Terminale *** que de la classe de Nicolas, où Nicolas est un élève de cette terminale qui joue le rôle de représentant de sa classe.

Choisir un système de représentants de E modulo $\mathcal{R}$, c'est choisir une famille $(x_\alpha)_{\alpha \in E/\mathcal{R}}$ d'éléments de E ou, de manière équivalente, une section $s : E/\mathcal{R} \rightarrow E$ de π . On vient de voir que, en général, cela suppose d'appliquer l'axiome de choix.

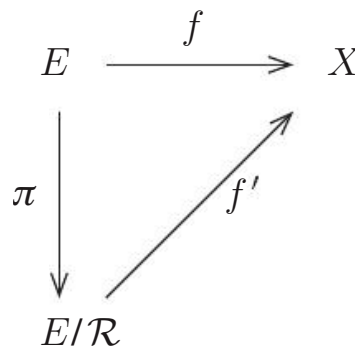
L'idée de quotient est une idée naturelle : on identifie entre eux certains éléments qui ont une propriété commune. On la retrouve dans de nombreux domaines des mathématiques :

- en topologie, on identifie des points du bord d'un rectangle pour définir un ruban de Möbius ou une bouteille de Klein ;
- en géométrie, on identifie des vecteurs colinéaires pour définir les espaces projectifs ; etc.

Pour pouvoir utiliser l'ensemble quotient $E/\mathcal{R}$, il faut pouvoir construire des applications l'ayant pour source. C'est très simple, mais pour bien mettre en évidence ce qu'on demande quand on aborde la notion de quotient, nous allons l'énoncer.

Propriété universelle de l'ensemble quotient. Soient E un ensemble, $\mathcal{R}$ une relation d'équivalence sur E , $\pi : E \rightarrow E/\mathcal{R}$ la projection canonique.

- 1) Propriété de factorisation unique. Pour toute application $f : E \rightarrow X$ constante sur les classes d'équivalence modulo $\mathcal{R}$, il existe une unique application $f' : E/\mathcal{R} \rightarrow X$ rendant commutatif le diagramme suivant :

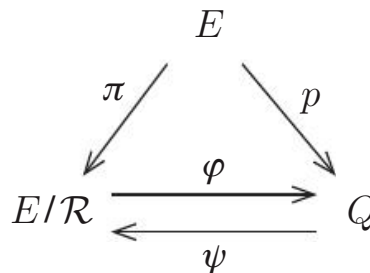


On dit que f se factorise par π (ou par $E/\mathcal{R}$) ou que f' est la factorisation de f par π (ou par $E/\mathcal{R}$).

- 2) Propriété d'unicité de $\pi : E \rightarrow E/\mathcal{R}$ à isomorphisme unique près. Soit $p : E \rightarrow Q$ une application possédant la même propriété que π , ce qui signifie que p est constante sur les classes d'équivalence modulo $\mathcal{R}$ et que, pour toute application $f : E \rightarrow X$ constante sur les classes d'équivalence modulo $\mathcal{R}$, il existe une unique application $f'' : Q \rightarrow X$ telle que $f = f'' \circ p$. Alors il existe un unique isomorphisme $\varphi : E/\mathcal{R} \rightarrow Q$ tel que $\varphi \circ \pi = p$.

Démonstration. 1) Soient α dans $E/\mathcal{R}$ et x dans E un représentant de $\alpha : \alpha = \pi(x)$. On définit f' par $f'(\alpha) = f(x)$, ce qui ne dépend pas du choix du représentant x de α puisque f est constante sur les classes d'équivalence.

- 2) Les propriétés de p et π permettent de construire de façon unique des applications $\varphi : E/\mathcal{R} \rightarrow Q$ et $\psi : Q \rightarrow E/\mathcal{R}$ telles que $\varphi \circ \pi = p$ et $\psi \circ p = \pi$, autrement dit, rendant commutatif le diagramme suivant.



Comme $\psi \circ \varphi \circ \pi = \pi = \text{id}_{E/\mathcal{R}} \circ \pi$ et $\varphi \circ \psi \circ p = p = \text{id}_Q \circ p$, les propriétés d'unicité donnent $\psi \circ \varphi = \text{id}_{E/\mathcal{R}}$ et $\varphi \circ \psi = \text{id}_Q$; par conséquent, φ et ψ sont des isomorphismes inverses l'un de l'autre. $\square$

Commentaires sur la propriété universelle du quotient.

- 1) Nous avons utilisé le mot isomorphisme à plusieurs reprises pour bien marquer le parallélisme entre cet énoncé et ceux sur les quotients de groupes ou d'anneaux que nous allons donner plus loin. Il s'agit ici d'isomorphismes d'ensembles, autrement dit de bijections.
- 2) On notera que f' est surjective si et seulement si f l'est.

- 3) On remarquera que la propriété universelle du quotient est une propriété du couple formé par l'ensemble $E/\mathcal{R}$ et l'application π . Les différents livres d'algèbre ne mettent pas toujours assez en évidence la propriété universelle du quotient (d'ensemble, de groupe, d'anneau, etc.), voire ne la donne pas. Il nous semble que c'est à tort. Elle permet de comprendre une fois pour toutes ce qu'il faut faire pour construire une application (un homomorphisme) de source un quotient. On peut aussi donner une image : c'est comme si la définition d'un quotient était statique et la propriété universelle dynamique.
- 4) La propriété universelle permet aussi de reconnaître l'égalité de deux homomorphismes de source le quotient : il suffit que leurs composés avec la surjection π soient égaux.
- 5) La deuxième partie de la propriété universelle permet de reconnaître que deux ensembles sont en bijection en vérifiant qu'ils satisfont à la même propriété universelle. C'est un résultat quasiment automatique dans ces propriétés universelles et que nous ne répéterons pas dans la suite ; le langage de la théorie des catégories (voir, par exemple, le livre de Lang) permettrait de le comprendre dans un cadre unificateur.

18.2 NOTION DE SOUS-GROUPE DISTINGUÉ

Notations générales pour les groupes. Nous utiliserons souvent les lettres G , G' , Γ pour noter des groupes.

L'élément neutre d'un groupe G est noté e_G et abrégé en e quand le contexte précise suffisamment de quel groupe il s'agit, ce qui est le cas le plus fréquent.

Soient A et B deux parties d'un groupe G . On note AB l'ensemble des produits d'un élément de A par un élément de B : $AB = \{ab, a \in A, b \in B\}$. Quand A est réduit à un seul élément a , on note aB plutôt que $\{a\}B$: $aB = \{ab, b \in B\}$; de même, si B est réduit à un seul élément b , on note $Ab = \{ab, a \in A\}$. De façon analogue, on note $A^{-1} = \{a^{-1}, a \in A\}$ l'ensemble des inverses des éléments de A . On a $(AB)C = A(BC)$ pour trois parties de G , mais on notera, qu'en général, $A A^{-1}$ n'est pas réduit à l'élément neutre de G .

En partant de l'idée de quotient. La notion de quotient en théorie des groupes remonte aux travaux de Galois. Mais la présentation de Galois, dans le cadre des groupes de permutations, est difficile à expliquer en peu de mots. C'est Otto Hölder (1859-1937) qui donnera une définition abstraite du quotient d'un groupe par un sous-groupe en 1889.

Que doit être une relation d'équivalence $\mathcal{R}$ sur un groupe G pour définir un quotient ? Par analogie avec le cas ensembliste, elle doit permettre de définir un groupe quotient $G/\mathcal{R}$, d'élément neutre e' , et un homomorphisme surjectif $\pi : G \rightarrow G/\mathcal{R}$ possédant une propriété universelle s'énonçant comme ci-dessus, en remplaçant les applications ensemblistes par des homomorphismes de groupes. Voyons ce que cela implique.

Posons $H = \ker(\pi)$; H est un sous-groupe de G .

Dire que deux éléments x et y de G sont équivalents modulo $\mathcal{R}$, équivaut à $\pi(x) = \pi(y)$, soit $\pi(x^{-1}y) = e'$ ou $x^{-1}y \in H$; ce qu'on peut encore exprimer en disant que y est de la forme xh avec $h \in H$, ou que : $y \in xH$.

On montre de même, à partir de $\pi(yx^{-1}) = e'$ que $x\mathcal{R}y$ équivaut à $y \in Hx$ ou $yx^{-1} \in H$.

La relation d'équivalence $\mathcal{R}$ peut donc se définir à l'aide d'un sous-groupe H de G qui doit posséder une propriété particulière : pour tout $x \in G$, on a $xH = Hx$.

Cette propriété équivaut à $xHx^{-1} = H$ pour tout x de G et aussi à $xHx^{-1} \subset H$ pour tout x de G ; en effet, en écrivant cette dernière relation pour x^{-1} , on obtient $x^{-1}Hx \subset H$, d'où $H \subset xHx^{-1}$ et l'égalité $xHx^{-1} = H$ en résulte.

On peut aussi énoncer la propriété de H sous la forme : pour tout $x \in G$ et tout $h \in H$, on a $xhx^{-1} \in H$. En utilisant le vocabulaire de 11.2, cela signifie que tout conjugué d'un élément de H est dans H . On est conduit à poser la définition suivante.

Définition : Sous-groupe distingué. Soient G un groupe et H un sous-groupe de G . On dit que H est sous-groupe distingué (on dit aussi normal) de G si tout conjugué d'un élément de H est dans H , autrement dit si, pour tout $x \in G$ et tout $h \in H$, on a $xhx^{-1} \in H$.

On dit aussi que le sous-groupe H est distingué dans G . On utilise souvent la notation $H \triangleleft G$ pour noter que H est un sous-groupe distingué de G .

On a vu ci-dessus qu'on peut énoncer plusieurs conditions équivalentes pour que H soit un sous-groupe distingué de G :

- $xHx^{-1} = H$ pour tout $x \in G$;
- $xH = Hx$ pour tout $x \in G$;
- $xHx^{-1} \subset H$ pour tout $x \in G$;
- pour tout $x \in G$ et tout $h \in H$, il existe $h' \in H$ tel que $xh = h'x$, etc.

Exemples dans S_3 . Le groupe S_3 est le groupe des permutations de l'ensemble $\{1, 2, 3\}$. Notons $\tau = (1\ 2)$ et $\sigma = (1\ 2\ 3)$. Le sous-groupe $H = \{e, \tau\}$ n'est pas distingué dans S_3 car le conjugué $\sigma \circ \tau \circ \sigma^{-1} = (2\ 3)$ n'appartient pas à H . Par contre, le sous-groupe $K = \{e, \sigma, \sigma^2\}$ est un sous-groupe distingué de S_3 puisque le conjugué d'un 3-cycle par un élément quelconque de S_3 est un 3-cycle donc un élément de K .

Résultats immédiats.

- 1) Dans un groupe quelconque G , les sous-groupes triviaux $\{e\}$ et G sont distingués.
- 2) Dans un groupe abélien, tout sous-groupe est distingué.
- 3) Le noyau d'un homomorphisme de groupes est un sous-groupe distingué : si $f: G \rightarrow G'$ est un homomorphisme de groupes, $\ker(f) \triangleleft G$.

Proposition

- 1) Soient $f: G \rightarrow G'$ un homomorphisme de groupes et K un sous-groupe distingué de G' . Alors $f^{-1}(K)$ est un sous-groupe distingué de G .

2) Soient $f: G \rightarrow G'$ un homomorphisme de groupes surjectif et H un sous-groupe distingué de G . Alors $f(H)$ est un sous-groupe distingué de G' .

Démonstration. 1) On sait que $f^{-1}(K)$ est un sous-groupe de G . Soient $x \in G$ et $h \in f^{-1}(K)$. On a $f(xhx^{-1}) = f(x)f(h)f(x)^{-1} \in K$, d'où $xhx^{-1} \in f^{-1}(K)$; $f^{-1}(K)$ est bien un sous-groupe distingué de G .

2) Si $h' \in f(H)$ et $y \in G'$, il existe $h \in H$ et $x \in G$ tels que $f(h) = h'$ et $f(x) = y$.

On voit que $yh'y^{-1} = f(xhx^{-1})$ est un élément de $f(K)$, d'où la conclusion. $\square$

L'image d'un sous-groupe distingué par un homomorphisme de groupes n'est pas nécessairement un sous-groupe distingué : prendre, par exemple, un sous-groupe H non distingué d'un groupe G et l'injection $i: H \rightarrow G$; H est un sous-groupe distingué de lui-même et $i(H)$ n'est pas distingué dans G .

Sous-groupe conjugué. Soient G un groupe, H un sous-groupe de G . Un sous-groupe de G de la forme gHg^{-1} où g est un élément de G est appelé conjugué de H . Deux sous-groupes H et H' de G sont dits conjugués s'il existe un élément $g \in G$ tel que $H' = gHg^{-1}$. Cette relation est une relation d'équivalence sur l'ensemble des sous-groupes de G . Un sous-groupe distingué est égal à tous ses conjugués.

Groupe simple. Un groupe G est dit simple si ses seuls sous-groupes distingués sont les deux sous-groupes triviaux $\{e\}$ et G .

Les groupes abéliens simples finis sont les groupes $(\mathbb{Z}/p\mathbb{Z}, +)$ pour p premier. La première famille de groupes simples finis non abéliens connue est celle des groupes alternés A_n pour $n \geq 5$ (voir exercice 18.13 pour le cas de A_5); c'est un résultat d'Évariste Galois qui passe pour être l'un des résultats les plus importants d'algèbre du XIX^e siècle. Depuis, la recherche des groupes simples finis s'est poursuivie. Elle semblait avoir abouti au début des années 1980 à une connaissance complète des groupes simples finis, avec des constructions magnifiques comme celle du groupe simple appelé le *Monstre*, ayant 808 017 424 794 512 875 886 459 904 961 710 757 005 754 368 000 000 000 éléments dont les merveilleuses propriétés ont été étudiées depuis; les démonstrations forment un ensemble de 500 articles totalisant 15 000 pages, travaux de plus de 100 chercheurs durant 30 ans, un record ! Dans la dernière édition de son *Algèbre*, Serge Lang indique qu'il resterait des problèmes dans ces 15 000 pages.

Classes à gauche et classes à droites. Plus généralement, un sous-groupe H d'un groupe G définit deux relations d'équivalence sur G : $\mathcal{R}_g$ et $\mathcal{R}_d$. Deux éléments x et y de G sont équivalents modulo $\mathcal{R}_g$ si $x^{-1}y \in H$, ce qui équivaut à $y \in xH$, et sont équivalents modulo $\mathcal{R}_d$ si $yx^{-1} \in H$, ce qui équivaut à $y \in Hx$. On vérifie facilement qu'il s'agit bien de relations d'équivalence sur H . La classe d'équivalence de x modulo $\mathcal{R}_g$ est xH ; elle est appelée classe à gauche modulo H . De même, la classe d'équivalence de x modulo $\mathcal{R}_d$ est Hx ; elle est appelée classe à droite modulo H .

Un sous-groupe H d'un groupe G est donc distingué dans G quand les classes à gauche et les classes à droite modulo H coïncident.

Si G est un groupe fini, les classes à gauche (ou à droite) ont toutes le même nombre d'éléments que H , ce qui montre que $|H|$ divise $|G|$.

Indice d'un sous-groupe. Avec les notations précédentes, l'application $x \mapsto x^{-1}$ induit une bijection de l'ensemble des classes à gauche sur l'ensemble des classes à droite : à la classe xH , on associe la classe Hx^{-1} . La vérification est facile (on notera que l'application $xH \mapsto Hx$ n'est pas nécessairement bien définie : prendre l'exemple du sous-groupe H engendré par $(1\ 2)$ dans S_3).

Par conséquent, le nombre (éventuellement infini) de classes à gauche est égal au nombre de classes à droite. Ce nombre est appelé indice du sous-groupe. Si H est un sous-groupe distingué de G , son indice est l'ordre de $|G/H|$.

18.3 GROUPE QUOTIENT

Soient G un groupe et $H \triangleleft G$. On a vu en 18.2 que la relation d'équivalence $x\mathcal{R}y$ définie par H peut s'exprimer de différentes façons : $y \in xH$, $x^{-1}y \in H$, $y \in Hx$, $yx^{-1} \in H$. On note G/H l'ensemble quotient de G par $\mathcal{R}$ et $\pi : G \rightarrow G/H$ la projection canonique.

On définit une loi de groupe sur G/H de la façon suivante. On prend deux éléments de G/H ; ils sont de la forme $\pi(x)$ et $\pi(x')$: on définit leur produit comme étant $\pi(xx')$. Il faut vérifier que cette définition ne dépend pas des choix de x et x' représentant les deux éléments de G/H . Si y et y' représentent les mêmes éléments, il existe h et h' dans H tels que $y = xh$, $y' = x'h'$; comme H est distingué dans G , il existe $h'' \in H$ tel que $hx' = x'h'' \in H$ et on a $yy' = xx'h''h'$, ce qui prouve que $\pi(yy') = \pi(xx')$.

On vérifie alors facilement qu'on a défini une loi de groupe sur G/H ; l'élément neutre est la classe H de chacun des éléments de H , π est un homomorphisme surjectif de groupes et $\ker(\pi) = H$. On peut dire qu'on *tue* les éléments de H en les identifiant à l'élément neutre.

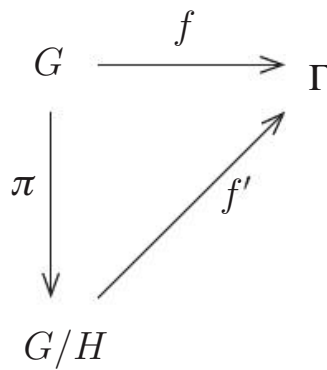
Le théorème de Lagrange (voir 11.4.2) permet d'affirmer que, si G est un groupe fini, on a

$$\frac{|G|}{|H|} = \frac{|G|}{|H|}.$$

On notera que, si $H = G$, alors $G/H = \{e_{G/H}\}$ et si $H = \{e_G\}$, on peut identifier G/H à G .

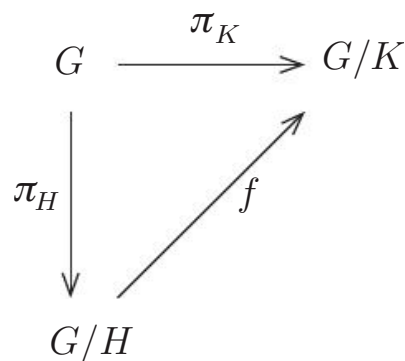
Pour construire des homomorphismes de source G/H , établissons maintenant la propriété universelle de la projection canonique $\pi : G \rightarrow G/H$.

Théorème : Propriété universelle du quotient d'un groupe par un sous-groupe distingué. Soient G un groupe, $H \triangleleft G$ et $\pi : G \rightarrow G/H$ la projection canonique. *Propriété de factorisation unique.* Pour tout groupe Γ et tout homomorphisme de groupes $f : G \rightarrow \Gamma$ tel que $f(H) = \{e_\Gamma\}$, il existe un unique homomorphisme de groupes $f' : G/H \rightarrow \Gamma$ rendant commutatif le diagramme :



Démonstration. Soit $\alpha \in G/H$; il existe $x \in G$ tel que $\pi(x) = \alpha$; on pose $f'(\alpha) = f(x)$. On vérifie facilement que f' ainsi défini est un homomorphisme de groupes rendant le diagramme précédent commutatif. L'unicité de f' résulte de la surjectivité de π . $\square$

Corollaire 1. Soient G un groupe, H et K deux sous-groupes distingués de G . On suppose $H \subset K$; alors il existe un unique homomorphisme de groupes $f : G/H \rightarrow G/K$ rendant commutatif le diagramme :



Démonstration. L'application de la propriété universelle est immédiate, puisque $K \supset H$ implique $\pi_K(H) = \{e\}$. Pour décrire f , on peut remarquer que c'est l'application $x \bmod H \mapsto x \bmod K$. On peut aussi dire qu'on *tue* plus quand on quotiente par K . $\square$

Corollaire 2. Soient $f : G \rightarrow \Gamma$ un homomorphisme de groupes et $\pi : G \rightarrow G/\ker(f)$ la projection canonique. La factorisation $f' : G/\ker(f) \rightarrow \Gamma$ de f par π est injective et définit un isomorphisme de groupes φ de $G/\ker(f)$ sur $\text{im}(f)$.

Démonstration. Il suffit de montrer que f' est injective, autrement dit que $\ker(f')$ est réduit à l'élément neutre de $G/\ker(f)$. Si $\alpha = \pi(x) \in G/\ker(f)$ et si $f'(\alpha) = e_\Gamma$, on a $f(x) = f'(\pi(x)) = e_\Gamma$, donc $x \in \ker(f)$ et α est l'élément neutre de $G/\ker(f)$. $\square$

Conséquences

- 1) Le résultat précédent permet de connaître, on pourrait dire *a priori*, la structure de l'image d'un homomorphisme à partir de sa source et de son noyau.
- 2) Il donne une factorisation d'un homomorphisme de groupes :

$$\begin{array}{ccc} G & \xrightarrow{f} & \Gamma \\ \pi \downarrow & & \text{incl} \uparrow \\ G / \ker(f) & \xrightarrow{\varphi} & \text{im}(f) \end{array}$$

où φ est un isomorphisme d'après le corollaire 2.

- 3) Si G est un groupe fini, on a les égalités entre cardinaux :

$$|\text{im}(f)| = \left| \frac{G}{\ker(f)} \right| = \frac{|G|}{|\ker(f)|}.$$

Corollaire 3. Soit $\varphi : G \rightarrow G'$ un isomorphisme de groupes, $H \triangleleft G$ et $H' = \varphi(H)$, $\pi : G \rightarrow G/H$ et $\pi' : G' \rightarrow G'/H'$. L'isomorphisme φ permet de définir un isomorphisme $\psi : G/H \rightarrow G'/H'$ tel que le diagramme

$$\begin{array}{ccccc} G & \xrightarrow{\varphi} & G' & \xrightarrow{\varphi^{-1}} & G \\ \pi \downarrow & & \pi' \downarrow & & \pi \downarrow \\ G/H & \xrightarrow{\psi} & G'/H' & \xrightarrow{\psi^{-1}} & G/H \end{array}$$

commute.

Démonstration. Le résultat paraît naturel, mais il n'est pas tout à fait évident et c'est une bonne occasion de s'exercer à manier la propriété universelle du quotient. On écrit souvent que ψ est défini par $x \bmod H \mapsto \varphi(x) \bmod H'$; c'est vrai, mais cela demande des vérifications qui sont données automatiquement par l'usage de la propriété universelle.

L'existence d'un homomorphisme $\psi : G/H \rightarrow G'/H'$ rendant commutatif le carré de gauche résulte de la propriété universelle puisque $(\pi' \circ \varphi)(H) = \pi'(H')$ est l'élément neutre de G'/H' .

L'existence d'un homomorphisme $\psi' : G'/H' \rightarrow G/H$ rendant commutatif le carré de droite résulte de la propriété universelle puisque $\pi(\varphi^{-1}(H')) = \pi(H)$ est l'élément neutre de G/H . Mais on ne sait pas à ce stade que ψ' est l'inverse de ψ .

Pour comparer les homomorphismes $\psi' \circ \psi$ et $\text{id}_{G/H}$, on remarque qu'ils rendent commutatifs les deux diagrammes suivants :

18.4 • Correspondance entre sous-groupes d'un groupe et sous-groupes

$$\begin{array}{ccc}
 G & \xrightarrow{\varphi^{-1} \circ \varphi} & G \\
 \pi \downarrow & & \pi \downarrow \\
 G/H & \xrightarrow{\psi' \circ \psi} & G/H
 \end{array}
 \qquad
 \begin{array}{ccc}
 G & \xrightarrow{\varphi^{-1} \circ \varphi} & G \\
 \pi \downarrow & & \pi \downarrow \\
 G/H & \xrightarrow{\text{id}_{G/H}} & G/H
 \end{array}$$

où $\varphi^{-1} \circ \varphi = \text{id}_G$. L'unicité de l'homomorphisme dans la propriété universelle montre qu'ils sont donc égaux.

On montre de même l'égalité $\psi \circ \psi' = \text{id}(G'/H')$, ce qui permet de conclure que ψ est un isomorphisme d'inverse ψ' . $\square$

Attention, si H et H' sont deux sous-groupes distingués isomorphes d'un groupe G , les quotients G/H et G/H' ne sont pas nécessairement isomorphes. Par exemple, si $G = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, $H = \text{Gr}((1,0))$, $H' = \text{Gr}((0,2))$, on a $G/H \simeq \mathbb{Z}/4\mathbb{Z}$ et $G/H' \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ qui ne sont pas isomorphes.

Application. $a\mathbb{Z}/ab\mathbb{Z} \simeq \mathbb{Z}/b\mathbb{Z}$

Pour tout entier n , on a construit, en 12.3 et 12.4, l'anneau $\mathbb{Z}/n\mathbb{Z}$ et la projection $\pi : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$. Il s'agit ici de groupes additifs. Soient a et b des entiers non nuls. L'application $k \mapsto ak$ est un isomorphisme $\mathbb{Z} \rightarrow a\mathbb{Z}$; l'image du groupe $b\mathbb{Z}$ est le groupe $ab\mathbb{Z}$. Le résultat précédent donne l'isomorphisme $\mathbb{Z}/b\mathbb{Z} \simeq a\mathbb{Z}/ab\mathbb{Z}$ qui se décrit simplement par $k \bmod b \mapsto ka \bmod ab$.

Par exemple, si $b = 5$ et $a = 3$, les éléments de $\mathbb{Z}/15\mathbb{Z}$ sont représentés par les entiers $k, 0 \leq k \leq 14$, les éléments de $3\mathbb{Z}/15\mathbb{Z}$ sont représentés par les entiers $0, 3, 6, 9, 12$, les éléments de $\mathbb{Z}/5\mathbb{Z}$ sont représentés par les entiers $0, 1, 2, 3, 4$; l'isomorphisme $\mathbb{Z}/5\mathbb{Z} \rightarrow 3\mathbb{Z}/15\mathbb{Z}$ associe la classe de $3k$ modulo 15 à la classe de k modulo 5.

18.4 CORRESPONDANCE ENTRE SOUS-GROUPES D'UN GROUPE ET SOUS-GROUPES D'UN DE SES QUOTIENTS

Présentation de la correspondance entre sous-groupes distingués. Soient G un groupe, $H \triangleleft G$ et $\pi : G \rightarrow G/H$ la projection canonique. Cette section a pour but d'étudier la correspondance définie par π entre sous-groupes de G et sous-groupes de G/H .

Notons :

- $\mathcal{A}$ l'ensemble des sous-groupes de G contenant H ;
- $\mathcal{A}'$ le sous-ensemble de $\mathcal{A}$ formé des sous-groupes distingués de G contenant H ;
- $\mathcal{B}$ l'ensemble des sous-groupes de G/H ;
- $\mathcal{B}'$ le sous-ensemble de $\mathcal{B}$ formé des sous-groupes distingués de G/H .

On définit une application $\Phi : \mathcal{A} \rightarrow \mathcal{B}$ par $\Phi(K) = \pi(K)$ pour tout K de $\mathcal{A}$ et une application $\Psi : \mathcal{B} \rightarrow \mathcal{A}$ par $\Psi(L) = \pi^{-1}(L)$ pour tout L de $\mathcal{B}$.

Proposition : Propriétés de la correspondance

- 1) Les applications Φ et Ψ sont inverses l'une de l'autre.
- 2) Elles se restreignent en des applications $\Phi' : \mathcal{A}' \rightarrow \mathcal{B}'$ et $\Psi' : \mathcal{B}' \rightarrow \mathcal{A}'$ inverses l'une de l'autre.
- 3) Pour tout $K \in \mathcal{A}$, on a :

$$K/H \simeq \pi(K).$$

- 4) Pour tout $K \in \mathcal{A}'$, on a :

$$(G/H)/(K/H) \simeq G/K.$$

- 5) Si K est un sous-groupe distingué quelconque de G , $\pi(K)$ est un sous-groupe distingué de G/H et $(G/H)/\pi(K) \simeq G/L$, où L est le sous-groupe distingué de G engendré par $H \cup K$, égal à $\pi^{-1}(\pi(K))$.

Démonstration. 1) Soit $K \in \mathcal{A}$. Si $g \in \pi^{-1}(\pi(K))$, on a $\pi(g) = \pi(k)$ avec $k \in K$; il existe $h \in H$ tel que $g = kh$; comme $H \subset K$, on a $g \in K$. D'où $\pi^{-1}(\pi(K)) \subset K$. L'inclusion opposée étant ensembliste, on a $\Psi \circ \Phi = \text{id}_{\mathcal{A}}$. L'égalité $\Phi \circ \Psi = \text{id}_{\mathcal{B}}$ est également ensembliste.

- 2) Voir la proposition de 18.2.

- 3) Considérons la restriction $\pi|_K : K \rightarrow G/H$. Son image est $\pi(K)$; son noyau est $K \cap H = H$. D'où $K/H \simeq \pi(K)$.

- 4) Notons $\pi_K : G \rightarrow G/K$ la projection canonique. Comme $H \subset K$, $\pi_K(H) = \{e_{G/K}\}$ et il existe un unique homomorphisme de groupes $p : G/H \rightarrow G/K$ rendant commutatif le diagramme :

$$\begin{array}{ccc} G & \xrightarrow{\pi_K} & G/K \\ \pi \downarrow & \nearrow p & \\ & G/H & \end{array}$$

La description de p est très simple : pour tout x de G , on a $p(x \text{ mod } H) = x \text{ mod } K$. Le noyau de p est donc $\pi(K) \simeq K/H$. Comme π_K est surjectif, p est surjectif et on obtient le résultat avec le corollaire 2 de 18.3.

- 5) On a $L \in \mathcal{A}'$, $\pi(L) = L/H$ d'après le 3) ; comme $\pi(L) = \pi(K)$, le 4) donne le résultat. $\square$

Exemple des quotients de $\mathbb{Z}$. Montrons comment les résultats précédents s'appliquent aux structures de groupes de ces anneaux.

Comme $\mathbb{Z}/n\mathbb{Z}$ est commutatif, ses sous-groupes sont distingués. Ils correspondent aux sous-groupes de $\mathbb{Z}$ contenant $n\mathbb{Z}$; ils sont donc de la forme $d\mathbb{Z}$ où d doit être un diviseur de n .

Soient $n = ab$ un produit de deux entiers. Le 4) du théorème donne un isomorphisme de groupes additifs $(\mathbb{Z}/ab\mathbb{Z})/(a\mathbb{Z}/ab\mathbb{Z}) \simeq \mathbb{Z}/a\mathbb{Z}$.

En reprenant les groupes de l'exemple de 18.3, l'isomorphisme $\varphi : (\mathbb{Z}/15\mathbb{Z})/(3\mathbb{Z}/15\mathbb{Z}) \rightarrow \mathbb{Z}/3\mathbb{Z}$ associe à la classe définie par un entier k dans $(\mathbb{Z}/15\mathbb{Z})/(3\mathbb{Z}/15\mathbb{Z})$ la classe de k modulo 3 en utilisant le diagramme commutative ci-dessus. Par exemple, 7 et 10 définissent le même élément α de $(\mathbb{Z}/15\mathbb{Z})/(3\mathbb{Z}/15\mathbb{Z})$ car la différence de leurs classes dans $\mathbb{Z}/15\mathbb{Z}$ est la classe de 3 qui appartient à $3\mathbb{Z}/15\mathbb{Z}$; l'image par φ de α est la classe de 1 dans $\mathbb{Z}/3\mathbb{Z}$, etc.

Remarque. On pourrait envisager une autre situation : celle d'un homomorphisme surjectif de groupes $f : G \rightarrow G'$ et la correspondance entre sous-groupes de G et sous-groupes de G' qu'il définit en prenant l'image directe ou réciproque par f .

Cette situation se ramène à celle étudiée dans cette section, puisqu'il existe un isomorphisme $\varphi : G' \rightarrow G/\ker(f)$ tel que $\varphi \circ f = \pi$.

18.5 PRODUITS DE GROUPES

La décomposition d'un groupe en produit permet, quand elle est possible, de mieux comprendre la structure du groupe et de ramener les calculs dans le groupe en calculs dans chacune des composantes.

Produit de deux groupes. Soient $(G_1, *_1)$ et $(G_2, *_2)$ deux groupes. On appelle produit de G_1 et G_2 le groupe $(G, *)$, noté $G_1 \times G_2$, dont l'ensemble des éléments est le produit $G = G_1 \times G_2$ et dont la loi est définie par $(x_1, x_2) * (y_1, y_2) = (x_1 *_1 y_1, x_2 *_2 y_2)$, autrement dit la multiplication se fait composante par composante.

Les projections $p_1 : G_1 \times G_2 \rightarrow G_1$ et $p_2 : G_1 \times G_2 \rightarrow G_2$ définissent des homomorphismes de groupes.

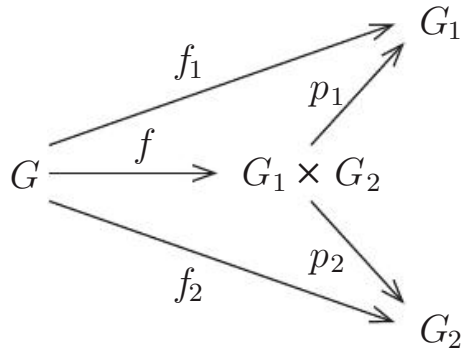
Les vérifications sont immédiates.

Si les lois de G_1 et G_2 sont notées additivement, la loi sur le produit est notée additivement ; elle est définie par $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$; de même, si les deux lois sont notées multiplicativement, la loi sur le produit est notée multiplicativement ; elle est définie par $(x_1, x_2)(y_1, y_2) = (x_1 y_1, x_2 y_2)$.

On définit de même le produit $\prod_{i \in I} G_i$ d'une famille $(G_i)_{i \in I}$ de groupes indexée par un ensemble I , fini ou infini : la loi du produit est définie par $(x_i)_{i \in I} (y_i)_{i \in I} = (x_i y_i)_{i \in I}$. Pour tout $i \in I$, la projection $p_i : \prod_{i \in I} G_i \rightarrow G_i$ est un homomorphisme de groupes. Si I est vide, le produit est un groupe réduit à un élément.

Les propriétés universelles suivantes, qui remontent à un article de Mac Lane de 1948, permettent de construire des homomorphismes dont le but est un produit. Les lois des différents groupes seront toutes notées multiplicativement.

Propriété universelle du produit de deux groupes. Soient G_1 et G_2 deux groupes. Pour tout groupe G , tout homomorphisme $f_1 : G \rightarrow G_1$ et tout homomorphisme $f_2 : G \rightarrow G_2$, il existe un unique homomorphisme $f : G \rightarrow G_1 \times G_2$, noté $f = (f_1, f_2)$, tel que le diagramme suivant commute :

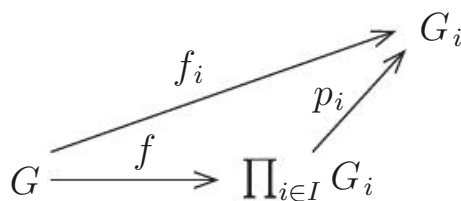


Démonstration. La seule façon possible de définir f est de poser $f(x) = (f_1(x), f_2(x))$ et tout marche bien. $\square$

La propriété universelle est une propriété du groupe $G_1 \times G_2$ et des projections p_1 et p_2 , autrement dit, du triplet $(G_1 \times G_2, p_1, p_2)$.

Propriété universelle du produit d'une famille de groupes. Soit $(G_i)_{i \in I}$ une famille de groupes.

Pour tout groupe G et toute famille $(f_i : G \rightarrow G_i)_{i \in I}$ d'homomorphismes, il existe un homomorphisme unique $f : G \rightarrow \prod_{i \in I} G_i$, noté $f = (f_i)_{i \in I}$, tel que, pour tout $i \in I$, le diagramme suivant commute :



Démonstration. On doit poser $f(x) = (f_i(x))_{i \in I}$. $\square$

La propriété universelle est une propriété du groupe $\prod_{i \in I} G_i$ et de la famille de projections $(p_i)_{i \in I}$.

Décomposition du groupe $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}, +\right)$ en produit. Soulignons que nous ne nous intéressons ici qu'à la structure de groupe additif de $\mathbb{Z}/n\mathbb{Z}$. Les résultats suivants s'étendent à l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ sous le nom de théorème chinois (voir 19.8) et c'est sous cette dernière forme qu'ils seront en général appliqués.

Théorème chinois pour le groupe $\mathbb{Z}/n\mathbb{Z}$: cas où n est le produit de deux entiers premiers entre eux. Soient a et b des entiers premiers entre eux. Notons

$f_1 : \frac{\mathbb{Z}}{ab\mathbb{Z}} \rightarrow \frac{\mathbb{Z}}{a\mathbb{Z}}$ et $f_2 : \frac{\mathbb{Z}}{ab\mathbb{Z}} \rightarrow \frac{\mathbb{Z}}{b\mathbb{Z}}$ les homomorphismes de groupes définis par $f_1 : k \bmod ab \mapsto k \bmod a$ et $f_2 : k \bmod ab \mapsto k \bmod b$ (voir 18.3, corollaire 1).

L'homomorphisme

$$f = (f_1, f_2) : \frac{\mathbb{Z}}{ab\mathbb{Z}} \rightarrow \frac{\mathbb{Z}}{a\mathbb{Z}} \times \frac{\mathbb{Z}}{b\mathbb{Z}}$$

est un isomorphisme de groupes.

Démonstration. Comme la source et le but de f ont le même cardinal, ab , il suffit de montrer que f est injective. Si k est un entier tel que $f(k \bmod ab) = (0 \bmod a, 0 \bmod b)$, k est un multiple de a et de b . Comme a et b sont premiers entre eux, k est donc un multiple de ab , donc $k = 0 \bmod ab$, ce qui prouve l'injectivité de f . $\square$

La proposition donne, par exemple : $\mathbb{Z}/60\mathbb{Z} \simeq \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z}$ et, en l'appliquant une seconde fois, $\mathbb{Z}/60\mathbb{Z} \simeq \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$.

La démonstration précédente est un peu frustrante et on sent bien que dans les applications, la connaissance explicite de f^{-1} est souvent nécessaire. L'identité de Bézout donne (à l'aide de l'algorithme d'Euclide, voir 12.12.2 et l'exercice 12.3) des entiers u et v tels que $ua + vb = 1$. Posons $y_1 = vb \bmod ab$ et $y_2 = ua \bmod ab$. On a donc $f(y_1) = (1 \bmod a, 0 \bmod b)$ et $f(y_2) = (0 \bmod a, 1 \bmod b)$; par conséquent $f^{-1}(k \bmod a, l \bmod b) = ky_1 + ly_2$. La connaissance de y_1 permet de calculer y_2 , puisque $y_1 + y_2 = 1 \bmod n$. Nous reviendrons sur ces calculs en 19.8.

Théorème chinois pour le groupe $\mathbb{Z}/n\mathbb{Z}$: cas où n est le produit de r entiers premiers entre eux. Soient $a_1, \dots, a_r$ des entiers premiers entre eux deux à deux.

Posons $n = \prod_{1 \leq i \leq r} a_i$ et, pour $i = 1, \dots, r$, notons $f_i : \frac{\mathbb{Z}}{n\mathbb{Z}} \rightarrow \frac{\mathbb{Z}}{a_i\mathbb{Z}}$ l'homomorphisme de groupes défini comme ci-dessus ; on a $f_i : k \bmod n \mapsto k \bmod a_i$.

L'homomorphisme

$$f = (f_1, \dots, f_r) : \frac{\mathbb{Z}}{n\mathbb{Z}} \rightarrow \prod_{1 \leq i \leq r} \frac{\mathbb{Z}}{a_i\mathbb{Z}}$$

est un isomorphisme de groupes.

La démonstration s'obtient en généralisant celle de la propriété précédente. $\square$

L'isomorphisme réciproque s'obtient en déterminant des entiers y_i , $1 \leq i \leq r$ tels que $y_i = 1 \bmod a_i$ et $y_i = 0 \bmod a_j$ pour $j \neq i$. Pour déterminer y_1 , par exemple, on détermine les u et v de l'identité de Bézout tels que $ua_1 + va_2 \dots a_r = 1$ et on pose $y_1 = va_2 \dots a_r$. La connaissance de $r - 1$ des y_i permet de calculer le dernier, puisque la somme des y_i est égale à $1 \bmod n$. Là encore, renvoyons en 19.8 pour plus de détails.

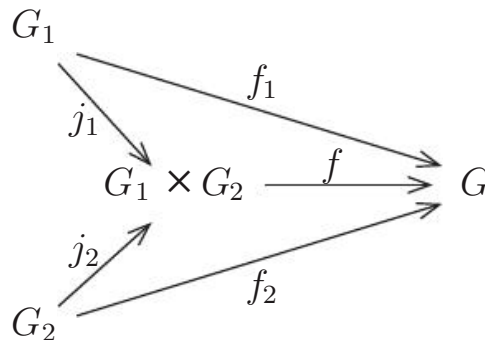
Somme de groupes abéliens. Le produit fini de groupes abéliens (attention, ce qui suit ne vaut pas pour les groupes quelconques) possède une autre propriété universelle, différente de la propriété du produit en ce sens que les homomorphismes vont en sens inverse comme on va le voir. On parle, quand on utilise cette propriété du produit de groupes abéliens, de la somme de groupes abéliens.

Soient G_1, G_2 deux groupes abéliens, $G_1 \times G_2$ leur produit, $j_1 : G_1 \rightarrow G_1 \times G_2, j_2 : G_2 \rightarrow G_1 \times G_2$, les applications définies par $j_1(x_1) = (x_1, 0), j_2(x_2) = (0, x_2)$ pour $x_1 \in G_1, x_2 \in G_2$.

Il est clair que j_1 et j_2 sont des homomorphismes de groupes.

Propriété universelle de la somme de deux groupes abéliens

Avec les données précédentes, pour tout groupe abélien G , tout homomorphisme $f_1 : G_1 \rightarrow G$ et tout homomorphisme $f_2 : G_2 \rightarrow G$, il existe un unique homomorphisme $f : G_1 \times G_2 \rightarrow G$ tel que le diagramme suivant commute :



Démonstration. Pour tout $(x_1, x_2) \in G_1 \times G_2$, on doit avoir

$$f(x_1, x_2) = f((x_1, 0) + (0, x_2)) = f(j_1(x_1)) + f(j_2(x_2)) = f_1(x_1) + f_2(x_2).$$

Comme il est clair qu'en définissant f ainsi on obtient un homomorphisme de groupes rendant le diagramme commutatif, la propriété est démontrée. $\square$

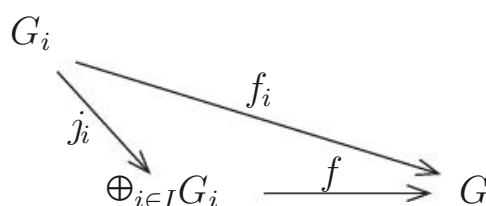
La propriété universelle est une propriété du groupe $G_1 \times G_2$ et des injections j_1 et j_2 , autrement dit, du triplet $(G_1 \times G_2, j_1, j_2)$.

Cette propriété se généralise aux produits finis de groupes abéliens, mais ne s'étend pas aux produits infinis de tels groupes. Voici le résultat qu'on peut obtenir.

Soit $(G_i)_{i \in I}$ une famille (finie ou infinie) de groupes abéliens. Notons $\bigoplus_{i \in I} G_i$ l'ensemble des familles $(x_i)_{i \in I}$ telles que $x_i \in G_i$ pour tout $i \in I$, avec la condition supplémentaire que les x_i non nuls de la famille sont *en nombre fini*. Il est clair que $\bigoplus_{i \in I} G_i$ est un groupe abélien : la somme de deux familles $(x_i)_{i \in I}, (y_i)_{i \in I}$ de $\bigoplus_{i \in I} G_i$ est la famille $(x_i + y_i)_{i \in I}$ dont les termes non nuls sont en nombre fini. On vérifie aussi que $\bigoplus_{i \in I} G_i$ est un sous-groupe de $\prod_{i \in I} G_i$; ce groupe est, en général, distinct de $\prod_{i \in I} G_i$ si I est infini (l'égalité des deux ensembles a lieu si les G_i non réduits à

l'élément neutre sont en nombre fini). On définit, pour tout $i \in I$, un homomorphisme $j_i: G_i \rightarrow \bigoplus_{i \in I} G_i$ en définissant $j_i(x_i)$ comme la famille dont tous les éléments sont nuls sauf celui d'indice i égal à x_i .

Propriété universelle de la somme d'une famille de groupes abéliens. Avec les données précédentes, pour tout groupe abélien G , toute famille $f_i: G_i \rightarrow G$ d'homomorphismes, il existe un unique homomorphisme $f: \bigoplus_{i \in I} G_i \rightarrow G$ tel que le diagramme commute pour tout $i \in I$.



Démonstration. Reprendre la démonstration précédente en calculant

$$f((x_i)_{i \in I}) = f\left(\sum_{i \in I} j_i(x_i)\right) = \sum_{i \in I} f_i(x_i).$$

La définition de f précédente a un sens car les sommes sur I se ramènent à des sommes finies puisque les familles $(x_i)_{i \in I}$ ne comportent qu'un nombre fini de termes non nuls. $\square$

18.6 GROUPES MONOGÈNES ET GROUPES CYCLIQUES

Soient G un groupe et x un élément de G . Rappelons (11.4.2) qu'on note $\text{Gr}(x)$ le sous-groupe de G engendré par x et que $\text{Gr}(x)$ est l'image de l'homomorphisme de groupes $f: \mathbb{Z} \rightarrow G$ défini par $f(1) = x$ (propriété universelle du groupe $(\mathbb{Z}, +)$, 11.5, prop. 3).

Définition. Soit G un groupe. On dit que G est un groupe monogène s'il existe un élément x de G tel que $G = \text{Gr}(x)$.

Une définition équivalente est de dire que G est un groupe monogène s'il existe un homomorphisme de groupes surjectif $f: (\mathbb{Z}, +) \rightarrow G$.

Deux situations peuvent se produire.

- 1) L'homomorphisme f est injectif. Alors f est un isomorphisme, G est isomorphe à $(\mathbb{Z}, +)$ et est infini.
- 2) L'homomorphisme f n'est pas injectif. Son noyau est un sous-groupe de $(\mathbb{Z}, +)$ non réduit à $\{0\}$; il est de la forme $n\mathbb{Z}$ avec $n > 0$, d'après 12.10. Le corollaire 2 de 18.3 montre que G est isomorphe au quotient $(\mathbb{Z}/n\mathbb{Z}, +)$ et est fini.

Les remarques précédentes classifient complètement les groupes monogènes en donnant $(\mathbb{Z}, +)$ comme modèle des groupes monogènes infinis et les $(\mathbb{Z}/n\mathbb{Z}, +)$, $n \geq 1$, comme modèles des groupes monogènes finis. Selon un usage assez général, on appellera groupe cyclique un groupe monogène fini (mais on trouve des textes qui appellent groupe cyclique tout groupe monogène, fini ou infini).

Proposition. *Tout sous-groupe d'un groupe monogène est un groupe monogène. En particulier, tout sous-groupe d'un groupe cyclique est un groupe cyclique.*

Démonstration. Soit H un sous-groupe d'un groupe monogène G .

- 1) Si G est infini, on sait que G est isomorphe à $\mathbb{Z}$; les sous-groupes de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$ d'après 12.10. Si $n = 0$, H est cyclique, réduit à un élément. Si $n \neq 0$, $n\mathbb{Z}$ est isomorphe à $\mathbb{Z}$, donc H est monogène infini.
- 2) Si G est fini, on sait que G est isomorphe à $\mathbb{Z}/n\mathbb{Z}$ avec $n > 0$; un sous-groupe H de G est donc isomorphe, comme on l'a vu en 18.4, à un groupe de la forme $d\mathbb{Z}/n\mathbb{Z}$ où d est un diviseur de n . Si on pose $n = ud$, alors $H \simeq \mathbb{Z}/u\mathbb{Z}$, toujours d'après 18.4. $\square$

18.7 ACTION D'UN GROUPE SUR UN ENSEMBLE

Définition 1 : action d'un groupe sur un ensemble. Soient G un groupe et E un ensemble. Une action (on dit aussi opération) de G sur E est la donnée d'un homomorphisme de groupes $\varphi : G \rightarrow S_E$ où S_E désigne le groupe des permutations de E , c'est-à-dire le groupe des bijections de E dans E .

On dit qu'on a défini un G -ensemble et que G agit (ou opère) sur E . Un G -ensemble est donc, au sens de cette définition, un triplet (E, G, φ) formé d'un ensemble E , d'un groupe G et d'un homomorphisme de groupes $\varphi : G \rightarrow S_E$. Par abus de langage, on parlera du G -ensemble E si cela ne crée pas de confusion.

Cette définition décrit bien l'action des éléments de G sur E comme des bijections de E : si $g \in G$ et si $x \in E$, l'action de g envoie x sur $\varphi(g)(x)$. Cette notation est assez lourde et on a envie d'une notation plus explicite comme $g.x$, pour l'action de $g \in G$ sur $x \in E$; cela conduit à reformuler la définition 1 en la définition équivalente suivante.

Définition 2 : action d'un groupe sur un ensemble. Soient G un groupe (dont la loi est notée par juxtaposition) et E un ensemble. Une action (on dit aussi opération) de G sur E est la donnée d'une application $G \times E \rightarrow E$, notée $(g, x) \mapsto g.x$ (ou $(g, x) \mapsto gx$ si cela ne crée pas de confusion), ayant les propriétés suivantes :

- 1) pour tout $x \in E$, on a $e.x = x$;
- 2) pour tout $x \in E$, tous $g, g' \in G$, on a $g.(g'.x) = (gg').x$.

L'équivalence des deux définitions est facile à vérifier : si on part de la définition 1, on a $\varphi(e_G) = \text{id}_E$ et $\varphi(g) \circ \varphi(g') = \varphi(gg')$; la définition 2 est vérifiée. Réciproquement, si on définit $\varphi(g)$ par $\varphi(g)(x) = g.x$, on montre que φ est un homomorphisme de groupes, puis que $\varphi(g)$ est inversible ; la définition 1 est vérifiée.

Donnons, même si nous n'en faisons pas usage dans la suite, la définition de morphisme de G -ensembles.

Définition 3 : G-ensembles. Soient deux G -ensembles E et F définis par $\varphi : G \rightarrow S_E$ et $\psi : G \rightarrow S_F$. Un morphisme (ou homomorphisme) du G -ensemble E vers le G -ensemble F est une application $f : E \rightarrow F$ telle que, pour tout $g \in G$, le diagramme

$$\begin{array}{ccc} E & \xrightarrow{\varphi(g)} & E \\ f \downarrow & & \downarrow f \\ F & \xrightarrow{\psi(g)} & F \end{array}$$

commute, autrement dit, telle que pour tout $g \in G$ et tout $x \in E$, $f(g.x) = g.(f(x))$.

Les exemples d'action de groupes sont nombreux et variés. Citons-en un avant de donner les définitions d'orbite et de stabilisateur.

Notons E le plan affine $\mathbb{R}^2$ et Ω un point de E . Le groupe G des rotations de centre Ω agit sur les points de E par $r.M = r(M)$, avec $r \in G$ et $M \in E$.

Définition 4 : orbite d'un élément. On appelle orbite d'un élément x d'un G -ensemble E et on note O_x l'ensemble des images de x par l'action des éléments de G :

$$O_x = \{g.x, g \in G\}.$$

On dira que x est fixé par G (où que x est un point fixe pour l'action de G) si son orbite n'a qu'un seul élément, x lui-même. On dira que G agit transitivement sur E s'il n'existe qu'une seule orbite, autrement dit si, pour tous $x, y \in E$, il existe $g \in G$ tel que $y = g.x$.

La relation $\mathcal{R}$ définie par $x\mathcal{R}y$ si x et y appartiennent à la même orbite est une relation d'équivalence. Les classes d'équivalence pour cette relation sont les orbites et elles forment une partition de E . On aura besoin pour certains exemples de choisir un système de représentants des classes d'équivalence modulo $\mathcal{R}$, ce qui revient à choisir un point par orbite (ici encore, l'axiome de choix peut être utile).

Définition 5 : stabilisateur d'un élément. On appelle stabilisateur d'un élément x d'un G -ensemble E et on note S_x l'ensemble des éléments de G laissant x fixé :

$$S_x = \{g \in G, g.x = x\}.$$

Le stabilisateur d'un élément de E est un sous-groupe de G comme il est facile de le vérifier.

Dans l'exemple ci-dessus, l'orbite de Ω est réduite à Ω et l'orbite d'un autre élément est un cercle de centre Ω , ce qui explique le choix du mot orbite. Le stabilisateur de Ω est G et le stabilisateur d'un autre élément est réduit à l'identité. On peut aussi dire que Ω est fixé par tout élément de G .

Proposition 1. *Soit G un groupe agissant sur un ensemble E . Si x et y sont deux éléments appartenant à une même orbite de E , alors les groupes S_x et S_y sont conjugués :*

$$\text{Si } y = \gamma.x \text{ alors } S_y = \gamma S_x \gamma^{-1}.$$

Démonstration. Soit $g \in S_x$. On a $(\gamma g \gamma^{-1}).y = (\gamma g \gamma^{-1}).(\gamma.x) = (\gamma g).x = \gamma.x = y$, ce qui prouve que $\gamma S_x \gamma^{-1} \subset S_y$. De même, comme $x = \gamma^{-1}y$, on a $\gamma^{-1}S_y \gamma \subset S_x$, d'où $S_y \subset \gamma S_x \gamma^{-1}$; les deux inclusions prouvent la proposition. $\square$

Proposition 2. *Soit G un groupe fini agissant sur un ensemble E . Pour tout $x \in E$, on a :*

$$|G| = |S_x| |O_x|$$

Démonstration. Soit $x \in E$. Considérons l'application $f : g \mapsto g.x$ de G dans O_x . Cette application est surjective et $f^{-1}(\{x\}) = S_x$. Soit $y = \gamma.x$ dans O_x ; la condition $g \in f^{-1}(\{y\})$ équivaut à $g.x = \gamma.x$, donc à $\gamma^{-1}g \in S_x$ ou encore à $g \in \gamma S_x$. Pour tout y de O_x , les ensembles $f^{-1}(\{y\})$ ont donc tous le même cardinal, égal à $|S_x|$. Comme ils forment une partition de G , la proposition en résulte. $\square$

Formule des classes. La formule suivante, dont la démonstration est élémentaire, est très utile. On l'appelle aussi *Équation des classes* ou *Formule des orbites*.

Proposition 3 : Formule des classes. *Soit G un groupe fini agissant sur un ensemble fini E . On suppose que cette action définit r orbites distinctes $O_1, \dots, O_r$ et on choisit un système de représentants $x_1, \dots, x_r$ de ces orbites ($x_i \in O_i$ pour $1 \leq i \leq r$). On a :*

$$|E| = \sum_{1 \leq i \leq r} |O_i| = \sum_{1 \leq i \leq r} \frac{|G|}{|S_{x_i}|}.$$

Démonstration. La première égalité résulte de la partition de E en orbites disjointes, la seconde de la proposition 2 précédente. $\square$

Soient n un entier non nul et $E = \{1, \dots, n\}$. Le groupe S_n agit sur E de manière naturelle par : $\sigma.k = \sigma(k)$. Cette action se restreint à tout sous-groupe de S_n . Elle permet de donner une démonstration de la proposition 3 de 11.6.

Proposition 4 : Décomposition d'une permutation en produit de cycles dis-joints. *Toute permutation se décompose en un produit de cycles disjoints (de manière unique à l'ordre des cycles près).*

Démonstration. Soit $\sigma \in S_r$. Le groupe $\text{Gr}(\sigma)$ agit sur E et E est une union d'orbites disjointes sous cette action. Si O_i est l'orbite d'un $i \in E$ et que $|O_i| = r$, la restriction de σ à O_i définit le cycle $(i \ \sigma(i) \dots \sigma^{r-1}(i))$. À chaque orbite correspond donc un cycle ; ces cycles commutent entre eux deux à deux et leur produit est σ . D'autre part, si σ est décomposé en produit de cycles disjoints, chaque élément i de E n'apparaît que dans un seul cycle, le cycle $(i \ \sigma(i) \dots)$; ceci prouve l'unicité de la décomposition (à l'ordre des cycles près). $\square$

Théorème de Cauchy. *Soient G un groupe fini d'ordre n et p un diviseur premier de n . Alors, il existe un élément d'ordre p dans G .*

Commentaire. Ce théorème a un petit air de réciproque du théorème de Lagrange (voir 11.4.2) : dans un groupe fini, l'ordre d'un élément divise l'ordre du groupe, mais il est loin d'être évident. Il a été énoncé par Cauchy, qui ne disposait pas encore du langage de la théorie des groupes, en 1815.

Le théorème est une conséquence directe des théorèmes de Ludwig Sylow (1832-1918), que nous n'avons pas la place d'étudier ici. Nous allons donner une démonstration dans le cas des groupes abéliens (qui peut conduire à une démonstration du cas général), puis une démonstration astucieuse du cas général proposée par James Mac Kay en 1959.

Démonstration du théorème de Cauchy pour les groupes abéliens. Raisonnons par récurrence sur n . Le résultat est vrai si $n = 1$. Supposons $n > 1$, choisissons un élément $x \neq e_G$ dans G et notons m son ordre. Si m est divisible par p , $x^{m/p}$ est d'ordre p , ce qui montre le théorème dans ce cas. Si m n'est pas divisible par p , posons $H = \text{Gr}(x)$. Comme G est abélien, H est un sous-groupe distingué de G et G/H est un groupe ; on a $|G/H| = n/m$, nombre divisible par p . On peut donc appliquer l'hypothèse de récurrence au groupe G/H : il existe un élément γ d'ordre p dans G/H . Posons $\gamma = y \bmod H$ et notons u l'ordre de y . Comme $\gamma^u = e_{G/H}$, l'ordre p de γ divise u . Donc l'ordre de y est un multiple de p et on est ramené au premier cas. $\square$

Démonstration de Mac Kay du théorème de Cauchy. On définit le sous-ensemble E de G^p par $E = \{(x_1, \dots, x_p), x_1 \dots x_p = e\}$. Tout choix de $p-1$ éléments $x_1, \dots, x_{p-1}$ de G définit un x_p unique vérifiant $x_1 \dots x_p = e$; par conséquent, $|E| = n^{p-1}$. Si $x_1 \dots x_p = e$, alors $x_1 \dots x_{p-1} = x_p^{-1}$, donc $x_p x_1 \dots x_{p-1} = e$, ce qui montre qu'on peut définir une application $f : E \rightarrow E$ en posant $f : (x_1, \dots, x_p) = (x_p, x_1, \dots, x_{p-1})$. On définit alors une action du sous-groupe $\text{Gr}(f)$ de S_E sur E par $f.(x_1, \dots, x_p) = (x_p, x_1, \dots, x_{p-1})$. Le groupe $\text{Gr}(f)$ est d'ordre p . Les orbites pour son action sur E ont donc pour ordre 1 ou p . Notons s le nombre de points fixes et t le nombre d'orbites à p éléments. L'équation

des classes s'écrit $n^{p-1} = s + pt$. On en déduit que s est divisible par p . Comme $(e, \dots, e) \in E$, on a $s \geq 1$, ce qui prouve que $s \geq p$. Mais un point fixe de E autre que $(e, \dots, e)$ est de la forme $(x, \dots, x)$ avec $x \neq e$ et $x^p = e$, ce qui démontre le théorème de Cauchy. $\square$

Actions du groupe linéaire et de ses sous-groupes. Le groupe linéaire $GL(E)$ d'un espace vectoriel E opère naturellement sur E : si $f : E \rightarrow E$ est un isomorphisme d'espace vectoriel et si $x \in E$, on pose $f.x = f(x)$. Cette action se restreint à tout sous-groupe de $GL(E)$. Souvent, on considérera des actions de $GL(E)$ sur des sous-ensembles bien choisis de E .

De même, le groupe affine d'un espace affine agit sur les points de l'espace, ce qui donne de nombreuses applications géométriques.

Actions d'un groupe sur lui-même par translation. Soit G un groupe. La multiplication à gauche par les éléments de G définit une action du groupe G sur l'ensemble G : pour $g \in G$ et $x \in G$, on pose $g.x = gx$. Chaque $g \in G$ définit ainsi une bijection de G et on définit un G -ensemble par l'injection $G \rightarrow S_G$ définie par $g \mapsto (x \mapsto gx)$. Ainsi G est-il représenté comme un sous-groupe d'un groupe de permutation.

Actions d'un groupe sur lui-même par conjugaison. Une autre action d'un groupe G sur lui-même est l'action par conjugaison. Pour $g \in G$ et $x \in G$, on pose $g.x = gxg^{-1}$ (on notera que cette action ne peut être notée par juxtaposition, la loi de G ayant ce privilège). Cette action est donnée par l'homomorphisme $G \rightarrow S_G$ qui associe à $g \in G$ l'automorphisme dit intérieur $x \mapsto gxg^{-1}$ appelé aussi conjugaison par g . Le stabilisateur d'un élément x de G pour cette action est appelé centralisateur de x dans G ; il a pour éléments les g de G tels que $gxg^{-1} = x$, autrement dit $gx = xg$; c'est l'ensemble des éléments de G commutant avec x .

Définition : p-groupe. Soit p un nombre premier. On dit qu'un groupe fini est un p -groupe si son ordre est une puissance de p (en particulier, un groupe réduit à l'élément neutre est un p -groupe).

Théorème de Burnside. *Le centre d'un p -groupe non trivial n'est jamais réduit à l'élément neutre.*

Démonstration. Soient G un p -groupe, $Z(G)$ son centre et faisons agir G sur lui-même par conjugaison. Notons r le nombre d'orbites pour cette action et $x_1, \dots, x_r$ un système de représentants de ces orbites. L'orbite d'un élément x du centre est réduite à x . En particulier l'orbite de e_G est réduite à e_G ; comme G est non trivial, on a donc $r > 1$. Toute orbite non réduite à un élément a pour ordre un diviseur de $|G|$ donc une puissance de p , d'après la proposition 2 de cette section. Posons $|Z(G)| = s$. La formule des classes s'écrit :

$$|G| = \sum_{1 \leq i \leq r} |O_i| = s + \sum_{s+1 \leq i \leq r} |O_i|.$$

On voit que s est la différence de deux multiples de p , donc est un multiple de p . Comme $e_G \in Z(G)$, on a $s \geq 1$, donc $s \geq p$ et le résultat. $\square$

Le mathématicien anglais William Burnside (1852-1927) a publié de nombreux travaux sur la théorie des groupes finis et le premier livre (en anglais) consacré entièrement à la théorie des groupes, en 1897.

Exercices

Pour les exercices sur les groupes de permutation, il est important de connaître la formule montrée dans l'exercice 11.3 donnant le conjugué d'un cycle $c = (x_1 \dots x_r)$ par une permutation σ de S_n :

$$\sigma \circ c \circ \sigma^{-1} = (\sigma(x_1) \dots \sigma(x_r)).$$

18.1 Étude des groupes diédraux D_3 et D_4

On note D_n l'ensemble des isométries du plan affine euclidien qui laissent l'ensemble des sommets d'un polygone régulier de n côtés ($n \geq 3$) globalement invariant (les groupes d'isométries de deux polygones réguliers à n côtés sont isomorphes).

- 1) Montrer que D_n est un groupe d'ordre $2n$ appelé groupe diédral.
- 2) Montrer que D_3 est engendré par deux éléments r et s qui vérifient les relations $r^3 = e$, $s^2 = e$, $sr = r^{-1}s$ et que D_3 est isomorphe à S_3 .
- 3) Montrer que D_4 est engendré par deux éléments r et s qui vérifient les relations $r^4 = e$, $s^2 = e$, $sr = r^{-1}s$. Décrire ses éléments et donner sa table de multiplication.
- 4) Montrer que D_4 est isomorphe à un sous-groupe G de S_4 et décrire toutes les permutations qui appartiennent à G .
- 5) Donner tous les sous-groupes de D_4 . Préciser s'ils sont distingués ou non. Lorsqu'ils le sont, calculer les quotients de D_4 par ces sous-groupes.
- 6) Donner un exemple de deux sous-groupes H et K de D_4 , tels que H soit distingué dans K , K distingué dans D_4 et H non distingué dans D_4 .
- 7) Montrer que D_n est engendré par deux éléments r et s tels que $r^n = s^2 = e$ et $sr^k = r^{-k}s$ pour $1 \leq k \leq n-1$.

18.2 Exercices généraux sur les groupes

- 1) Soient E un ensemble, $(G, \cdot)$ un groupe, A l'ensemble des applications de E dans G . Montrer que A peut-être muni d'une structure de groupe en posant, pour $f, f' \in A$ et pour tout $x \in E$: $(f \cdot f')(x) = f(x) \cdot f'(x)$.
- 2) Soit G un groupe dont tous les éléments (sauf l'élément neutre) sont d'ordre 2. Montrer que G est abélien.
- 3) Déterminer les différentes structures possibles de groupes à n éléments pour $n \leq 7$ (le sens de l'expression *structure de groupe* est donné dans l'introduction à ce chapitre). Quelles structures de groupes à 8, 9, 10, 11, 12 éléments pouvez-vous citer ?
- 4) Pour tout nombre premier p , on considère le groupe G_p des matrices carrées inversibles d'ordre 2 à coefficients dans le corps $\mathbb{Z}/p\mathbb{Z}$. On note Γ_p le sous-groupe de G_p engendré par les deux matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Quel est l'ordre de Γ_p ?

18.3 Exemples de S_3 et S_4

1) Exemple de S_3

- a) On note τ la transposition (2 3) de S_3 et H le sous-groupe engendré par τ . Décrire les classes à gauche et à droite modulo H .
- b) Donner tous les sous-groupes de S_3 . Préciser s'ils sont distingués ou non. Lorsqu'ils le sont, calculer les quotients de S_3 par ces sous-groupes.

2) Exemple de S_4

- a) Dresser la liste des éléments de S_4 .
- b) Montrer que S_4 contient trois sous-groupes isomorphes à D_4 . Sont-ils distingués dans S_4 ?
- c) Soit V le sous-groupe engendré par les trois doubles transpositions. Montrer que V est un sous-groupe de A_4 , distingué dans A_4 et dans S_4 .
- d) Donner un système de représentants des classes de S_4 modulo V . À quels groupes S_4/V et A_4/V sont-ils isomorphes ?

18.4 Groupes cycliques

- 1) Les groupes suivants sont-ils cycliques : $\mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/30\mathbb{Z}$, $\mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/31\mathbb{Z}$?
- 2) Soient G un groupe monogène et $f : G \rightarrow G'$ un homomorphisme surjectif de groupes. Montrer que G' est un groupe monogène.

- 3)** Soient n un entier, $n \geq 1$. On note $D(n)$ l'ensemble des diviseurs de n .
- a)** Soit $d \in D(n)$. Montrer que le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ contient un unique sous-groupe d'ordre d et donner sa structure.
 - b)** Montrer que $a \bmod n$ est un générateur de $(\mathbb{Z}/n\mathbb{Z}, +)$ si et seulement si $\text{pgcd}(a, n) = 1$.
 - c)** Pour $d \geq 2$, on note $\varphi(d)$ le nombre d'entiers de $\{1, \dots, d-1\}$ premiers avec d ; on pose $\varphi(1) = 1$. Montrer que :

$$n = \sum_{d \in D(n)} \varphi(d).$$

Les propriétés de $\varphi(n)$, appelé indicateur d'Euler, seront étudiées dans l'exercice 21.1.

- d)** Appliquer les questions précédentes au cas $n = 15$.
- 4)** Soient $(G, +)$ un groupe cyclique, a et b deux éléments d'ordre $2p$ dans G , $p > 2$ premier. De quel ordre peut être $a + b$?

18.5 Exemple de $\text{GL}(2, \mathbb{R})$

On pose $G = \text{GL}(2, \mathbb{R})$.

- 1)** Montrer que l'ensemble $H = \{a \text{ id}, a \in \mathbb{R}\}$ des homothéties de $\mathbb{R}^2$ forme un sous-groupe distingué de G et que c'est le centre de G .
- 2)** Soit $S = \text{SL}(2, \mathbb{R}) = \{M \in G, \det(M) = 1\}$. Montrer que S est un groupe et qu'il est distingué dans G . Le groupe S est appelé groupe spécial linéaire.
- 3)** Soient G^+ (resp. G^-) l'ensemble des matrices de G à déterminant positif (resp. négatif). Les ensembles G^- et G^+ sont-ils des sous-groupes de G ? Sont-ils distingués ?
- 4)** Montrer que S est distingué dans G^+ . Trouver des groupes isomorphes à G/S et à G^+/S .

18.6 Sous-groupes distingués et groupes quotients

1) Sous-groupes d'indice deux

- a)** Montrer que, dans un groupe, tout sous-groupe d'indice 2 est distingué.
- b)** Que peut-on en déduire pour le sous-groupe alterné A_n de S_n ?

2) Quotient par le centre

- a)** Soit G un groupe de centre C . Montrer que, soit G est un groupe abélien, soit G/C n'est pas monogène.
- b)** En déduire que, si p est un nombre premier, un groupe d'ordre p^2 est abélien.

3) Sous-groupes engendrés. Soient G un groupe, H et K deux sous-groupes de G .

- a) Si L un sous-groupe de $H \cap K$ distingué dans H et dans K , montrer que L est distingué dans le sous-groupe de G engendré par $H \cup K$.
- b) Si H et K sont distingués dans G , montrer que le sous-groupe de G engendré par $H \cup K$ est distingué dans G .

4) Théorème du losange. Soient G un groupe, H un sous-groupe de G , K un sous-groupe distingué de G .

- a) Montrer que HK est un sous-groupe de G .
- b) On note $j : H \rightarrow HK$ l'injection, $\pi' : HK \rightarrow G/K$ la restriction de la projection canonique $G \rightarrow G/K$. Montrer que $H/(H \cap K)$ et HK/K sont isomorphes.

5) Soit G un groupe fini d'ordre n , $m < n$ un diviseur de n . Montrer que, s'il existe un unique sous-groupe H d'ordre m dans G , alors $H \triangleleft G$.

6) Soit $n > 2$ un entier. Existe-t-il un homomorphisme surjectif $S_n \rightarrow A_n$?

18.7 Automorphismes intérieurs d'un groupe

Soit G un groupe. On note $\text{Aut}(G)$ le groupe des automorphismes de G . Pour tout $a \in G$, on définit l'automorphisme intérieur associé à a , $\varphi(a) : G \rightarrow G$ par : $x \mapsto axa^{-1}$.

- 1) Vérifier que, pour tout a de G , $\varphi(a)$ est un automorphisme de G .
- 2) Montrer que φ ainsi défini est un homomorphisme de groupes de G dans $\text{Aut}(G)$.
- 3) On note $\text{Int}(G)$ l'ensemble des automorphismes intérieurs de G . Montrer que $\text{Int}(G)$ est un sous-groupe distingué de $\text{Aut}(G)$.
- 4) On note C le centre de G . Montrer que $\text{Int}(G)$ est isomorphe à G/C .

18.8 Commutateurs ; groupe dérivé d'un groupe

Soit G un groupe. On appelle commutateur d'un couple (a, b) d'éléments de G , l'élément, noté $[a, b]$, égal à $a^{-1}b^{-1}ab$. On note $D(G)$ le sous-groupe de G engendré par l'ensemble des commutateurs de G . On l'appelle groupe dérivé de G .

1) Montrer que, pour tout triplet (g, a, b) d'éléments de G , on a :

$$[gag^{-1}, bgb^{-1}] = g[a, b]g^{-1}.$$

En déduire que $D(G)$ est un sous-groupe distingué de G .

Quelle est la propriété de G équivalente à : $D(G) = \{e\}$?

2) Soit H un sous-groupe distingué de G . Montrer que H contient $D(G)$ si, et seulement si, G/H est abélien.

18.9 Décomposition en produit

Soient G un groupe, H et K des sous-groupes de G . On suppose que :

- H et K sont des sous-groupes distingués de G ;
- $H \cap K = \{e\}$;
- $H.K = G$.

On définit l'application $f: H \times K \rightarrow G$ par $f(h, k) = hk$.

- 1) Montrer que f est une application injective.
- 2) Montrer que si h est un élément de H et k un élément de K , h et k commutent entre eux.
- 3) Montrer que f est un isomorphisme de groupes.
- 4) On suppose que G est un groupe fini et on remplace l'hypothèse $H.K = G$ par $|H| \cdot |K| = |G|$. Montrer que f est encore un isomorphisme de groupes.
- 5) Soit p premier. Montrer que les seules structures de groupe d'ordre p^2 sont celles de $\mathbb{Z}/p^2\mathbb{Z}$ et de $(\mathbb{Z}/p\mathbb{Z})^2$ (utiliser l'exercice 18.6 2)).
- 6) a) Soit G un groupe d'ordre pq avec p et q premiers, $p > q$. Montrer que G possède un unique sous-groupe H d'ordre p et que celui-ci est distingué dans G (on pourra montrer que, sinon, G posséderait au moins p^2 éléments).
b) On reprend les notations précédentes et on suppose que q ne divise pas $p - 1$. Montrer que G est cyclique (choisir un élément a d'ordre p , un élément b d'ordre q ; montrer qu'il existe k tel que $bab^{-1} = a^k$ et calculer $b^q ab^{-q}$ de deux manières différentes).
c) Soit G un groupe d'ordre 77. Quelle sont les structures possibles pour G ?

Commentaire. C'est souvent le résultat du 4) qui est utile pour les groupes finis.

18.10 Classes de conjugaison dans S_n et A_n

Soit n un entier, $n \geq 1$. On appelle partition de n une suite finie u_σ décroissante d'entiers naturels > 0 de somme n . À toute permutation σ de S_n on associe une partition de n de la manière suivante : le nombre de termes de la suite est le nombre de cycles dans la décomposition de σ en produit de cycles disjoints (y compris les cycles de longueur 1), les termes de u_σ sont les longueurs de ces cycles rangées dans l'ordre décroissant.

Par exemple, la permutation $\sigma = (1 \ 12 \ 10 \ 4 \ 3 \ 7)(2 \ 5 \ 9 \ 6)$ de S_{12} définit la suite $u_\sigma = (6, 4, 1, 1)$.

- 1) Montrer que deux permutations σ et σ' de S_n sont conjuguées dans S_n si et seulement si : $u_\sigma = u_{\sigma'}$.
- 2) En déduire que le nombre de classes de conjugaison de S_n est égal à $p(n)$.
- 3) Calculer $p(n)$ pour $n = 1, \dots, 6$.

Commentaire. L'étude des partitions d'un entier fait l'objet du chapitre XVI de *l'Introduction à l'Analyse Infinitésimale* d'Euler publiée en 1748. Euler donne des relations permettant de calculer par récurrence les nombres $p(n)$... et bien d'autres.

18.11 Groupes opérant sur les ensembles

1) Exemples d'orbites. Soit G un sous-groupe de S_4 opérant sur $\{1, 2, 3, 4\}$ par l'action induite par l'action naturelle de S_4 . Pour $i = 1, 2, 3, 4$ on note O_i l'orbite de i et S_i le stabilisateur de i . Déterminer O_i et S_i pour $i = 1, 2, 3, 4$ dans chacun des cas suivants :

- a) $G = \text{Gr}((1\ 2\ 3))$;
- b) $G = \text{Gr}((1\ 2\ 3\ 4))$;
- c) G est le groupe engendré par les doubles transpositions ;
- d) $G = A_4$.

2) Exemples d'orbites dans $\mathbb{R}^2$. Soit G un sous-groupe de $\text{GL}(2, \mathbb{R})$. On fait agir G sur le plan affine euclidien P en choisissant un point O de cet espace et en identifiant $\mathbb{R}^2$ et les vecteurs d'origine O . Décrire l'orbite d'un point A de P quand G est le sous-groupe engendré par :

- a) une symétrie par rapport à une droite D passant par O ;
- b) une rotation d'angle $\pi/2$ de centre O ;
- c) une rotation d'angle $2\pi/n$ (n entier > 0) de centre O et une symétrie par rapport à une droite D passant par O .

3) Un groupe de 35 éléments agit sur un ensemble de 19 éléments sans fixer aucun d'eux. Combien y a-t-il d'orbites pour cette action ?

4) Comment présenter autrement la donnée d'un $\mathbb{Z}$ -ensemble ?

5) Montrer que $O^+(2, \mathbb{R})$ agit transitivement sur le cercle unité de $\mathbb{R}^2$ et que $O^+(3, \mathbb{R})$ agit transitivement sur la sphère unité de $\mathbb{R}^3$.

6) Montrer que l'action de $\text{GL}(n, \mathbb{R})$ sur $\mathbb{R}^n$ n'est pas une action transitive, mais qu'elle définit sur l'ensemble des bases de $\mathbb{R}^n$ une action transitive.

18.12 Cardinal de $GL(n, \mathbb{Z}/p\mathbb{Z})$ et de $SL(n, \mathbb{Z}/p\mathbb{Z})$

Soient n un entier ≥ 1 , p un nombre premier ; on note u_n le cardinal de $G_n = GL(n, \mathbb{Z}/p\mathbb{Z})$ et $(e_i)_{1 \leq i \leq n}$ la base canonique de $(\mathbb{Z}/p\mathbb{Z})^n$.

- 1) Déterminer le cardinal de l'orbite de e_1 pour l'action naturelle de G_n sur $(\mathbb{Z}/p\mathbb{Z})^n$.
- 2) Déterminer, en fonction de u_{n-1} , le cardinal du stabilisateur de e_1 . En déduire u_n et l'exposant N de la plus grande puissance de p divisant u_n . Retrouver ce résultat en comptant les bases de $(\mathbb{Z}/p\mathbb{Z})^n$.
- 3) Trouver un sous-groupe H de G_n d'ordre p^N (ce sous-groupe est appelé p -sous-groupe de Sylow de G).
- 4) Quel est le cardinal v_n de $SL(n, \mathbb{Z}/p\mathbb{Z})$?

18.13 Classes de conjugaison

Soit G un groupe fini. On fait agir G sur lui-même par conjugaison : $g.x = gxg^{-1}$. La classe de conjugaison de x est l'orbite de x pour cette action ; c'est l'ensemble des conjugués de x dans G . Cet exercice utilise au 4) la notion de partition d'un entier définie dans l'exercice 18.10.

- 1) Montrer que deux éléments d'une même classe de conjugaison de G ont même ordre.
- 2) Former l'équation des classes lorsque G est abélien.
- 3) On suppose que G est le groupe diédral D_n ; on rappelle (exercice 18.1) que D_n est engendré par deux éléments r et s tels que $r^n = s^2 = e$ et $sr^k = r^{-k}s$.
 - a) Quels sont les conjugués des éléments de la forme r^k ou sr^k suivant les valeurs de k et la parité de n ?
 - b) Déterminer les classes de conjugaison et former l'équation des classes pour D_3 , D_4 , D_5 et D_6 .
- 4) Déterminer les classes de conjugaison et former l'équation des classes pour S_3 , S_4 , S_5 .
- 5) Déterminer les classes de conjugaison et former l'équation des classes pour A_4 , A_5 . En déduire que A_5 est simple.

18.14 Détermination des sous-groupes finis de $O^+(3, \mathbb{R})$

On note $O^+(3, \mathbb{R})$ le sous-groupe de $GL(3, \mathbb{R})$ dont les éléments sont les transformations orthogonales de déterminant 1, c'est-à-dire les rotations de $\mathbb{R}^3$. On note S la sphère unité de $\mathbb{R}^3$.

Soit G un sous-groupe fini d'ordre $n > 1$ de $O^+(3, \mathbb{R})$. On définit les notations suivantes.

- Pour tout g de $G \setminus \{\text{id}\}$, $F(g) = \{x \in S, g(x) = x\}$, (g étant une rotation, ses points fixes sont deux points diamétralement opposés de S).
- $X = \cup_{g \in G \setminus \{\text{id}\}} F(g)$, (X est donc l'ensemble des points fixes de tous les éléments de G autres que l'identité).
- $\Gamma = \{(g, x), g \in G \setminus \{\text{id}\}, x \in S, g(x) = x\}$.

On considère l'action naturelle de G sur X . Soient $O_1, \dots, O_k$ les différentes orbites dans X pour cette action, $n_1, \dots, n_k$ l'ordre des stabilisateurs des éléments de chacune de ces orbites (l'ordre ne dépend pas de l'élément de l'orbite considéré).

1) Montrer que : $|\Gamma| = 2n - 2$.

2) Montrer que : $|\Gamma| = \sum_{1 \leq i \leq k} (n_i - 1) \frac{n}{n_i}$; en déduire que les entiers n et k vérifient l'équation :

$$2 - \frac{2}{n} = \sum_{1 \leq i \leq k} \left(1 - \frac{1}{n_i}\right) \quad (E)$$

3) Montrer que k ne peut être égal qu'à 2 ou 3 ; trouver les différentes valeurs possibles *a priori* de n dans chaque cas.

4) Montrer qu'à chaque solution de l'équation précédente correspond une classe de sous-groupes de $O^+(3, \mathbb{R})$ qu'on précisera.

18.15 Groupe du tétraèdre régulier

Soient T un tétraèdre régulier de $\mathbb{R}^3$, de centre O , $E = \{A, B, C, D\}$ l'ensemble des sommets de T , G le groupe des isométries affines de $\mathbb{R}^3$ laissant E invariant et H le sous-groupe de G formé des isométries positives.

1) On définit un homomorphisme $\varphi : G \rightarrow S_4$ en associant à un élément de G la permutation de E qu'il induit. Montrer que φ est injective puis que φ est surjective (se rappeler que S_4 est engendré par les transpositions).

2) Montrer que $\varphi' = \varphi|_H$ définit un isomorphisme de H sur A_4 .

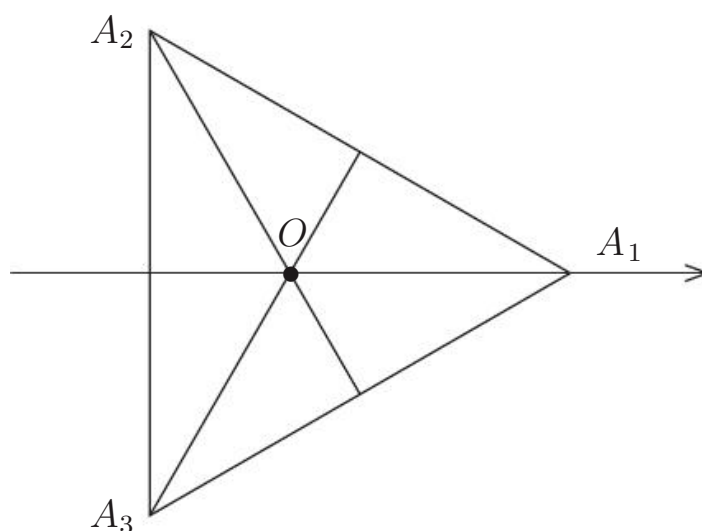
Solutions

18.1 1) Il est clair que le composé d'isométries de D_n est une isométrie de D_n et que l'inverse d'une isométrie de D_n est une isométrie de D_n ; par conséquent, D_n est un sous-groupe du groupe des isométries du plan.

Notons $A_1, \dots, A_{n-1}$ les sommets du polygone et O son centre. L'image de O , isobarycentre des n sommets, est O lui-même. L'image de A_1 peut-être n'importe lequel des n sommets ; si l'image de A_1 est A_k , l'image de A_2 peut-être A_{k+1} ou A_{k-1} ; comme l'image d'un repère définit une isométrie, ce sont les seules possibilités ; on vérifie qu'elles peuvent être réalisées, la première avec la rotation R amenant A_1 en A_k , la seconde avec le composé $R \circ S$ de R avec la symétrie orthogonale S par rapport à OA_1 . On obtient donc $2n$ éléments.

2) Notons A_1, A_2, A_3 les sommets d'un triangle équilatéral.

Toute isométrie f de D_3 définit une permutation $\varphi(f)$ de l'ensemble $\{1, 2, 3\}$; on définit donc une application $\varphi : D_3 \rightarrow S_3$ où $\varphi(f)(i)$ est l'indice du sommet $f(A_i)$. Il est clair que φ est un homomorphisme injectif ; comme sa source et son but ont même cardinal, c'est un isomorphisme. On peut remarquer que les transpositions correspondent à des symétries et les deux 3-cycles à des rotations de $\pm 2\pi/3$. Une symétrie et une rotation engendrent D_3 et vérifient les relations indiquées.

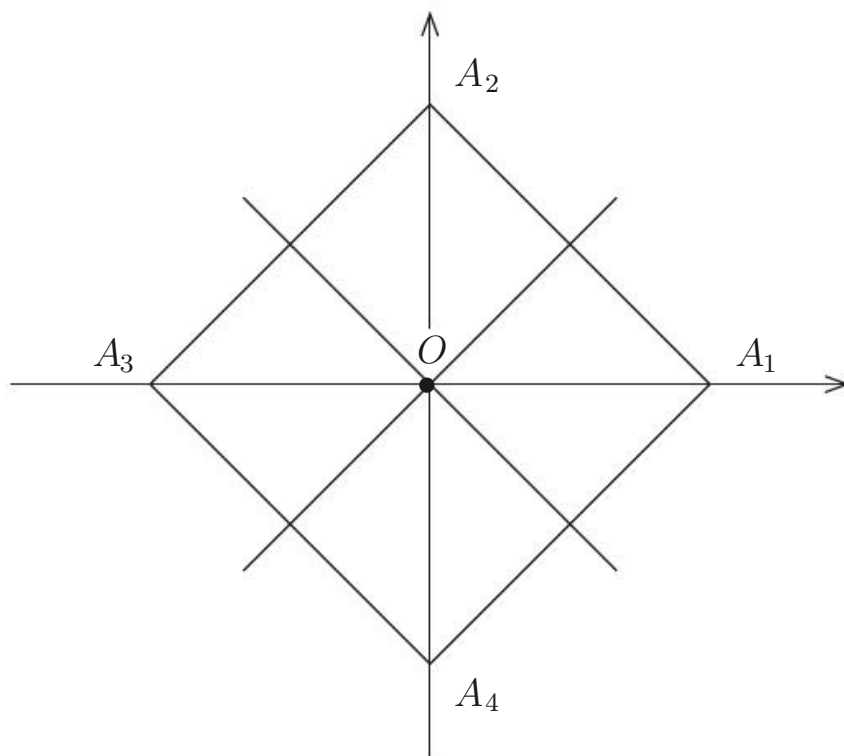


3) Les 8 éléments de D_4 sont les quatre rotations d'angles $0, \pi/2, \pi$ et $3\pi/2$, et les quatre symétries ayant pour axe une diagonale ou une médiane du carré.

Notons r la rotation d'angle $\pi/2$ et s la symétrie par rapport à l'une des diagonales. Alors les symétries par rapport aux médianes sont rs et r^3s et la symétrie par rapport à l'autre diagonale est r^2s .

On a donc $D_4 = \{\text{id}, r, r^2, r^3, s, rs, r^2s, r^3s\}$; on voit que r et s engendrent D_4 et on vérifie que $sr = r^3s = r^{-1}s$, ce qui permet d'écrire les produits d'éléments de D_4 sous l'une des huit formes indiquées.

4) Notons $A_1, A_2 = r(A_1), A_3 = r(A_2), A_4 = r(A_3)$ les sommets du carré et définissons $\varphi : D_4 \rightarrow S_4$ de manière analogue au **2)**.



Les huit éléments de G sont :

$\text{id}, \varphi(r) = (1\ 2\ 3\ 4), \varphi(r^2) = (1\ 3)(2\ 4), \varphi(r^3) = (1\ 4\ 3\ 2), \varphi(s) = (2\ 4),$
 $\varphi(rs) = (1\ 2)(3\ 4), \varphi(r^2s) = (1\ 3), \varphi(r^3s) = (1\ 4)(2\ 3).$

Une autre numérotation des sommets du carré peut donner un autre sous-groupe de S_4 (voir exercice 18.3 2b)).

5) Les sous-groupes de D_4 doivent avoir comme ordre un diviseur de 8 ; il n'est pas sûr, *a priori*, qu'il en existe pour chacun des diviseurs de 8.

- Sous-groupe à 1 élément : $\{\text{id}\}$, distingué dans D_4 ; $D_4/\{\text{id}\} = D_4$.
- Sous-groupes à 2 éléments : on cherche les éléments d'ordre 2 de D_4 et on trouve les 4 symétries et le demi-tour r^2 ; seul $r^2 = -\text{id}$ engendre un sous-groupe H distingué de D_4 , car on a $r^{-1}sr = r^{-2}s \notin \text{Gr}(s)$, etc. Le quotient D_4/H a 4 éléments d'ordre 2, les classes de id, r, s, sr ; il est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.
- Sous-groupes à 4 éléments : on en trouve 3 : $\text{Gr}(r)$ et les groupes engendrés par les symétries par rapport à deux droites orthogonales : $\{\text{id}, s, r^2, sr^2\}$ et $\{\text{id}, sr, r^2, sr^3\}$. Étant d'indice 2 dans D_4 , ils sont distingués dans D_4 (voir exercice 18.6

1a) pour un argument général, mais on peut aussi faire des vérifications par des calculs). Les quotients sont d'ordre 2, donc isomorphes à $\frac{\mathbb{Z}}{2\mathbb{Z}}$.

- Sous-groupe à 8 éléments : D_4 , distingué dans D_4 ; $D_4/D_4 = \{e\}$.

6) Prendre les groupes $H = \text{Gr}(s)$ et $K = \{\text{id}, s, r^2, sr^2\}$.

7) On prend pour r la rotation de centre O et d'angle $2\pi/n$ et pour s la symétrie par rapport à (OA_1) . Les relations $r^n = s^2 = e$ sont évidentes. Pour établir la troisième relation, il suffit de montrer l'égalité de sr^k et $r^{-k}s$ en A_1 et en A_2 puisque (O, A_1, A_2) est un repère affine : on vérifie que $sr^k(A_1) = A_{n-k} = r^{-k}s(A_1)$ et que $sr^k(A_2) = A_{n-k-1} = r^{-k}s(A_2)$. Les éléments de D_n peuvent s'écrire r^k ou sr^k , $0 \leq k \leq n-1$.

18.2 1) Vérifications sans difficulté.

2) Pour tous $a, b \in G$, on a $aabb = a^2b^2 = e = (ab)^2 = abab$; on en déduit $ab = ba$.

3) À ce stade de leur lecture, ceux et celles qui ont ce livre sous les yeux connaissent probablement différentes structures de groupe, mais n'ont peut-être jamais fait la démarche de se demander s'ils connaissaient toutes les structures de groupes d'un ordre donné.

- Groupes à p éléments, p premier. L'ordre d'un élément divisant l'ordre du groupe, un élément x autre que l'élément neutre a pour ordre p et engendre le groupe ; la seule structure de groupe à p éléments quand p est premier est donc celle du groupe $\frac{\mathbb{Z}}{p\mathbb{Z}}$ (avec la loi $+$) ; ceci est vrai en particulier pour $p = 2, 3, 5, 7, 11$.

Pour tout n , il existe une structure de groupe cyclique d'ordre n , celle de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ (avec la loi $+$). Il s'agit maintenant d'en chercher d'autres si n n'est pas premier.

- Groupes à 4 éléments : $\frac{\mathbb{Z}}{4\mathbb{Z}}$ et $\left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^2$. Si un groupe G a un élément d'ordre 4, il est cyclique et isomorphe à $\frac{\mathbb{Z}}{4\mathbb{Z}}$. Sinon, G a trois éléments a, b, c d'ordre 2 et est abélien d'après le 2). On ne peut avoir que $ab = c, bc = a, ca = b$ et la structure de G est celle de $\left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^2$.
- Groupes à 6 éléments : $\frac{\mathbb{Z}}{6\mathbb{Z}}$ et S_3 . Soit G un groupe d'ordre 6. Si G n'est pas cyclique, il a un élément a d'ordre 2 et un élément b d'ordre 3 d'après le théorème de Cauchy. On a $ab \neq ba$, sinon ab est un élément d'ordre 6, ce qui a été exclu. On en déduit que les éléments de G sont nécessairement e, b, b^2, a, ab, ab^2 et que

$ba = ab^2$; cette unique possibilité donne bien un groupe : on connaît S_3 , c'est lui ! On pourrait aussi penser au groupe à 6 éléments D_3 , mais ce qui précède montre qu'il est sûrement isomorphe à S_3 (on a construit un isomorphisme de D_3 sur S_3 dans l'exercice 18.1 2)).

- Groupes à 8 éléments : $\left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^3$, $\frac{\mathbb{Z}}{4\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}}$, $\frac{\mathbb{Z}}{8\mathbb{Z}}$, et D_4 ; les ordres des éléments du premier sont 1 ou 2, ceux du second 1, 2 ou 4, ceux du troisième 1, 2, 4, 8 ; ils ne sont donc pas isomorphes entre eux, ni à D_4 qui n'est pas abélien. Il existe encore une autre structure de groupe à 8 éléments, celle du groupe quaternionique Q_2 formé par les huit matrices $\pm 1, \pm i, \pm j, \pm k$ (voir 23.10) : on a $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$; Q_2 a 6 éléments d'ordre 4 et n'est pas commutatif. On peut montrer qu'il n'existe pas d'autre structure de groupe à 8 éléments.
- Groupes à 9 éléments : $\left(\frac{\mathbb{Z}}{3\mathbb{Z}}\right)^2$, $\frac{\mathbb{Z}}{9\mathbb{Z}}$.
- Groupes à 10 éléments : $\frac{\mathbb{Z}}{10\mathbb{Z}} \simeq \frac{\mathbb{Z}}{5\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}}$ et D_5
- Groupes à 12 éléments : $\left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^2 \times \frac{\mathbb{Z}}{3\mathbb{Z}}$, $\frac{\mathbb{Z}}{12\mathbb{Z}}$, A_4 et D_6 sont connus. Les deux derniers ne sont pas isomorphes, par exemple parce que D_6 a des éléments d'ordre 6 et que A_4 n'en a pas. On pourrait aussi penser à $\frac{\mathbb{Z}}{6\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}}$, mais l'isomorphisme du théorème chinois (voir 18.5) montre que $\frac{\mathbb{Z}}{6\mathbb{Z}} \simeq \frac{\mathbb{Z}}{3\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}}$, d'où $\frac{\mathbb{Z}}{6\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \simeq \frac{\mathbb{Z}}{3\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \simeq \left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^2 \times \frac{\mathbb{Z}}{3\mathbb{Z}}$.

Les groupes cités ne sont pas isomorphes entre eux : cela se voit par des arguments de cardinaux ou d'ordre d'éléments, le fait d'être abélien ou non. Savoir s'il existe d'autres structures de groupe à 9, 10, 12 éléments que celles que nous avons évoquées n'est pas évident ; en fait, il existe encore une autre structure de groupe à 12 éléments.

4) On voit d'abord que $A^k = \begin{pmatrix} 1 & 0 \\ 0 & 2^k \end{pmatrix}$, donc A engendre un sous-groupe d'ordre

égal à l'ordre t de 2 dans le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$ (sur l'ordre de 2 dans ce groupe, voir le commentaire de 21.1.e). On voit aussi que $B^2 = I_2$ et que

$$BA^k B = \begin{pmatrix} 2^k & 0 \\ 0 & 1 \end{pmatrix}. \text{ En continuant ces calculs, on s'aperçoit que les éléments de } \Gamma_p$$

sont de l'une des deux formes suivantes $\begin{pmatrix} 2^k & 0 \\ 0 & 2^l \end{pmatrix}$ ou $\begin{pmatrix} 0 & 2^k \\ 2^l & 0 \end{pmatrix}$. L'ordre de Γ_p est donc 2^{ℓ} . Voici un tableau des ordres des Γ_p pour de petites valeurs de p .

3	5	7	11	13	17	19	23	29	31	37	41	127
8	32	18	200	288	128	648	242	1568	50	2592	800	98

18.3 1) Posons $\sigma = (1\ 2\ 3)$; on a $\sigma^2 = (1\ 3\ 2)$.

a) Chaque classe compte deux éléments comme H . Les classes à gauche modulo H sont $H, \sigma H = \{\sigma, (1\ 2)\}$ et $\sigma^2 H = \{\sigma^2, (1\ 3)\}$. Les classes à droite modulo $H, H\sigma = \{\sigma, (1\ 3)\}$ et $H\sigma^2 = \{\sigma^2, (1\ 2)\}$.

b) Les classes à gauche et à droite modulo H ne sont pas égales, ce qui montre que H n'est pas un sous-groupe distingué de S_3 . Il en est de même pour les deux autres sous-groupes de S_3 engendrés par les transpositions $(1\ 2)$ et $(1\ 3)$.

Les autres sous-groupes de S_3 : e, S_3 et $\text{Gr}(\sigma) = \{e, \sigma, \sigma^2\}$ sont distingués, les deux premiers trivialement, le dernier, qui est égal à A_3 , soit en remarquant que le conjugué d'un 3-cycle est un 3-cycle (d'après l'exercice **11.3**), soit en remarquant que A_3 est un sous-groupe d'indice 2 de S_3 et en appliquant le résultat de l'exercice **18.6 1a**.

2) a) On trouve id, six transpositions, trois doubles transpositions, huit 3-cycles, six 4-cycles, soit en tout vingt quatre permutations.

b) En numérotant les sommets dans l'exercice **18.1 5)** avec les suites d'indices 2 1 3 4 et 4 2 3 1, on obtient deux autres plongements de D_4 dans S_4 . Aucune des images de D_4 n'est distinguée dans S_4 , puisque, comme les transpositions sont conjuguées dans S_4 (en utilisant la formule de l'exercice 11.3), un sous-groupe distingué de S_4 qui contient une transposition les contient toutes et est donc égal à S_4 .

c) Il suffit de noter que le conjugué d'une double transposition est une double transposition (l'exercice 11.3 est toujours utile).

d) On peut considérer le groupe S_3 comme un sous-groupe de S_4 , celui des permutations de S_4 laissant 4 fixe. On voit alors que, si σ et σ' sont deux permutations distinctes de S_3 , elles ne sont pas égales modulo V car $\sigma' \circ \sigma^{-1}$, qui fixe 4 sans être l'identité, n'est pas dans V . La projection $S_4 \rightarrow S_4/V$ se restreint donc à un isomorphisme $S_3 \rightarrow S_4/V$. Si on restreint ce dernier homomorphisme à A_3 , on obtient un isomorphisme $A_3 \rightarrow A_4/V$; comme $A_3 \simeq \mathbb{Z}/3\mathbb{Z}$, on peut aussi dire que $A_4/V \simeq \mathbb{Z}/3\mathbb{Z}$.

18.4 1) L'ordre de $\mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/30\mathbb{Z}$ est $15 \times 30 = 450$; comme un élément (x, y) de ce groupe a pour ordre maximal 30, car $30(x, y) = (30x, 30y) = (0, 0)$, le groupe n'est pas cyclique. Par contre, le second groupe est cyclique : comme 15 et 31 sont premiers entre eux, il est isomorphe à $\mathbb{Z}/465\mathbb{Z}$, d'après le théorème chinois (voir **18.5**).

2) Si G est monogène, il existe un élément $x \in G$ qui engendre G ; $\text{im}(f)$, qui est égal à G' , est engendré par $f(x)$, donc G' est monogène.

On peut également dire que, si G est monogène, il existe un homomorphisme surjectif $g : \mathbb{Z} \rightarrow G$; le composé $f \circ g : \mathbb{Z} \rightarrow G'$ est surjectif, ce qui prouve que G' est monogène.

3) a) Posons $n = ud$. Comme $u \bmod n$ est d'ordre d dans $(\mathbb{Z}/n\mathbb{Z})$, le groupe $H = u\mathbb{Z}/n\mathbb{Z}$ est un sous-groupe cyclique d'ordre d de $(\mathbb{Z}/n\mathbb{Z}, +)$, engendré par $u \bmod n$ et isomorphe à $(\mathbb{Z}/d\mathbb{Z}, +)$. Si K est un sous-groupe d'ordre d de $\mathbb{Z}/n\mathbb{Z}$, il est cyclique. Soit v un générateur de K . On a $dv = 0 \bmod n$, donc $dv = \lambda ud$, avec λ entier, ce qui prouve que u divise v . Par conséquent, $K \subset H$. D'où $K = H$.

b) Si $\text{pgcd}(a, n) = d > 1$, il existe des entiers $u, v < n$ tels que $a = ud$ et $n = vd$. On a donc $va = vud = nu = 0 \bmod n$, ce qui prouve que l'ordre de $a \bmod n$ est strictement inférieur à n . Si $\text{pgcd}(a, n) = 1$ et si k est l'ordre de a dans $(\mathbb{Z}/n\mathbb{Z}, +)$, l'égalité $ka = 0 \bmod n$ montre que n divise ka ; comme n est premier avec a , il divise k , ce qui prouve que $k = n$ et que a engendre $(\mathbb{Z}/n\mathbb{Z}, +)$.

c) Considérons la fonction $w : \{0, \dots, n-1\} \rightarrow D(n)$ qui associe à chaque classe $k \bmod n$ son ordre dans $(\mathbb{Z}/n\mathbb{Z}, +)$. Pour $d \in D(n)$, $w^{-1}(d)$ est l'ensemble des générateurs de l'unique sous-groupe d'ordre d de $(\mathbb{Z}/n\mathbb{Z}, +)$; son cardinal est $\varphi(d)$ d'après ce qui précède. Comme les $w^{-1}(d)$ forment une partition de $\mathbb{Z}/n\mathbb{Z}$, on obtient la formule demandée.

d) Pour $n = 15 = 3 \times 5$, on trouve comme générateurs de $\mathbb{Z}/15\mathbb{Z}$ les classes de 1, 2, 4, 7, 8, 11, 13, 14 et $\varphi(15) = 8$. Les éléments d'ordre 5 sont les classes de 3, 6, 9, 12 et les éléments d'ordre 3 sont les classes de 5 et 10. Comme $\varphi(1) = 1$, $\varphi(3) = 2$, $\varphi(5) = 4$, on a bien $15 = \sum_{d \in D(15)} \varphi(d)$.

4) D'après **3) a)**, si G contient des éléments d'ordre $2p$, G contient un unique sous-groupe d'ordre $2p$, isomorphe à $(\mathbb{Z}/2p\mathbb{Z}, +)$, dans lequel on peut se placer. L'ordre de $a + b$ peut être *a priori* 1, 2, p ou $2p$. Il peut être 1 si $b = -a$, p si $a = b$. Les éléments d'ordre 2 ou $2p$ sont des classes d'entiers impairs donc $a + b$ est la classe d'un entier pair et son ordre n'est ni 2 ni $2p$.

18.5) Cet exercice utilise la propriété $\det(g \circ f) = \det(g)\det(f)$ (14.6 proposition 5).

1) Voir 11.9.

2) La propriété rappelée montre que le déterminant est un homomorphisme $G \rightarrow (\mathbb{R})^*$; son noyau S est donc un sous-groupe distingué de G .

3) La propriété rappelée montre que G^- n'est pas stable par multiplication (le produit de deux matrices dont les déterminants sont strictement négatifs a un déterminant positif) et que G^+ est un sous-groupe distingué de G (si $\det(f) > 0$, alors $\det(g \circ f \circ g^{-1}) = \det(f) > 0$).

4) Comme S est un sous-groupe distingué de G et que $S \subset G^+ \subset G$, S est un sous-groupe distingué de G^+ .

Comme $\det : G \rightarrow (\mathbb{R}^*, \cdot)$ est surjectif de noyau S , on a $G/S \simeq (\mathbb{R}^*, \cdot)$. De même, $G^+/S \simeq (\mathbb{R}^{+*}, \cdot)$.

18.6 1) a) Soit H un sous-groupe d'indice 2 d'un groupe G . Soit $x \in G$. Si $x \in H$, on a $xH = H = Hx$. Si $x \notin H$, comme H est d'indice 2 dans G , la classe à gauche xH de x , qui est distincte de H , est le complémentaire de H dans G et, de même, la classe à droite Hx de x est le complémentaire de H dans G ; donc $xH = Hx$. Comme les classes à gauche et à droite sont égales dans tous les cas, H est un sous-groupe distingué de G .

b) Le sous-groupe alterné A_n , étant d'indice 2 dans S_n est un sous-groupe distingué de S_n .

On peut aussi dire que A_n est le noyau de l'homomorphisme de signature $S_n \rightarrow \{0, 1\}$ ou remarquer que le conjugué d'une transposition est une transposition (exercice **11.3**) et raisonner directement : un élément de A_n est produit d'un nombre pair de transpositions, son conjugué par une permutation quelconque sera le produit des conjugués de ces transpositions et sera donc une permutation paire.

2) a) Notons $\pi : G \rightarrow G/C$ la projection canonique. Supposons que G ne soit pas abélien (donc $C \neq G$) et que G/C soit monogène. Notons $\alpha = \pi(g)$ un générateur de G/C et x, x' deux éléments de G ne commutant pas. Il existe des entiers k et k' tels que $\pi(x) = \alpha^k$ et $\pi(x') = \alpha^{k'}$, donc il existe $c, c' \in C$ tels que $x = cg^k$ et $x' = c'g^{k'}$; on a alors

$$xx' = cg^k c'g^{k'} = cc'g^k g^{k'} = c'cg^{k'} g^k = c'g^{k'} cg^k = x'x,$$

ce qui donne une contradiction.

b) Si G est un groupe d'ordre p^2 , un sous-groupe de G a pour ordre 1, p ou p^2 . D'après le théorème de Burnside, comme le centre C de G n'est pas réduit à e_G , son ordre est p ou p^2 . Si c'est p^2 , alors $C = G$ et G est commutatif. Si c'est p , comme C est distingué dans G , G/C est un groupe d'ordre p donc cyclique contrairement au résultat du **a**).

3) a) Un élément x de $\text{Gr}(H \cup K)$ est de la forme $h_1 k_1 \dots h_r k_r$ où r est un entier et $h_i \in H, k_i \in K$ pour $i = 1 \dots r$. Par récurrence sur r , on vérifie que $xLx^{-1} \subset L$, d'où

le résultat. On peut aussi vérifier que $N = \{x, xLx^{-1} \subset L\}$ est un sous-groupe de G qui contient $H \cup K$ donc $\text{Gr}(H \cup K)$.

b) Pour tout $x \in G$, le produit

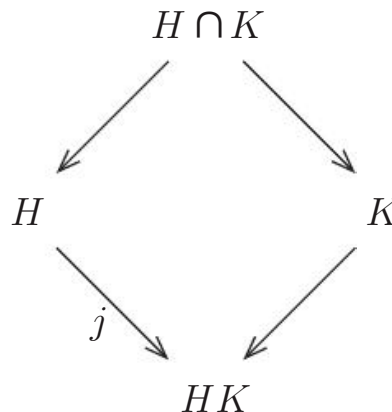
$$x(h_1 k_1 \dots h_r k_r)x^{-1} = (xh_1 x^{-1})(xk_1 x^{-1}) \dots (xh_r x^{-1})(xk_r x^{-1})$$

est un produit fini d'éléments de H et d'éléments de K , d'où le résultat.

4) a) Si $h, h' \in H$ et $k, k' \in K$, $(hk)(h'k') = hh'(h'^{-1}kh')k'$ est un élément de HK car, comme K est un sous-groupe distingué de G , $h'^{-1}kh' \in K$. On procède de manière analogue pour montrer que HK est stable par inversion : $(hk)^{-1} = k^{-1}h^{-1} = h^{-1}hk^{-1}h^{-1}$, etc.

b) L'isomorphisme classique donne $\text{im}(\pi' \circ j) \simeq H/\ker(\pi' \circ j)$. On vérifie que $\ker(\pi' \circ j) = H \cap K$ et que $\pi'(j(H)) = HK/K$, d'où le résultat. On peut voir que cet isomorphisme se décrit sur les classes par la simple formule $h \bmod H \cap K \mapsto h \bmod K$, le h étant considéré comme élément de h au départ, de HK à l'arrivée.

Le nom de losange vient de la forme du diagramme :



5) Tout conjugué de H , étant un sous-groupe d'ordre m de G , est égal à H , donc $H \triangleleft G$.

6) Si un tel homomorphisme existe, son noyau est un sous-groupe distingué H de S_n d'ordre 2. Le sous-groupe H est engendré par un élément σ d'ordre 2 de S_n ; la décomposition de σ en produit de cycles disjoints montre que σ est un produit de transpositions à supports disjoints et, comme $n > 2$, il est facile de construire un conjugué de σ différent de σ . On obtient une contradiction.

18.7 1) Pour tous $a \in G$, $x, x' \in G$, on a

$$\varphi(a)(xx') = a(xx')a^{-1} = (axa^{-1})(ax'a^{-1}) = \varphi(a)(x)\varphi(a)(x'),$$

ce qui prouve que $\varphi(a)$ est un homomorphisme. L'inverse de $\varphi(a)$ est $\varphi(a^{-1})$.

2) Pour tous $a, b \in G$ et $x \in G$, $\varphi(ab)(x) = (ab)x(ab)^{-1} = a(bxb^{-1})a^{-1} = [\varphi(a) \circ \varphi(b)](x)$, donc $\varphi(ab) = \varphi(a) \circ \varphi(b)$.

3) Pour tout $\psi \in \text{Aut}(G)$ et tous $a, x \in G$, on a

$$[\psi \circ \varphi(a) \circ \psi^{-1}](x) = \psi(a\psi^{-1}(x)a^{-1}) = \psi(a)x(\psi(a))^{-1} = [\varphi(\psi(a))](x);$$

donc $\psi \circ \varphi(a) \circ \psi^{-1} \in \text{Int}(G)$ et la conclusion.

4) L'homomorphisme $\varphi : G \rightarrow \text{Aut}(G)$ a pour image $\text{Int}(G)$; on voit facilement que son noyau est C , d'où le résultat.

18.8 1) Les calculs pour montrer l'égalité sont simples. On peut aussi vérifier que $[a, b]^{-1} = [b, a]$. Un élément de $D(G)$ est donc un produit fini de commutateurs et tout conjugué d'un produit fini de commutateurs est un produit de commutateurs. Par conséquent, $D(G)$ est un sous-groupe distingué de G .

L'égalité $D(G) = \{e\}$ a lieu si et seulement si, pour tous $a, b \in G$, $a^{-1}b^{-1}ab = e$, ce qui équivaut à $ab = ba$ ou à la commutativité de G .

2) Notons $\pi : G \rightarrow G/H$ la projection canonique. L'image par π d'un commutateur de G est un commutateur de G/H . La commutativité de G/H équivaut à ce que tout commutateur de G/H soit égal à $\pi(e)$. Si l'image par π de tout commutateur de G est $\pi(e)$, c'est que $D(G) \subset H$; la réciproque est évidente.

18.9 1) Si $f(h, k) = f(h', k')$ on a $hk = h'k'$, donc $h'^{-1}h = k'k^{-1}$. Le membre de gauche de cette égalité est dans H , le membre de droite est dans K . Comme $H \cap K = \{e\}$, on a $h = h'$ et $k = k'$.

2) On remarque que $x = hkh^{-1}k^{-1}$ est le produit de h par $kh^{-1}k^{-1}$, conjugué d'un élément de H qui est dans H puisque H est un sous-groupe distingué de G ; x est donc un élément de H . De même, x est un élément de K en l'écrivant comme produit de hkh^{-1} et de k^{-1} . C'est donc un élément de $H \cap K = \{e\}$; on en déduit $hkh^{-1}k^{-1} = e$, donc $hk = kh$.

3) Le **2)** permet de montrer que f est un homomorphisme de groupes, le **1)** montre que f est injectif, la troisième condition qu'il est surjectif.

4) Comme précédemment, f est un homomorphisme injectif ; l'égalité des cardinaux montre qu'il est surjectif.

5) L'exercice **18.6 2)** montre qu'un groupe G d'ordre p^2 est abélien. Tout sous-groupe de G est donc distingué dans G . Si G a un élément d'ordre p^2 , G est cyclique et isomorphe à $\mathbb{Z}/p^2\mathbb{Z}$. Sinon, tout élément non nul est d'ordre p . On choisit a d'ordre p engendrant un sous-groupe H et $b \notin H$ engendrant un sous-groupe K ; on a $H \cap K = \{e\}$, car si $x \neq e$ appartient à cette intersection, c'est un générateur de H et de K , donc $H = K$; on peut alors appliquer le **4)**. On a donc $G \simeq H \times K \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

6) a) Le théorème de Cauchy montre que G possède au moins un sous-groupe d'ordre p . Supposons que G en possède deux, H et K , engendrés par a et b respectivement. Alors $H \cap K = \{e\}$ (même raisonnement qu'au **5**), donc G possède les p^2 éléments $a^k b^l$, $0 \leq k, l < p$; en effet, ces éléments sont tous distincts car si $a^k b^l = a^r b^s$, on a $a^{k-r} = b^{s-l}$ et cet élément de $H \cap K$ est nécessairement l'élément neutre, d'où $k = r$ et $l = s$. Comme $p^2 > pq$, ceci est impossible. Mais si G possède un seul sous-groupe d'ordre p , celui-ci est distingué, car les conjugués de G , qui sont des sous-groupes d'ordre p lui sont tous égaux (ce dernier résultat a été vu dans l'exercice **18.6 5**).

b) Si a est d'ordre p , il engendre l'unique sous-groupe H d'ordre p de G . Comme bab^{-1} est d'ordre p , on a donc $bab^{-1} \in H - \{e\}$, donc il existe k , $1 \leq k < p$ tel que $bab^{-1} = a^k$. On a, d'une part, $b^q = e$ donc $b^q ab^{-q} = a$; d'autre part, $b^s ab^{-s} = a^{k^s}$ pour tout entier s (par récurrence), donc $b^q ab^{-q} = a^{k^q}$. On en déduit $a^{k^q - 1} = e$, ce qui montre que $k^q - 1 \equiv 0 \pmod{p}$, ce qui impose $k = 1$ puisque q ne divise pas $p - 1$. On a donc $ab = ba$; G est un groupe cyclique engendré par ab .

c) Il suffit d'appliquer ce qui précède pour montrer que $G \simeq \mathbb{Z}/77\mathbb{Z}$.

18.10 1) Soit $\rho \in S_n$. Si $\sigma = c_1 \dots c_r$ est la décomposition de σ en produit de cycles disjoints et si $\sigma' = \rho \sigma \rho^{-1}$, on a $\sigma' = \rho c_1 \rho^{-1} \dots \rho c_r \rho^{-1}$; l'exercice **11.3** montre que $\rho c_i \rho^{-1}$ est, pour $i = 1, \dots, r$, un cycle de même longueur que c_i , d'où $u\sigma = u\sigma'$.

Réciproquement, si $u\sigma = u\sigma'$, on définit, en écrivant l'une au-dessous de l'autre les décompositions des deux permutations σ et σ' en produit de cycles disjoints classés

n	1	2	3	4	5	6
$p(n)$	1	2	3	5	7	11

dans l'ordre décroissant de leurs longueurs, une permutation $\rho : x_i' \mapsto x_i$ telle que $\sigma' = \rho \sigma \rho^{-1}$, d'après l'exercice **11.3**.

2) Conséquence immédiate du **1**).

3) On trouve les résultats suivants.

Par exemple, $p(4) = 5$ car : $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$; les différents cas correspondent respectivement à un 4-cycle, un 3-cycle, une double transposition, une transposition et l'identité de S_4 .

18.11 1) Pour la question **a)**, on trouve $O_1 = O_2 = O_3 = \{1, 2, 3\}$, $O_4 = \{4\}$, $S_1 = S_2 = S_3 = \{\text{id}\}$, $S_4 = G$. Pour les questions **b**, **c**, **d**), on trouve $O_i = \{1, 2, 3, 4\}$ pour $i = 1, \dots, 4$, $S_i = \{\text{id}\}$ pour $i = 1, \dots, 4$.

- 2) a)** L'orbite est formée de deux points en général, de A seul si $A \in D$.
- b)** L'orbite est formée des quatre sommets d'un carré de centre O en général, de A seul si $A = O$.
- c)** L'orbite est formée des n sommets d'un polygone régulier de centre O et de leurs n symétriques par rapport à la droite D ; quand la droite passe par un des sommets du polygone ou est médiatrice d'un côté du polygone, l'orbite ne compte que n points ; dans le cas du point O , l'orbite est réduite à un point.
- 3)** Le cardinal d'une orbite est un diviseur de 35 ; ce n'est ni 1 ni 35, donc c'est 5 ou 7. Notons m et n les nombres d'orbites à 5 et 7 éléments ; on doit avoir : $19 = 5m + 7n$; la seule possibilité est $m = 1$ et $n = 2$; le nombre d'orbites est 3.
- 4)** Un $\mathbb{Z}$ -ensemble E est défini par un homomorphisme $\varphi : \mathbb{Z} \rightarrow S_E$. Pour définir φ , il suffit de définir $\varphi(1) = f : E \rightarrow E$. Autrement dit, la donnée d'un $\mathbb{Z}$ -ensemble équivaut à la donnée d'une bijection de E .
- 5)** Si A et B sont deux points du cercle unité de $\mathbb{R}^2$, il suffit de considérer la rotation de centre O et d'angle $(\overrightarrow{OA}, \overrightarrow{OB})$.

Soient deux points A et B de la sphère unité, P un plan contenant O , A et B (un tel plan est unique sauf si A et B sont diamétralement opposés) et D une droite perpendiculaire à P en O . Il existe une rotation r d'axe D telle que $r(A) = B$.

6) L'action de $GL(n, \mathbb{R})$ n'est pas transitive sur $\mathbb{R}^n$ car l'orbite de 0 est 0 ; elle est transitive sur l'ensemble des bases de $\mathbb{R}^n$: c'est la propriété universelle 7.4.

18.12 1) Tout élément non nul de $(\mathbb{Z}/p\mathbb{Z})^n$ peut être dans l'orbite de e_1 . Pour le voir, soit f_1 un vecteur non nul de $(\mathbb{Z}/p\mathbb{Z})^n$. On peut compléter les familles $\{e_1\}$ et $\{f_1\}$ en bases $(e_i)_{1 \leq i \leq n}$ et $(f_i)_{1 \leq i \leq n}$ de $(\mathbb{Z}/p\mathbb{Z})^n$ et définir une application linéaire $\varphi : (\mathbb{Z}/p\mathbb{Z})^n \rightarrow (\mathbb{Z}/p\mathbb{Z})^n$ par $\varphi(e_i) = f_i$ pour $1 \leq i \leq n$; on a $\varphi \in GL(n, \mathbb{Z}/p\mathbb{Z})$ et $\varphi(e_1) = f_1$. Le cardinal de l'orbite de e_1 est donc $p^n - 1$.

2) Si une application linéaire stabilise e_1 , la première ligne de sa matrice par rapport à la base canonique est de la forme $(1 \ a_2 \dots a_n)$, la première colonne est $(1 \ 0 \dots 0)$ et le reste est une matrice inversible d'ordre $n - 1$. Comme $a_2, \dots, a_n$ sont quelconques, le cardinal du stabilisateur de e_1 est : $p^{n-1}u_{n-1}$; comme l'orbite de e_1 a $p^n - 1$ éléments, $u_n = (p^n - 1) p^{n-1}u_{n-1}$; comme $u_1 = p - 1$,

$$u_n = p^{n(n-1)/2} \prod_{1 \leq k \leq n} (p^k - 1) = \prod_{1 \leq k \leq n} (p^n - p^{n-k})$$

et $N = n(n - 1)/2$.

La seconde forme du résultat peut être obtenue directement si l'on remarque que se donner une application linéaire inversible f , c'est se donner une base de $(\mathbb{Z}/p\mathbb{Z})^n$ dont

les vecteurs sont les images par f des vecteurs de la base canonique de $(\mathbb{Z}/p\mathbb{Z})^n$. On peut choisir le premier vecteur de base de $p^n - 1$ façons (il faut exclure le vecteur nul), le second vecteur de base de $p^n - p$ façons (il faut exclure les vecteurs colinéaires au précédent), etc.

3) Prendre pour H le sous-groupe des matrices triangulaires supérieures n'ayant que des 1 sur la diagonale.

4) La formule de récurrence est la même, avec $v_1 = 1$, d'où $v_n = u_n / (p - 1)$. On peut aussi remarquer que $\text{SL}(n, \mathbb{Z}/p\mathbb{Z})$ est le noyau de $\det : \text{GL}(n, \mathbb{Z}/p\mathbb{Z}) \rightarrow \mathbb{Z}/p\mathbb{Z}$. Comme l'image de cet homomorphisme a $p - 1$ éléments, on retrouve le même résultat :

$$v_n = p^{n(n-1)/2} \prod_{1 \leq k \leq n} (p^k - 1).$$

18.13 1) Voir exercice 11.1 f).

2) Il y a une classe de conjugaison par élément de G et l'équation des classes s'écrit $|G| = \sum_{g \in G} 1$.

3) a) Comme $sr^k s = r^{-k}$ et que le conjugué de r^k par un élément de la forme sr^l est aussi r^{-k} , r^k a pour conjugués lui-même et r^{-k} (si $r^{-k} \neq r^k$).

Les conjugués de sr^k sont de la forme $r^l sr^k r^{-l} = sr^{k-2l}$ et $sr^l sr^k r^{-l} s = sr^{2l-k}$.

Si n est impair, tous les éléments de la forme sr^k sont conjugués entre eux ; si n est pair, il y a deux classes de conjugaison, une pour k pair, l'autre pour k impair.

b) D_3 : les classes de conjugaison sont $\{\text{id}\}$, $\{r, r^2\}$, $\{s, sr, sr^2\}$, d'où l'équation : $6 = 1 + 2 + 3$.

D_4 : les classes de conjugaison sont $\{\text{id}\}$, $\{r^2\}$, $\{r, r^3\}$, $\{s, sr^2\}$, $\{sr, sr^3\}$, d'où l'équation : $8 = 1 + 1 + 2 + 2 + 2$.

D_5 : les classes de conjugaison sont $\{\text{id}\}$, $\{r^2, r^3\}$, $\{r, r^4\}$, $\{s, sr, sr^2, sr^3, sr^4\}$, d'où l'équation : $10 = 1 + 2 + 2 + 5$.

D_6 : les classes de conjugaison sont $\{\text{id}\}$, $\{r^3\}$, $\{r, r^5\}$, $\{r^2, r^4\}$, $\{s, sr^2, sr^4\}$, $\{sr, sr^3, sr^5\}$, d'où l'équation : $12 = 1 + 1 + 2 + 2 + 3 + 3$.

4) Pour S_3 :

Comme $D_3 \simeq S_3$, ce qui précède donne l'équation des classes de S_3 . On peut préciser la correspondance entre partitions de 3 et éléments de S_3 .

La partition donnée par $3 = 1 + 1 + 1$ correspond à trois orbites à un élément ; c'est la classe de l'identité.

La partition donnée par $3 = 1 + 2$ correspond à une orbite à un élément et une orbite à deux éléments qui est la classe des trois transpositions.

La partition donnée par $3 = 3$ correspond à une orbite à trois éléments ; c'est la classe des deux permutations circulaires.

Pour S_4 :

$4 = 1 + 1 + 1 + 1$ correspond à la classe de id,

$4 = 1 + 1 + 2$ correspond à la classe des six transpositions,

$4 = 1 + 3$ correspond à la classe des huit 3-cycles,

$4 = 2 + 2$ correspond à la classe des trois doubles transpositions,

$4 = 4$ correspond à la classe des six 4-cycles.

On a bien : $24 = 1 + 6 + 8 + 3 + 6$.

Pour S_5 :

$5 = 1 + 1 + 1 + 1 + 1$ correspond à la classe de id,

$5 = 1 + 1 + 1 + 2$ correspond à la classe des dix transpositions,

$5 = 1 + 1 + 3$ correspond à la classe des vingt 3-cycles,

$5 = 1 + 2 + 2$ correspond à la classe des quinze doubles transpositions,

$5 = 2 + 3$ correspond à la classe des vingt « 3-cycles et transpositions »,

$5 = 1 + 4$ correspond à la classe des trente 4-cycles

$5 = 5$ correspond à la classe des vingt-quatre 5-cycles.

Et 120 est bien la somme des cardinaux des différentes classes.

5) Deux éléments conjugués dans A_n sont conjugués dans S_n mais la réciproque n'est pas vraie comme le montre l'exemple de $(1\ 2\ 3)$ et $(1\ 3\ 2)$ non conjugués dans A_3 .

Cas de A_4 . Soit $\sigma = (a\ b\ c)$ un 3-cycle de A_4 , $a, b, c \in \{1, 2, 3, 4\}$ et d le quatrième élément. Le stabilisateur de σ doit fixer d ; ses éléments sont $e, (a\ b\ c), (a\ c\ b)$. L'orbite a donc quatre éléments.

Comme il y a huit 3-cycles, on a deux classes de conjugaison : celle de $(1\ 2\ 3)$, $(1\ 4\ 2)$, $(1\ 3\ 4)$, $(2\ 4\ 3)$ et l'autre. Les doubles transpositions sont conjuguées : on vérifie facilement que la classe de $(1\ 2)(3\ 4)$ contient les deux autres. D'où l'équation des classes : $12 = 1 + 4 + 4 + 3$.

Cas de A_5 . Maintenant les 3-cycles sont conjugués, car si on se donne deux 3-cycles $\sigma = (a\ b\ c)$ et $\sigma' = (a'\ b'\ c')$, une permutation s telle que $s(a) = a', s(b) = b', s(c) = c'$ donne $s \circ \sigma \circ s^{-1} = \sigma'$; la condition $s \in A_5$ s'obtient en jouant sur les images des deux derniers éléments : on peut composer s ou non avec la transposition échangeant ces deux éléments.

Soient $(a\ b)(c\ d)$ et $(a'\ b')(c'\ d')$ deux doubles transpositions ; si les quatre éléments forment le même ensemble, ces permutations sont conjugués puisque tout se passe comme ci-dessus dans A_4 ; si $\{a, b\} = \{a', b'\}$, on a par exemple $c = c'$ et $(d\ c\ d')$ permet de prouver qu'ils sont conjugués ; enfin $(a\ b)(c\ d)$ et $(a\ c)(e\ b)$ sont conjugués par $(d\ b\ a\ c\ e)$. Il n'y a qu'une classe de conjugaison, comme dans S_5 .

On trouve vingt-quatre 5-cycles partagés en deux classes. En effet, l'effectif d'une classe divise 60 et 24 ; d'autre part, on énumère facilement 12 conjugués de $(1\ 2\ 3\ 4\ 5)$.

L'équation est $60 = 1 + 20 + 15 + 12 + 12$.

On peut aussi raisonner comme dans le cas de A_4 avec les stabilisateurs. Comme dans A_4 , le stabilisateur d'un 3-cycle $(a\ b\ c)$ a trois éléments et le stabilisateur d'une double transposition a 4 éléments, les orbites correspondantes ont 20 et 15 éléments respectivement. Le stabilisateur d'un 5-cycle a cinq éléments, car si s stabilise $(a\ b\ c\ d\ e)$, l'image $s(a)$ détermine s ; donc l'orbite de ce 5-cycle a 12 éléments. On en déduit qu'il n'existe pas de sous-groupe distingué non trivial de A_5 , celui-ci devant être réunion de classes de conjugaison et d'ordre un diviseur de 60 : les nombres 1, 12, 12, 15, 20 ne le permettent pas.

18.14 1, 2) Comme Γ est un sous-ensemble d'un produit, on va compter ses éléments en comptant à partir de chacune des coordonnées.

Comptons à partir de G : chaque g de G distinct de l'identité est une rotation ayant deux points fixes sur S : les intersections de l'axe de la rotation avec S ; d'où la première formule.

Comptons à partir de X , en décomposant X en une union disjointe indexée par les éléments de chacune des orbites :

$$\Gamma = \{(g, x), g \in G \setminus \{\text{id}\}, x \in S, g(x) = x\} = \bigcup_{1 \leq i \leq k} \bigcup_{x \in O_i} \{(g, x), g \in S_x \setminus \{\text{id}\}\} ;$$

mais $|S_x| = n_i$ pour tout x de l'orbite O_i et $|O_i| = \frac{n}{n_i}$, d'où (E).

En égalant les deux expressions obtenues pour l'ordre de Γ et en divisant par n , on obtient (E).

3) Comme $n_i > 1$ (un point de X est stabilisé par la rotation qui le définit), on a : $1 - 1/n_i \geq 1/2$; la relation (E) impose donc $k \leq 3$, puisque $2 - 2/n < 2$; comme $n > 1$, on a : $2 - 2/n \geq 1$, donc $k > 1$. Étudions ces deux cas.

Cas de $k = 2$

$2 - 2/n = 1 - 1/n_1 + 1 - 1/n_2$ donne : $2/n = 1/n_1 + 1/n_2$; comme $n_1 \leq n$ et $n_2 \leq n$, $1/n_1 + 1/n_2 \geq 2/n$, la seule possibilité est : $n_1 = n_2 = n$; les deux orbites ont chacune un élément. C'est le cas d'un groupe cyclique engendré par une rotation d'angle $2\pi/n$.

Cas de $k = 3$

La relation (E) donne : $1 + 2/n = 1/n_1 + 1/n_2 + 1/n_3$; supposons $n_1 \leq n_2 \leq n_3$; on a alors : $1 + 2/n \leq 3/n_1$, donc $n_1 = 2$, ce qui donne : $1/2 + 2/n = 1/n_2 + 1/n_3$ donc $1/2 + 2/n \leq 2/n_2$, ce qui impose $n_2 \leq 3$.

Si $n_2 = 2$, on a : $n = 2n_3$.

Si $n_2 = 3$, on a : $1/6 + 2/n = 1/n_3$, ce qui impose $n_3 \leq 5$, donc $n_3 = 3, 4$ ou 5 qui donne $n = 12, 24$ ou 60 . Récapitulons :

n	n_1	n_2	n_3
$2n_3$	2	2	n_3
12	2	3	3
24	2	3	4
60	2	3	5

4) La première ligne correspond au cas du groupe diédral D_n (engendré par r, s comme dans l'exercice **18.16**) considéré comme le groupe laissant stable l'ensemble des sommets d'un polygone régulier plan de n côtés, les r^k étant les rotations d'angle $2k\pi/n$ autour de l'axe du polygone, les sr^k étant les symétries par rapport aux n axes de symétries du polygone. Les trois dernières lignes du tableau correspondent aux groupes du tétraèdre régulier, du cube (ou de l'octaèdre régulier), du dodécagone (ou de l'icosaèdre régulier). Les vérifications sont longues. Nous ne traiterons que le premier cas, dans l'exercice suivant.

18.15 1) Il est clair que φ définit un homomorphisme de groupes. Une isométrie f de G conserve l'isobarycentre O des sommets de T . Si $\varphi(f) = e$, f laisse fixe trois sommets A, B, C de T , donc laisse fixe le repère affine (O, A, B, C) ; c'est l'identité ; donc φ est injective. Toute transposition est dans l'image de φ ; par exemple, la transposition $(A B)$ est l'image de la symétrie par rapport au plan contenant C, D et le milieu de $[AB]$. Comme S_4 est engendré par les transpositions, φ est surjective ; c'est donc un isomorphisme de groupes.

2) Les 3-cycles sont dans l'image de φ' ; par exemple, le 3-cycle $(A B C)$ est l'image de la rotation d'angle $2\pi/3$ et d'axe la perpendiculaire au plan (ABC) passant par D ; les doubles transpositions sont également dans l'image de φ' ; par exemple, $(A B)(C D)$ est l'image de la symétrie par rapport à la droite joignant les milieux de $[AB]$ et $[C D]$ (c'est la perpendiculaire commune à ces deux segments). Donc φ' est surjective.

OUVERTURES SUR LES ANNEAUX COMMUTATIFS UNITAIRES

19

Les anneaux ont été présentés au chapitre 12 pour donner un cadre théorique à quelques résultats d'arithmétique dans l'anneau $\mathbb{Z}$: notions de divisibilité et de pgcd, identité de Bézout.

Nous y avons défini rapidement aux sections 12.6, 12.8 et 12.9 les notions d'homomorphisme et d'isomorphisme d'anneaux, d'éléments inversibles, d'anneau intègre, d'idéal, d'idéal engendré par une partie d'un anneau, d'idéal principal et d'anneau principal. Il est nécessaire de revenir à ces pages (que nous ne pouvons répéter ici) avant de lire ce chapitre.

Nous continuerons à dire dans ce chapitre *anneau* pour *anneau commutatif unitaire* puisque, sauf mention contraire, ce sont ces anneaux que nous rencontrerons ; un homomorphisme d'anneau envoie l'élément unité sur l'élément unité.

Rappelons que, dans un anneau A , nous notons (E) l'idéal engendré par une partie $E \subset A$ et que la notation n'indique pas que l'idéal engendré dépend de A . On note (x) ou xA l'idéal engendré par un élément $x \in A$. Rappelons aussi qu'un anneau nul (on dit souvent *l'anneau nul*) est un anneau ayant un élément qui joue à la fois le rôle du 0 et du 1.

19.1 SOUS-ANNEAU, EXTENSION DE CORPS

19.1.1 Sous-anneau

Un sous-anneau d'un anneau commutatif unitaire $(A, +, \cdot)$ est une partie de A stable par addition, par multiplication et contenant l'élément unité de A ; c'est un sous-groupe de $(A, +)$ et c'est un anneau.

Un idéal I de A n'est pas un sous-anneau de A , sauf dans le cas où $1 \in I$ puisque dans ce cas $I = A$. Par exemple, $n\mathbb{Z}$ est un sous-anneau de $\mathbb{Z}$ seulement si $n = \pm 1$.

Anneaux de fonctions. En analyse, on considère souvent, si I désigne un intervalle de $\mathbb{R}$, des sous-anneaux de l'anneau $\text{Fonct}(I, \mathbb{R})$ de toutes les fonctions de I dans $\mathbb{R}$; par exemple, les sous-anneaux $\mathcal{C}^0(I, \mathbb{R})$, $\mathcal{C}^1(I, \mathbb{R})$, $\mathcal{C}^n(I, \mathbb{R})$, $\mathcal{C}^\infty(I, \mathbb{R})$ des fonctions respectivement continues, continûment dérivables, n fois continûment dérivables, de classe $\mathcal{C}^\infty$ sur I . La propriété d'être un sous-anneau correspond ici aux propriétés de continuité ou de dérivabilité d'une somme ou d'un produit de fonctions.

Sous-anneau engendré. Soient A un anneau et E une partie de A . On appelle sous-anneau engendré par E le plus petit sous-anneau de A contenant E . Ce sous-anneau existe : c'est l'intersection de la famille de tous les sous-anneaux de A contenant E (il en existe au moins un : A lui-même).

La construction de l'anneau des polynômes à une indéterminée à coefficients dans un anneau a été indiquée en 13.2 ; la proposition suivante montre comment cet anneau permet une description du sous-anneau engendré par un élément.

Proposition 1. Soient B un sous-anneau d'un anneau A et x un élément de A .

- 1) Le sous-anneau de A engendré par B et x est l'image de l'homomorphisme d'anneaux $\varphi : B[X] \rightarrow A$ défini par $\varphi(X) = x$.
- 2) Si x est un élément de A , le sous-anneau de A engendré par x est l'image de l'homomorphisme $\varphi : \mathbb{Z}[X] \rightarrow A$ défini par $\varphi(X) = x$.

Démonstration. 1) Notons C le sous-anneau de A engendré par B et x . Tout élément de l'image de φ est de la forme $\sum_{0 \leq k \leq n} b_k x^k$ avec $b_k \in B$ pour $k = 0, \dots, n$; c'est donc un élément de C , d'où $\text{im}(\varphi) \subset C$. Comme $\text{im}(\varphi)$ est un sous-anneau de A contenant B et x , on a $C \subset \text{im}(\varphi)$; on en déduit $C = \text{im}(\varphi)$.

2) La seconde partie se déduit de la première en prenant pour B le sous-anneau image de $\mathbb{Z}$ par l'homomorphisme $\mathbb{Z} \rightarrow A$ (voir 12.6). $\square$

Les éléments de C étant des valeurs de polynômes à coefficients dans B , nous utiliserons la notation $B[x]$ pour désigner l'anneau engendré par B et x . La représentation d'un élément de $B[x]$ comme valeur d'un polynôme de $B[X]$ en x n'est pas, en général, unique.

19.1.2 Extension de corps

Soient K et L deux corps ; quand K est un sous-corps de L , on dit souvent que L est une extension du corps K ou simplement une extension de K .

Par extension (si on peut dire), s'il existe un homomorphisme $i : K \rightarrow L$, nécessairement injectif d'après 12.6, l'extension L de $i(K)$ est encore appelée extension de K .

Par exemple, $\mathbb{C}$ est une extension de $\mathbb{R}$, $\mathbb{R}$ est une extension de $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$ est une extension de $\mathbb{Q}$, etc.

Extension engendrée. Soient L une extension d'un corps K et E une partie de L . Il existe une extension de K , notée $K(E)$ et appelée extension de K engendrée par E , contenant E et minimum pour la relation d'inclusion entre extensions de K contenues dans L : c'est l'intersection des extensions de K contenues dans L et contenant E (il en existe au moins une : L).

Une description de l'extension engendrée par un élément à l'aide du corps des fractions rationnelles à une indéterminée est donnée en 19.15.

Degré d'une extension. Si L est une extension de K , L est munie d'une structure de K -espace vectoriel, comme il est très facile de le vérifier. La dimension de cet espace est appelé degré de l'extension L de K et notée $[L : K]$. Le degré d'une extension peut-être fini ou infini.

Proposition 2 : formule de multiplicativité des degrés. Soient K un corps, L une extension de K de degré fini l , $B = (e_i)_{1 \leq i \leq l}$ une base de L en tant que K -espace vectoriel, M une extension de L de degré fini m , $C = (f_j)_{1 \leq j \leq m}$ une base de M en tant que L -espace vectoriel. Alors :

1) M est une extension de K de degré fini lm , autrement dit :

$$[M : K] = [M : L][L : K]$$

2) $D = (e_i f_j)_{1 \leq i \leq l, 1 \leq j \leq m}$ est une base de M en tant que K -espace vectoriel.

Démonstration. La famille D est libre sur K car si $0 = \sum_{1 \leq i \leq l, 1 \leq j \leq m} \lambda_{ij} e_i f_j$ avec les

$$\lambda_{ij} \in K, \text{ on a } 0 = \sum_{1 \leq j \leq m} \left(\sum_{1 \leq i \leq l} \lambda_{ij} e_i \right) f_j, \text{ ce qui montre que } \sum_{1 \leq i \leq l} \lambda_{ij} e_i = 0 \text{ pour } 1 \leq j \leq m,$$

puisque C est une base de M sur L ; on en déduit que les λ_{ij} sont tous nuls puisque B est une base de L sur K .

La famille D est génératrice. D'abord, si $u \in M$, il existe des éléments $(u_j)_{1 \leq j \leq m}$ tels que $u = \sum_{1 \leq j \leq m} u_j f_j$ puisque C est une base de M sur L ; pour $1 \leq j \leq m$, il existe

des éléments $\lambda_{ij} \in K$, $1 \leq i \leq l$, tels que

$$u_j = \sum_{1 \leq i \leq l} \lambda_{ij} e_i ; \text{ on a } u = \sum_{1 \leq i \leq l, 1 \leq j \leq m} \lambda_{ij} e_i f_j \text{ et on conclut.} \quad \square$$

Pour comprendre l'intérêt pratique de cette formule simple, il faut attendre le chapitre suivant.

19.2 CARACTÉRISTIQUE

Caractéristique d'un anneau. Soit A un anneau. On a vu en 12.6 qu'il existe un unique homomorphisme d'anneaux $f: \mathbb{Z} \rightarrow A$. Le noyau de f est un idéal de $\mathbb{Z}$; il est donc de la forme $n\mathbb{Z}$ et on peut choisir $n \geq 0$. Si $n > 0$, c'est le plus petit des entiers $k > 0$ tels que $k \cdot 1 = 0$. On appelle n la caractéristique de l'anneau A et on la note $\text{car}(A)$.

Si B est un sous-anneau d'un anneau A , alors $\text{car}(B) = \text{car}(A)$; deux anneaux n'ayant pas même caractéristique ne sont pas isomorphes.

Si A est un anneau de caractéristique n , l'anneau de polynômes $A[X]$ est également de caractéristique n .

Caractéristique d'un corps. Une extension d'un corps a même caractéristique que lui. Si K est un corps de caractéristique n non nulle, alors n est premier ; en effet, si $n = pq$ avec $p, q < n$, on a $n \cdot 1 = (p \cdot 1)(q \cdot 1) = 0$, ce qui implique $p \cdot 1 = 0$ ou $q \cdot 1 = 0$ et contredit la définition de la caractéristique.

Exemples. Sont de caractéristique 0 : l'anneau $\mathbb{Z}$, les corps $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, etc.

L'anneau $\mathbb{Z}/n\mathbb{Z}$ est de caractéristique n et, si p est premier, le corps $\mathbb{Z}/p\mathbb{Z}$ est de caractéristique p .

Si $\text{car}(A) = 1$, A est l'anneau nul.

Proposition, définition : homomorphisme de Frobenius. Soit A un anneau de caractéristique p premier.

1) Pour tout x et y de A :

$$(x + y)^p = x^p + y^p.$$

2) L'application $F: A \rightarrow A$ définie par $F(x) = x^p$ est un homomorphisme d'anneau, appelé homomorphisme de Frobenius.

3) Pour tous $x, y \in A$ et tout $k \in \mathbb{N}$:

$$(x + y)^{p^k} = x^{p^k} + y^{p^k}.$$

Démonstration. 1) Cette égalité surprend d'abord puisqu'elle est complètement fautive en caractéristique 0. En voici la raison : pour tout entier k tel que $1 \leq k \leq p-1$, comme $p! = k!(p-k)!C_p^k$ et comme le nombre premier p ne divise aucun des facteurs de $k!(p-k)!$, il divise C_p^k ; autrement dit : $C_p^k \equiv 0 \pmod{p}$ (par exemple, $C_7^3 = 35 \equiv 0 \pmod{7}$). La formule du binôme montre alors que, dans A bien sûr :

$$(x + y)^p = x^p + C_p^1 x^{p-1} y + \dots + C_p^{p-1} x y^{p-1} + y^p = x^p + y^p.$$

- 2) Le 1) montre que $F(x + y) = F(x) + F(y)$. L'égalité $F(xy) = F(x)F(y)$ est évidente.
 3) L'homomorphisme $F^k = F \circ \dots \circ F$ vérifie donc $F^k(x + y) = F^k(x) + F^k(y)$.

Comme $F^k(x) = x^{p^k}$, on obtient la formule donnée. $\square$

19.3 QUOTIENT D'UN ANNEAU PAR UN IDÉAL

Au chapitre 12, nous avons montré que les idéaux de l'anneau $\mathbb{Z}$ étaient de la forme $n\mathbb{Z}$ avec n entier et nous avons défini les quotients de l'anneau $\mathbb{Z}$ à la main ; il ne nous paraissait pas utile alors de définir abstraitement ces quotients. Nous allons le faire maintenant.

Construction d'un anneau quotient. Soient donc A un anneau et I un idéal de A . Comme A est commutatif, I est, en particulier, un sous-groupe distingué de A , ce qui permet de définir le groupe quotient A/I et la projection canonique $\pi : A \rightarrow A/I$, $x \mapsto x \bmod I$. La relation d'équivalence $\mathcal{R}$ associée sur A s'écrit dans ce cadre avec une notation additive : $a \mathcal{R} b$ si et seulement si $a - b \in I$.

Montrons que A/I hérite en plus d'une structure d'anneau commutatif unitaire. Soient deux éléments de A/I ; ils sont de la forme $\pi(a)$ et $\pi(a')$; on définit leur produit comme étant $\pi(aa')$. Cette définition ne dépend pas des choix de a et a' représentant les éléments de A/I ; en effet, si b et b' représentent les mêmes éléments de A/I , il existe $x, x' \in I$ tels que $b = a + x$, $b' = a' + x'$ et alors $bb' = aa' + y$ avec $y = xa' + x'a + xx'$; comme y est un élément de I , on a $\pi(bb') = \pi(aa')$.

Il est alors facile de terminer de vérifier que A/I est ainsi muni d'une structure d'anneau (si l'anneau est unitaire, son élément neutre pour la multiplication est $\pi(1_A)$) appelé anneau quotient de A par I et que la projection canonique est un homomorphisme d'anneaux.

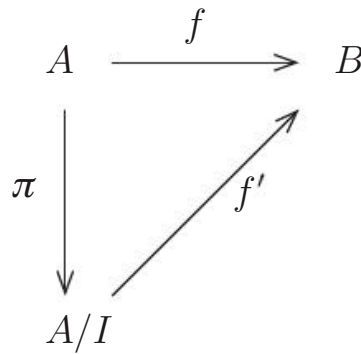
On notera que, si $I = A$, alors A/I est l'anneau nul et, si $I = (0)$, alors $A/I = A$.

Comme pour les groupes, on peut penser que quotienter par un idéal revient à *tuer* les éléments de l'idéal en les identifiant à 0. On dispose aussi d'une propriété universelle permettant de construire facilement des homomorphismes de source l'anneau quotient.

Proposition : propriété universelle du quotient d'un anneau par un idéal.

Soient A un anneau, I un idéal de A , $\pi : A \rightarrow A/I$ la projection canonique.

Pour tout homomorphisme d'anneaux $f : A \rightarrow B$ tel que $f(I) = 0$, il existe un unique homomorphisme d'anneaux $f' : A/I \rightarrow B$ tel que le diagramme suivant commute.



Démonstration. La propriété universelle du quotient de groupes nous montre déjà qu'il existe un unique homomorphisme de groupes, f' , rendant commutatif le diagramme précédent. Pour $\alpha \in A/I$ de représentant a dans A , on a $f'(\alpha) = f(a)$ et il est facile de vérifier que f' est aussi un homomorphisme d'anneaux : si α et β sont des éléments de A/I , de représentants a et b dans A , on a

$$f'(\alpha\beta) = f'(\pi(ab)) = f(ab) = f(a)f(b) = f'(\pi(a))f'(\pi(b)) = f'(\alpha)f'(\beta) ;$$

enfin $f'(1_{A/I}) = f(1_A) = 1_B$. □

Donnons deux corollaires analogues à ceux donnés en 18.3.

Corollaire 1. Soient $\varphi : A \rightarrow A'$ un isomorphisme d'anneaux, I un idéal de A . Alors $\varphi(I)$ est un idéal de A' et A/I est isomorphe à $A'/\varphi(I)$.

Démonstration. La démonstration est analogue à celle donnée pour le corollaire 3 de 18.3. □

Corollaire 2. Soit $f : A \rightarrow A'$ un homomorphisme d'anneaux. L'homomorphisme $\varphi : A/\ker(f) \rightarrow f(A)$ défini par la propriété universelle est un isomorphisme.

Démonstration. La démonstration est analogue à celle donnée pour le corollaire 2 de 18.3. □

On remarquera qu'on peut reprendre entièrement cette section en ne supposant pas les anneaux unitaires : les quotients d'anneaux se définissent de la même manière et ont les mêmes propriétés universelles.

19.4 EXEMPLES DE QUOTIENTS

Nous allons présenter dans cette section de nombreux exemples de constructions d'anneaux par quotient. Certains anneaux, comme $\mathbb{Q}[\sqrt{n}]$ (voir plus loin), peuvent également être présentés comme des sous-anneaux d'anneaux connus, de $\mathbb{C}$ par

exemple, mais leur présentation par quotient apporte un autre éclairage sur leurs propriétés.

Quotients de $\mathbb{Z}$. Soit $n \in \mathbb{Z}$. Le quotient de l'anneau $\mathbb{Z}$ par l'idéal $n\mathbb{Z}$ est bien l'anneau $\mathbb{Z}/n\mathbb{Z}$ défini à la main dans les sections 12.3 et 12.4, puisque nous avons alors quotienté $\mathbb{Z}$ par la relation d'équivalence $a - b \in I$ avec $I = n\mathbb{Z}$. Dans ce cas, pour travailler dans le quotient on prend, comme nous l'avons déjà dit, des entiers bien choisis représentant les classes d'équivalence, le plus souvent $0, 1, \dots, n-1$.

Anneaux et corps de caractéristique finie. Si A est un anneau de caractéristique n , le noyau de l'application $f : \mathbb{Z} \rightarrow A$ définie par $f(1) = 1$ est l'idéal $n\mathbb{Z}$ de $\mathbb{Z}$; son image est un sous-anneau de A , isomorphe à $\mathbb{Z}/n\mathbb{Z}$ par le corollaire 2 de 19.3. En particulier, un corps de caractéristique p (p premier d'après 19.2) contient un sous-corps isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

Constructions de $\mathbb{C}$. Voilà sans doute plusieurs années que nos lectrices et lecteurs utilisent les nombres complexes, mais certains et certaines d'entre vous n'ont peut-être jamais vraiment construit le corps formé par ces nombres à partir du corps des nombres réels. Ils sont en bonne compagnie : les mathématiciens n'ont donné de telles constructions qu'au dix-neuvième siècle, près de 300 ans après la découverte des premiers nombres complexes, par Cardan, en 1545 (voir [ESC1], pour des détails sur cette histoire).

Construction de $\mathbb{C}$ comme quotient : $\mathbb{R}[X]/(X^2 + 1)$. Considérons l'anneau des polynômes $\mathbb{R}[X]$ et l'idéal $(X^2 + 1)$ engendré par le polynôme $X^2 + 1$. Notons A le quotient de $\mathbb{R}[X]$ par $(X^2 + 1)$ et $\pi : \mathbb{R}[X] \rightarrow A$ la projection canonique.

Le composé de l'inclusion $j : \mathbb{R} \rightarrow \mathbb{R}[X]$ avec π définit une injection de $\pi \circ j : \mathbb{R} \rightarrow A$ car si a, a' sont deux nombres réels tels que $\pi(j(a)) = \pi(j(a'))$, on a $j(a') = j(a) + (X^2 + 1)P$ où $P \in \mathbb{R}[X]$, ce qui ne peut avoir lieu que si $P = 0$, d'où $a = a'$.

Soit $P \in \mathbb{R}[X]$; la division euclidienne de P par $X^2 + 1$ permet de trouver des polynômes Q et R uniques tels que $P = (X^2 + 1)Q + R$ et $\deg(R) \leq 1$, autrement dit, $R(X)$ est de la forme $a + bX$ avec $a, b \in \mathbb{R}$.

Comme $(X^2 + 1)Q$ est un élément de l'idéal $(X^2 + 1)$, on voit que $\pi(P) = \pi(R)$. Si on pose $\pi(X) = \iota$, tout élément de A s'écrit de façon unique sous la forme $a + b\iota$ avec $a, b \in \mathbb{R}$. Comme $\pi(X^2 + 1) = 0$, on a $\iota^2 = -1$.

Appelons désormais nombres complexes les éléments de A , notons $\mathbb{C}$ l'anneau A et vérifions que tout se passe comme nous en avons l'habitude. L'addition de deux nombres complexes s'écrit $(a + \iota b) + (a' + \iota b') = a + a' + \iota(b + b')$ par définition des lois du quotient. La multiplication de deux nombres complexes s'écrit : $(a + \iota b)(a' + \iota b') = aa' + \iota^2 bb' + \iota(ab' + a'b)$, d'où :

$$(a + \iota b)(a' + \iota b') = aa' - bb' + \iota(ab' + a'b).$$

Il ne reste qu'à vérifier que tout élément non nul est inversible. Soit $a + \iota b$ un élément non nul de A , c'est-à-dire tel que a et b ne soient pas tous deux nuls. L'égalité $(a + \iota b)(a - \iota b) = a^2 + b^2$ montre que l'inverse de $a + \iota b$ est donné par

$$(a + \iota b)^{-1} = \frac{a - \iota b}{a^2 + b^2}.$$

Le corps $\mathbb{R}$ s'identifie aux éléments de $\mathbb{C}$ correspondants par l'homomorphisme injectif $\pi \circ j$. On peut donc dire que $\mathbb{C}$ est une extension de degré 2 de $\mathbb{R}$. Une base du $\mathbb{R}$ -espace vectoriel $\mathbb{C}$ est $(1, \iota)$.

Histoire de la construction de $\mathbb{C}$ par quotient. Cauchy explique cette construction à l'Académie des sciences le 28 juin 1847. Il souligne d'abord que, à son époque, la théorie des imaginaires est fondée *sur des principes qui manquent de clarté*. Il rappelle l'équivalence définie par Gauss de deux nombres entiers modulo un troisième, l'équivalence définie par Kummer de deux polynômes φ et χ de $\mathbb{R}[X]$ modulo un troisième w (ce sont les notations de Cauchy). Pour indiquer cette dernière équivalence, Cauchy propose d'utiliser la lettre i des géomètres allemands en écrivant $\varphi(i) = \chi(i)$, où i est une *lettre symbolique* (pour nous, c'est $\pi(X)$). Quand w est de degré n , Cauchy explique qu'on peut toujours se ramener à une expression de degré $n - 1$ en i . Il décrit donc ce qui est pour nous le quotient de l'anneau $\mathbb{R}[X]$ par l'idéal engendré par w . Même si Cauchy ne parle ni de structure d'anneau, ni de structure de corps, il n'a pas de difficulté pour transporter les opérations de $\mathbb{R}[X]$ sur ces nouvelles expressions. Quand il a bien expliqué cela, il particularise au cas $w = X^2 + 1$.

Construction de $\mathbb{C}$ par matrices. La construction de $\mathbb{C}$ précédente n'est pas la seule possible. On peut aussi identifier les réels aux homothéties du plan $\mathbb{R}^2$ et introduire les complexes comme les similitudes directes dans ce plan, autrement dit comme des applications linéaires particulières. La matrice, dans la base canonique, d'une similitude directe est de la forme $k \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$, ce qui conduit à considérer dans

$M_2(\mathbb{R})$ l'ensemble E des matrices de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ avec $a, b \in \mathbb{R}$. On vérifie

facilement que E est un sous-anneau de $M_2(\mathbb{R})$, puis que E est un corps car tout élément non nul de E est inversible (l'inverse de la matrice (non nulle) $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ est

$$\frac{1}{a^2 + b^2} \begin{pmatrix} a & b \\ -b & a \end{pmatrix}).$$

En posant $1 = I_2$ et $\iota = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, on voit que $\iota^2 = -1$ et que tout élément de E s'écrit $a + \iota b$; on retrouve bien $\mathbb{C}$.

Construction de $\mathbb{R}$. Il ne peut être question de traiter complètement ici la construction du corps des réels à partir du corps $\mathbb{Q}$. On notera cependant que la construction des réels proposée par Georg Cantor en 1872 se fait en deux étapes et que l'une d'elle est algébrique : elle consiste à construire le quotient de l'anneau A des suites de Cauchy (voir un cours d'analyse) de rationnels (pour montrer que A est un anneau, on peut le voir comme sous-anneau de l'anneau des suites quelconques de rationnels) par l'idéal des suites de A convergeant vers 0. Pour comprendre ensuite que le quotient de A obtenu est bien un modèle des réels comme on le souhaite, il reste un certain nombre de vérifications techniques, sans difficultés mais fastidieuses.

Construction de $\mathbb{Z}[\iota]$ comme quotient. L'anneau $\mathbb{Z}[\iota]$ peut être présenté comme le sous-anneau de $\mathbb{C}$ engendré par ι , mais il est intéressant aussi de le présenter comme un quotient.

L'idéal $(X^2 + 1)$ de $\mathbb{Z}[X]$ est le noyau de l'homomorphisme $\varphi : \mathbb{Z}[X] \rightarrow \mathbb{C}$ défini par $X \mapsto \iota$. En effet, soit P est un polynôme de $\mathbb{Z}[X]$; la division euclidienne de P par $X^2 + 1$ s'écrit : $P = (X^2 + 1)Q + a + bX$ avec $a, b \in \mathbb{Z}$; on a $P(\iota) = 0$ si et seulement si $a + b\iota = 0$, soit $a = b = 0$, ce qui équivaut à $P \in (X^2 + 1)$. Par conséquent, $\mathbb{Z}[X]/(X^2 + 1)$ est isomorphe à $\text{im}(\varphi)$, c'est-à-dire au sous-anneau de $\mathbb{C}$ formé des nombres complexes de la forme $a + b\iota$ avec $a, b \in \mathbb{Z}$. Cet anneau est appelé anneau des entiers de Gauss.

Histoire. Gauss introduit l'anneau $\mathbb{Z}[\iota]$ pour ses recherches sur les lois de réciprocité en 1832. L'article est écrit en latin ; Gauss y décrit les unités de l'anneau, les éléments associés, les éléments premiers et établit la factorialité de l'anneau. Il propose également le mot complexe pour remplacer celui d'imaginaire utilisé jusque là.

Étude de $\mathbb{Q}[\sqrt{n}]$ et $\mathbb{Z}[\sqrt{n}]$. Soit $n \in \mathbb{Z}$. On suppose que n est sans facteur carré et que la notation $\sqrt{n}$ désigne la racine carrée de n si $n \geq 0$ et $\iota\sqrt{-n}$ si $n < 0$.

Notons $\varphi : \mathbb{Q}[X] \rightarrow \mathbb{C}$ l'homomorphisme d'évaluation en $\sqrt{n}$; son image est par définition $\mathbb{Q}[\sqrt{n}]$. Si $P \in \mathbb{Q}[X]$, la division euclidienne de P par $X^2 - n$ donne $P = Q(X^2 - n) + a + bX$ avec $a, b \in \mathbb{Q}$. Comme $\varphi(P) = a + b\sqrt{n}$, $\varphi(P) = 0$ équivaut à $a = b = 0$ et on voit que $\ker(\varphi) = (X^2 - n)$. On en déduit $\mathbb{Q}[\sqrt{n}] \simeq \mathbb{Q}[X]/(X^2 - n)$ et que tout élément de $\mathbb{Q}[\sqrt{n}]$ peut s'écrire de manière unique $a + b\sqrt{n}$ avec $a, b \in \mathbb{Q}$; on a une inclusion $\mathbb{Q} \subset \mathbb{Q}[\sqrt{n}]$.

On peut vérifier directement que $\mathbb{Q}[\sqrt{n}]$ est un corps : l'inverse d'un élément non nul $a + b\sqrt{n}$ est $\frac{a}{a^2 - nb^2} - \frac{b}{a^2 - nb^2}\sqrt{n}$. On voit que $\mathbb{Q}[\sqrt{n}]$ est l'extension du corps $\mathbb{Q}$ engendrée par $\sqrt{n}$. Une base du $\mathbb{Q}$ -espace vectoriel $\mathbb{Q}[\sqrt{n}]$ est $(1, \sqrt{n})$; le corps $\mathbb{Q}[\sqrt{n}]$ est une extension de degré 2 de $\mathbb{Q}$.

La division euclidienne d'un polynôme $P \in \mathbb{Z}[X]$ par $X^2 - n$ s'écrit

$$P = Q(X^2 - n) + a + bX \text{ avec } Q \in \mathbb{Z}[X], a, b \in \mathbb{Z}$$

et on montre comme ci-dessus que $\mathbb{Z}[\sqrt{n}] \simeq \mathbb{Z}[X]/(X^2 - n)$ (attention, la même notation $(X^2 - n)$ désigne ici un idéal de $\mathbb{Z}[X]$ et ci-dessus un idéal de $\mathbb{Q}[X]$).

Par exemple, le quotient $\mathbb{Z}[X]/(X^2 + 5)$ est isomorphe à l'image de l'homomorphisme $\mathbb{Z}[X] \rightarrow \mathbb{C}$ défini par $X \mapsto i\sqrt{5}$, c'est-à-dire au sous-anneau de $\mathbb{C}$ des nombres complexes de la forme $a + bi\sqrt{5}$ avec $a, b \in \mathbb{Z}$. Il existe un second plongement $\mathbb{Z}[X]/(X^2 + 5)$ dans $\mathbb{C}$, de même image, défini par $X \mapsto -i\sqrt{5}$.

On pourrait penser aussi à construire l'anneau engendré sur $\mathbb{Q}$ par la racine carrée d'une fraction r/s avec $r, s \in \mathbb{Z}$ et premiers entre eux, mais cela n'apporte rien de nouveau car $\mathbb{Q}[\sqrt{r/s}] = \mathbb{Q}[\sqrt{rs}]$, puisque $\sqrt{rs} = s\sqrt{r/s}$.

Norme d'un élément de $\mathbb{Q}[\sqrt{n}]$. Si $z = a + b\sqrt{n}$, on définit $\tilde{z} = a - b\sqrt{n}$ et on appelle norme l'application $N : \mathbb{Q}[\sqrt{n}] \rightarrow \mathbb{C}$ définie par $N(z) = z\tilde{z} = a^2 - nb^2$. Dans le cas où $n < 0$, $\tilde{z} = \bar{z}$ et $N(z) = |z|^2 = a^2 + (-n)b^2 \in \mathbb{N}$. Pour utiliser la norme, les propriétés suivantes, immédiates pour $n < 0$ d'après les propriétés du module des nombres complexes, se vérifient par des calculs simples pour $n > 0$.

Propriétés de la norme.

- 1) Pour tout $z \in \mathbb{Q}[\sqrt{n}]$, $\tilde{\tilde{z}} = z$.
- 2) Pour tous $z_1, z_2 \in \mathbb{Q}[\sqrt{n}]$, $\widetilde{z_1 + z_2} = \tilde{z}_1 + \tilde{z}_2$, $\widetilde{z_1 z_2} = \tilde{z}_1 \tilde{z}_2$.
- 3) Pour tous $z_1, z_2 \in \mathbb{Q}[\sqrt{n}]$, $N(z_1 z_2) = N(z_1)N(z_2)$.
- 4) $N(z) = 0$ équivaut à $z = 0$.
- 5) La restriction de N à $\mathbb{Z}[\sqrt{n}]$ prend des valeurs entières, toujours positives si $n < 0$.

Quotients $A[X]/(X - a)$. Soit A un anneau et $a \in A$. L'endomorphisme d'anneau $\varphi : A[X] \rightarrow A[X]$ définie par $\varphi(X) = X - a$ admet comme inverse l'application $\psi : A[X] \rightarrow A[X]$ définie par $\psi(X) = X + a$; c'est donc un isomorphisme. Comme l'image de l'idéal (X) par φ est l'idéal $(X - a)$, le corollaire 1 de 19.3 donne un isomorphisme $A[X]/(X) \simeq A[X]/(X - a)$. Pour trouver un autre anneau isomorphe à $A[X]/(X)$, on remarque que l'homomorphisme d'évaluation $\text{ev}_0 : A[X] \rightarrow A$ (voir Exemple 1 de 13.8) défini par $\text{ev}_0(P) = P(0)$ est surjectif, de noyau (X) , d'où $A[X]/(X) \simeq A$ par le corollaire 2 de 19.3; dans cet isomorphisme, la classe d'un polynôme P donne $P(0)$. L'isomorphisme $A[X]/(X - a) \simeq A$ est donc $P \bmod (X - a) \mapsto P(a)$. Tous les anneaux quotients $A[X]/(X - a)$ sont isomorphes à A .

Exemple de $\mathbb{Z}[X]/(2X - 1)$. La description de ce quotient est un peu plus délicate. On peut dire que la construction revient à ajouter à $\mathbb{Z}$ une lettre X qu'on force à être égale à $1/2$. On peut aussi décrire la construction de cet anneau comme l'adjonction d'inverses aux puissances de 2 (on parle d'anneau de fractions). Notons $f : \mathbb{Z}[X] \rightarrow \mathbb{Q}$ l'homomorphisme défini par $f(X) = 1/2$, autrement dit l'évaluation des polynômes de $\mathbb{Z}[X]$ en $1/2$. L'image de f est donc $\mathbb{Z}[1/2]$, le sous-anneau engendré par $1/2$ dans $\mathbb{Q}$.

19.5 • Correspondance entre idéaux d'un anneau et idéaux d'un de ses quotients

Soit $P \in \mathbb{Z}[X]$; la division euclidienne de P par $2X - 1$ dans $\mathbb{Q}[X]$ s'écrit $P = (2X - 1)Q + r$, où $Q \in \mathbb{Q}[X]$ et où $r \in \mathbb{Q}$; $P(1/2) = 0$ équivaut à $r = 0$. Vérifions à la main que $Q \in \mathbb{Z}[X]$ dans ce cas : si $Q = \sum_{0 \leq k \leq n} a_k X^k$, on a

$$(2X - 1)Q = -a_0 + (2a_0 - a_1)X + \dots + (2a_{n-1} - a_n)X^n + 2a_n X^{n+1} ;$$

comme ce polynôme est égal à P , ses coefficients sont entiers et une récurrence immédiate montre que les coefficients de Q sont entiers. On en déduit que $\ker(f)$ est l'idéal $(2X - 1)$ de $\mathbb{Z}[X]$ et que l'image de f est isomorphe à l'anneau quotient $\mathbb{Z}[X]/(2X - 1)$, autrement dit $\mathbb{Z}[X]/(2X - 1) \simeq \mathbb{Z}[1/2]$.

19.5 CORRESPONDANCE ENTRE IDÉAUX D'UN ANNEAU ET IDÉAUX D'UN DE SES QUOTIENTS

Proposition 1.

1) Soient $f: A \rightarrow A'$ un homomorphisme d'anneaux et I' un idéal de A' . Alors $f^{-1}(I')$ est un idéal de A .

2) Soient $f: A \rightarrow A'$ un homomorphisme d'anneaux surjectif et I un idéal de A . Alors $f(I)$ est un idéal de A' .

Démonstration. 1) On a déjà vu que $f^{-1}(I')$ est un sous-groupe de A . De plus, si $a \in A$ et $x \in f^{-1}(I')$. On a $f(ax) = f(a)f(x) \in I'$, d'où $ax \in f^{-1}(I')$.

2) D'abord, $f(I)$ est un sous-groupe de A' . Si $y \in f(I)$ et $a' \in A'$, il existe $x \in I$ et $a \in A$ tels que $y = f(x)$, $a' = f(a)$ et on a $a'y = f(a)f(x) = f(ax) \in f(I)$. $\square$

Par contre, l'image d'un idéal n'est pas, en général, un idéal : par exemple, $\mathbb{Z}$ est un idéal de lui-même, mais non son image par l'inclusion dans $\mathbb{Q}$ ou $\mathbb{R}$.

Soient A un anneau, I un idéal de A , $\pi: A \rightarrow A/I$ la projection canonique. Cette situation est celle d'un homomorphisme surjectif d'anneaux $f: A \rightarrow A'$, puisqu'alors $A' \simeq A/\ker(f)$ comme en 18.4.

La projection π permet de construire une correspondance entre l'ensemble $\mathcal{A}$ des idéaux de A contenant I et l'ensemble $\mathcal{B}$ des idéaux de A/I de la façon suivante. On définit une application $\Phi: \mathcal{A} \rightarrow \mathcal{B}$ par $\Phi(J) = \pi(J)$ pour tout J de $\mathcal{A}$ et une application $\psi: \mathcal{B} \rightarrow \mathcal{A}$ par $\psi(L) = \pi^{-1}(L)$ pour tout L de $\mathcal{B}$.

Proposition 2.

1) Les applications Φ et ψ sont inverses l'une de l'autre.

2) Pour tout $J \in \mathcal{A}$, on a un isomorphisme d'anneaux non unitaires :

$$J/I \simeq \pi(J).$$

3) Pour tout $\mathfrak{J} \in \mathcal{A}$, on a un isomorphisme d'anneaux :

$$(A/I)/(\mathfrak{J}/I) \simeq (A/I)/\pi(\mathfrak{J}) \simeq A/\mathfrak{J}.$$

4) Soit $\mathfrak{J}$ un idéal quelconque de A ; $\pi(\mathfrak{J})$ est un idéal de A/I et on a un isomorphisme d'anneaux :

$$(A/I)/\pi(\mathfrak{J}) \simeq A/(I + \mathfrak{J}).$$

Démonstration. Il suffit de compléter la démonstration donnée dans 18.4 : les isomorphismes de groupes sont déjà construits et on vérifie qu'ils conservent produits et unités. Pour le 4), on remarque que $\pi^{-1}(\pi(\mathfrak{J})) = I + \mathfrak{J}$. $\square$

Exemple des quotients de $\mathbb{Z}$. La seule différence avec le paragraphe correspondant de 18.4 est qu'ici il s'agit d'anneaux et d'isomorphismes d'anneaux.

Les idéaux de l'anneau $\mathbb{Z}/n\mathbb{Z}$ correspondent aux idéaux de $\mathbb{Z}$ contenant $n\mathbb{Z}$; ils sont donc de la forme $d\mathbb{Z}$ avec d diviseur de n .

Le 3) du théorème donne l'isomorphisme entre anneaux $(\mathbb{Z}/ab\mathbb{Z})/(\mathfrak{a}\mathbb{Z}/ab\mathbb{Z}) \simeq \mathbb{Z}/a\mathbb{Z}$.

19.6 PRODUITS D'ANNEAUX

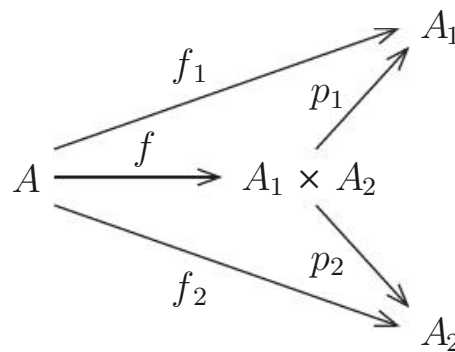
Soient A_1 et A_2 deux anneaux. En les considérant comme groupes, on définit le groupe produit $A_1 \times A_2$ avec la loi $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$ (voir 18.5). En posant $(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_2 b_2)$, on munit ce groupe d'une structure d'anneau, comme il est facile de le vérifier. L'élément neutre pour l'addition est $(0, 0)$, l'élément neutre pour la multiplication est $(1, 1)$. Les applications de projection $p_1 : A_1 \times A_2 \rightarrow A_1$ et $p_2 : A_1 \times A_2 \rightarrow A_2$ sont des homomorphismes d'anneaux.

On peut généraliser cette construction au produit d'une famille $(A_i)_{i \in I}$ d'anneaux.

Ces produits et ces projections vérifient des propriétés universelles qui sont analogues à celles vues en 18.5 pour les groupes.

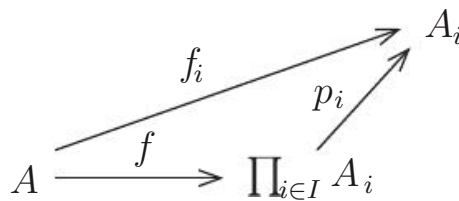
Propriété universelle des produits d'anneaux.

1) Soient A_1 et A_2 deux anneaux. Pour tout anneau A , tout homomorphisme d'anneaux $f_1 : A \rightarrow A_1$ et tout homomorphisme d'anneaux $f_2 : A \rightarrow A_2$, il existe un unique homomorphisme d'anneaux $f : A \rightarrow A_1 \times A_2$ noté $f = (f_1, f_2)$, tel que le diagramme



commute.

- 2) Soit $(A_i)_{i \in I}$ une famille d'anneaux. Pour tout anneau A et toute famille $(f_i : A \rightarrow A_i)_{i \in I}$ d'homomorphismes d'anneaux, il existe un homomorphisme d'anneaux unique $f : A \rightarrow \prod_{i \in I} A_i$ noté $f = (f_i)_{i \in I}$, tel que, pour tout $i \in I$, le diagramme



commute.

Démonstration. 1) On a montré en 18.5 l'existence d'un unique homomorphisme de groupes f tel que le diagramme de groupes commute. On vérifie que f est un homomorphisme d'anneaux.

- 2) On peut raisonner de manière analogue au 1). □

Remarques. Notons qu'un produit de deux anneaux n'est pas intègre car $(1, 0)(0, 1) = 0$. De même pour les produits de plusieurs anneaux.

Le produit de la famille vide est l'anneau nul.

Un produit fini d'anneaux peut également être vu comme une somme (voir la définition de somme de groupes abéliens en 18.5), mais les homomorphismes d'injection ne sont pas unitaires : par exemple, dans le cas du produit de deux anneaux A_1 et A_2 , l'application $j_1 : A_1 \rightarrow A_1 \times A_2$ est telle que $j_1(1) = (1, 0)$ qui est un diviseur de 0. Autrement dit, l'anneau $A_1 \times \{0\}$ n'est pas un sous-anneau de $A_1 \times A_2$, n'ayant pas le même élément unité.

19.7 OPÉRATIONS SUR LES IDÉAUX

Intersection d'idéaux. L'intersection de deux idéaux d'un anneau est un idéal de cet anneau.

Union d'idéaux. L'union de deux idéaux I_1, I_2 d'un anneau A n'est pas, en général, un idéal de A . On le voit en considérant les idéaux comme des sous-groupes : si aucun des deux idéaux n'est contenu dans l'autre, il existe des éléments $x \in I_1 \setminus I_2$ et $y \in I_2 \setminus I_1$ et alors $x + y \notin I_1 \cup I_2$. Cependant, l'union d'une suite finie d'idéaux, croissante pour l'inclusion, est un idéal, le plus grand d'entre eux. Plus généralement, on a le résultat suivant.

Proposition. Soient A un anneau et $(I_n)_{n \in \mathbb{N}}$ une suite croissante pour l'inclusion d'idéaux de A . Alors $I = \bigcup_{n \in \mathbb{N}} I_n$ est un idéal de A .

Démonstration. Soient $x, y \in I$. Il existe des entiers k, l tels que $x \in I_k$ et $y \in I_l$. Supposons $k \leq l$. Alors $x \in I_l$, d'où $x + y \in I_l$, d'où $x + y \in I$. D'autre part, pour tout $a \in A$, $ax \in I_k$, donc $ax \in I$. $\square$

Somme d'idéaux. Soient A un anneau, I_1 et I_2 deux idéaux de A . La somme de I_1 et I_2 est l'idéal, noté $I_1 + I_2$, engendré par $I_1 \cup I_2$ (la notion d'idéal engendré est décrite en 12.9). On vérifie que c'est l'ensemble des éléments de A de la forme $x_1 + x_2$ avec $x_1 \in I_1$ et $x_2 \in I_2$.

On définit de même la somme d'une famille $(I_j)_{j \in J}$ d'idéaux (où J est un ensemble fini ou infini) comme l'idéal engendré par la réunion des I_j . Un élément de cet idéal est une somme $x_{i_1} + \dots + x_{i_r}$ d'éléments d'un nombre fini d'idéaux de la famille. Si la famille est vide, elle engendre l'idéal $\{0\}$.

Dans $\mathbb{Z}$, la somme de deux idéaux (a) et (b) est l'idéal engendré par $\text{pgcd}(a, b)$ (voir 12.12.1) ; de manière analogue, pour un corps K , la somme des idéaux (S) et (T) de $K[X]$, est l'idéal engendré par $\text{pgcd}(S, T)$ (voir 13.4). Par contre, si K est un corps, dans l'anneau $K[X, Y]$ (voir 20.8 pour la définition de cet anneau), l'idéal $(X) + (Y)$ est l'idéal des polynômes sans terme constant alors que $\text{pgcd}(X, Y) = 1$.

Produit d'idéaux. Soient A un anneau, I_1 et I_2 deux idéaux de A . Le produit de I_1 et I_2 est l'idéal, noté $I_1 I_2$, engendré par l'ensemble des produits $x_1 x_2$ avec $x_1 \in I_1$ et $x_2 \in I_2$. Les éléments de $I_1 I_2$ sont des sommes finies de tels produits. On prendra garde à la notation : le produit de deux idéaux I_1 et I_2 est l'idéal engendré par les produits d'éléments de I_1 et de I_2 ; il n'est pas en général égal à l'ensemble de ces produits, ensemble qu'on pourrait également noter $I_1 I_2$.

On a $I_1 I_2 \subset I_1 \cap I_2$ car il est clair que $I_1 I_2 \subset I_1$ et que $I_1 I_2 \subset I_2$. Si $I_1 = A$, alors $I_1 I_2 = I_2$; l'anneau A joue le rôle d'élément neutre dans le produit d'idéaux.

Cette définition se généralise au cas d'une famille $(I_k)_{1 \leq k \leq n}$ finie d'idéaux de A . Les éléments de l'idéal $\prod_{1 \leq k \leq n} I_k$ sont des sommes finies de produits $a_1 \dots a_n$, avec $a_k \in I_k$ pour $1 \leq k \leq n$. Si la famille est vide, l'idéal produit est A .

La construction précédente permet de définir l'idéal I^n , puissance n -ième d'un idéal I , pour tout entier n : I^n a pour éléments les sommes finies de produits de n éléments de I ; on a $I^0 = A$.

Voici un exemple simple : le produit de deux idéaux principaux (a) et (b) d'un anneau A est l'idéal engendré par ab .

19.8 THÉORÈME CHINOIS

Histoire. Le livre de Jean-Claude Martzloff : *Histoire des mathématiques chinoises* (Masson, 1987) explique tout ce qu'on sait actuellement sur l'histoire ancienne de ce théorème. Il apparaît dans un livre d'arithmétique de Sunzi datant sans doute des années 400 de notre ère, puis dans des textes indiens. Voici une traduction du texte de Sunzi.

Soient des objets en nombre inconnu. Si on les compte par 3, il en reste 2 ; par 5, il en reste 3 ; par 7, il en reste 2. Combien y a-t-il d'objets ?

Réponse : 23.

Règle

En comptant par 3, il en reste 2 : pose 140.

En comptant par 5, il en reste 3 : pose 63.

En comptant par 7, il en reste 2 : pose 30.

Faire la somme de ces trois nombres : 233, soustraire 210, d'où la réponse.

En général, pour chaque unité restante d'un décompte par 3, poser 70 ; pour chaque unité restante d'un décompte par 5, poser 21 ; pour chaque unité restante d'un décompte par 7, poser 15. Si la somme vaut 106 ou plus, soustraire 105 pour trouver la réponse.

Que le texte chinois ne donne aucune justification ne doit pas surprendre, le style chinois de rédaction des mathématiques a toujours été de ne pas donner de démonstration et la traduction chinoise des *Éléments d'Euclide*, au XVII^e siècle, supprimait toutes les démonstrations ! On va voir que les nombres 70, 21, 15 donnés par Sunzi s'obtiennent naturellement (et non par bricolage). On remarquera que le texte de Sunzi vaut pour les restes modulo 3, 5, 7 et que rien n'est dit pour d'autres nombres.

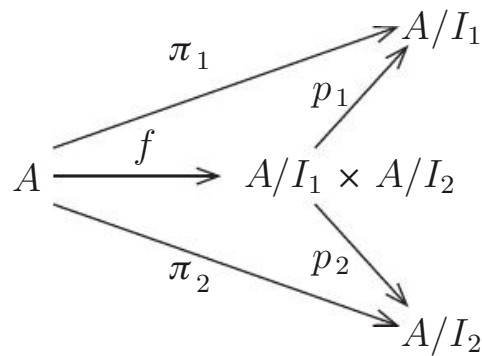
Définition. Soient A un anneau, I_1 et I_2 deux idéaux de A . On dit que I_1 et I_2 sont des idéaux étrangers si $I_1 + I_2 = A$, autrement dit si on peut écrire $1 = x_1 + x_2$ avec $x_1 \in I_1$ et $x_2 \in I_2$.

On dit que deux éléments a et b de A sont étrangers si les idéaux qu'ils engendrent sont étrangers, ce qui revient à dire qu'il existe des éléments u et v de A tels que $au + bv = 1$.

Exemple. Soient $n_1, n_2 \in \mathbb{Z}$. Notons $I_1 = (n_1)$ et $I_2 = (n_2)$. Les idéaux I_1 et I_2 sont étrangers si et seulement si n_1 et n_2 sont premiers entre eux (voir la définition 2 de 12.12.2). Mais on ne confondra pas la notion d'éléments étrangers avec celle d'éléments premiers entre eux qui sera définie en 19.10.

Théorème chinois, cas de deux idéaux. Soient A un anneau, I_1 et I_2 deux idéaux étrangers de A . On a les résultats suivants.

- 1) $I_1 I_2 = I_1 \cap I_2$.
- 2) Soient $\pi_1 : A \rightarrow A/I_1$, $\pi_2 : A \rightarrow A/I_2$, $p_1 : A/I_1 \times A/I_2 \rightarrow A/I_1$, $p_2 : A/I_1 \times A/I_2 \rightarrow A/I_2$ les projections canoniques et $f = (\pi_1, \pi_2) : A \rightarrow A/I_1 \times A/I_2$, $f(x) = (\pi_1(x), \pi_2(x)) = (x \bmod I_1, x \bmod I_2)$, l'unique homomorphisme d'anneaux rendant commutatif le diagramme



Alors f est surjectif et son noyau est $I = I_1 I_2$, autrement dit, f permet de définir par passage au quotient un isomorphisme d'anneaux

$$\varphi : A/I \rightarrow A/I_1 \times A/I_2;$$

φ est défini par $x \bmod I \mapsto (x \bmod I_1, x \bmod I_2)$.

Démonstration. 1) On a déjà remarqué que $I_1 I_2 \subset I_1 \cap I_2$. Réciproquement, soit $x \in I_1 \cap I_2$ et notons $a_1 \in I_1$ et $a_2 \in I_2$ deux éléments de somme 1. On a $x = x(a_1 + a_2) = xa_1 + xa_2$, ce qui prouve que x est la somme de deux éléments de $I_1 I_2$; c'est donc un élément de $I_1 I_2$; par conséquent $I_1 I_2 = I_1 \cap I_2$.

2) Le noyau de f est l'ensemble des $x \in A$ tels que $\pi_1(x) = 0$ et $\pi_2(x) = 0$, ce qui équivaut à $x \in I_1 \cap I_2$, d'où $\ker(f) = I$ d'après le 1).

La propriété universelle du quotient permet de définir l'homomorphisme φ et on sait que φ est injectif. Reste à montrer qu'il est surjectif. Comme il existe $a_1 \in I_1$ et $a_2 \in I_2$ tels que $a_1 + a_2 = 1$, on a $f(a_2) = f(1 - a_1) = (1, 0)$ et $f(a_1) = f(1 - a_2) = (0, 1)$. Pour tous y_1, y_2 de A , on a donc $f(y_1 a_1 + y_2 a_2) = (\pi_1(y_1), \pi_2(y_2))$, ce qui montre la surjectivité de f et implique celle de φ . $\square$

Théorème chinois, cas général. Soient A un anneau, $(I_k)_{1 \leq k \leq n}$ une famille finie d'idéaux de A étrangers deux à deux (c'est-à-dire que, pour tous k, l de $\{1, \dots, n\}$, $I_k + I_l = A$).

1) Notons I le produit des idéaux I_k , $1 \leq k \leq n$; alors I est égal à l'intersection des idéaux I_k , $1 \leq k \leq n$:

$$I = I_1 \dots I_n = \bigcap_{1 \leq k \leq n} I_k.$$

2) Soient $\pi_k : A \rightarrow A/I_k$, $p_k : \prod_{1 \leq k \leq n} A/I_k \rightarrow A/I_k$, $1 \leq k \leq n$, les projections canoniques et $f = (\pi_k)_{1 \leq k \leq n} : A \rightarrow \prod_{1 \leq k \leq n} A/I_k$, $f(x) = (\pi_1(x), \dots, \pi_n(x)) = (x \bmod I_1, \dots, x \bmod I_n)$, l'unique homomorphisme d'anneaux tel que $p_k \circ f = \pi_k$ pour $1 \leq k \leq n$. Alors f induit un isomorphisme

$$\varphi : A/I \rightarrow \prod_{1 \leq k \leq n} A/I_k.$$

Démonstration. 1) L'inclusion $I_1 \dots I_n \subset I_1 \cap \dots \cap I_n$ est immédiate comme pour le cas de deux idéaux. Il faut montrer l'inclusion réciproque.

Raisonnons par récurrence sur n . Montrons d'abord que I_1 et $\prod_{2 \leq k \leq n} I_k$ sont étrangers. Comme I_1 et I_k sont étrangers pour $2 \leq k \leq n$, il existe, pour $2 \leq k \leq n$, des éléments $x_k \in I_1$ et $y_k \in I_k$ tels que $x_k + y_k = 1$. On a donc $\prod_{2 \leq k \leq n} (x_k + y_k) = 1$. Le membre de gauche est la somme x de termes où apparaît un x_k et qui sont dans I_1 et du produit $z_1 = y_2 \dots y_n$ qui est dans $I_2 \dots I_n$, ce qui prouve que I_1 et $\prod_{2 \leq k \leq n} I_k$ sont étrangers. Le cas $n = 2$ du théorème montre que $I_1(I_2 \dots I_n) = I_1 \cap (I_2 \dots I_n)$ et l'hypothèse de récurrence permet de conclure.

2) Le noyau de f est $I_1 \cap \dots \cap I_n = I$. D'où l'existence de φ et son injectivité. On remarque que $f(z_1) = (1, 0, \dots, 0)$. En suivant une démarche analogue au 1), on peut montrer que pour $k = 2, \dots, n$, il existe $z_2, \dots, z_n$ tels que $z_k = 1 \bmod I_k$ et $z_k = 0 \bmod I_l$ pour $l \neq k$. Cela permet de montrer, comme dans le cas où $n = 2$, que f est surjectif ; par conséquent, φ est surjectif. $\square$

Prenons maintenant le cas particulier de l'anneau $\mathbb{Z}$.

Théorème chinois, cas de $\mathbb{Z}$. Soient $a_1, \dots, a_n$ des entiers premiers deux à deux.

L'application $f : \mathbb{Z} \rightarrow \prod_{1 \leq k \leq n} (\mathbb{Z}/a_k \mathbb{Z})$ telle que $x \mapsto (x \bmod a_1, \dots, x \bmod a_n)$ définit par passage au quotient un isomorphisme $\varphi : (\mathbb{Z}/(a_1 \dots a_n) \mathbb{Z}) \rightarrow \prod_{1 \leq k \leq n} (\mathbb{Z}/a_k \mathbb{Z})$, $x \bmod a_1 \dots a_n \mapsto (x \bmod a_1, \dots, x \bmod a_n)$. $\square$

La condition sur les a_k d'être premiers entre eux est essentielle. Par exemple, on a $\mathbb{Z}/12\mathbb{Z} \simeq (\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$, mais on n'a pas $\mathbb{Z}/12\mathbb{Z} \simeq (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/6\mathbb{Z})$; en tant que groupe, ce dernier n'a pas d'élément d'ordre 12 et n'est pas cyclique.

Pour appliquer le théorème chinois, deux situations différentes se présentent. Dans la première, il est nécessaire de calculer l'inverse ψ de φ . Dans la seconde, il s'agit juste de trouver l'antécédent x d'un n -uplet $(x_1, \dots, x_n)$, autrement dit de résoudre le système $x = x_k \bmod a_k$, $1 \leq k \leq n$. Expliquons comment faire l'un et l'autre.

Exemple d'un cas simple de calcul de φ . Pour $n = 15$, le théorème chinois donne l'isomorphisme $\varphi : \mathbb{Z}/15\mathbb{Z} \rightarrow (\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/5\mathbb{Z})$ défini par $x \bmod 15 \mapsto (x \bmod 3, x \bmod 5)$, par exemple $\varphi(4) = (1, -1)$, $\varphi(7) = (1, 2)$. Il suffit de remarquer que $\varphi(10) = (1, 0)$ et que $\varphi(6) = (0, 1)$ pour décrire ψ : on a $\psi(a \bmod 3, b \bmod 5) = 10a + 6b \bmod 15$, par exemple, $\psi(2, 3) = 20 + 18 = 8 \bmod 15$.

Détermination de l'inverse de l'isomorphisme en général. Pour déterminer l'inverse $\psi : \prod_{1 \leq k \leq n} (\mathbb{Z}/a_k\mathbb{Z}) \rightarrow (\mathbb{Z}/(a_1 \dots a_n)\mathbb{Z})$ de l'isomorphisme φ , il suffit de déterminer des entiers $z_1, \dots, z_n$ tels que $z_k = 1 \bmod a_k$ et $z_k = 0 \bmod a_l$ pour $l \neq k$. Comment obtenir les z_k ? Prenons l'exemple de z_1 : comme a_1 et le produit $a_2 \dots a_n$ sont premiers entre eux, on sait calculer (voir exercice 12.3) des entiers u et v tels que $ua_1 + va_2 \dots a_n = 1$ et on prend $z_1 = va_2 \dots a_n$. Le calcul de $z_2, \dots, z_{n-1}$ suit la même démarche et on obtient z_n en remarquant que $z_1 + \dots + z_n = 1 \bmod a_1 \dots a_n$.

Exemple de Sunzi. Pour $n = 3 \times 5 \times 7 = 105$, l'isomorphisme ψ est défini par $\psi(1, 0, 0) = 70$, $\psi(0, 1, 0) = 21$, $\psi(0, 0, 1) = 15$ et on a

$$\psi(a \bmod 3, b \bmod 5, c \bmod 7) = 70a + 21b + 15c \bmod 105$$

comme l'affirmait Sunzi.

Résolution d'un système de congruences. Si le problème consiste à trouver l'image d'un élément de $(r_1, \dots, r_n) \in \prod_{1 \leq k \leq n} (\mathbb{Z}/a_k\mathbb{Z})$, il s'agit d'une simple résolution d'un système de congruences : trouver un entier x tel que $x = r_k \bmod a_k$. Il est alors inutile de construire complètement ψ pour trouver x . On cherche x sous la forme $x = s_1 + a_1 s_2 + (a_1 a_2) s_3 + \dots + (a_1 \dots a_{n-1}) s_n$ où $s_1, \dots, s_n$ sont des entiers qu'on détermine de proche en proche, ce qui est moins coûteux en temps de calcul et permet d'ajouter une congruence supplémentaire sans avoir besoin de refaire tous les calculs. Par exemple, cherchons x tel que :

$$\begin{cases} x = 2 \bmod 3 \\ x = 3 \bmod 5 \\ x = 2 \bmod 7 \end{cases}$$

On part de $x = s_1 + 3s_2 + 15s_3$. La première équation donne $s_1 = 2$, la seconde donne $2 + 3s_2 = 3 \bmod 5$, d'où $3s_2 = 1 \bmod 5$ et en multipliant par

$2 = 3^{-1} \bmod 5$, $s_2 = 2 \bmod 5$; on arrive alors à $8 + 15s_3 = 2 \bmod 7$, d'où $s_3 = 1$, puis $x = 23 \bmod 105$.

Interpolation de Lagrange. Nous avons expliqué le problème de l'interpolation de Lagrange dans l'exercice 6.8. Il consiste à trouver un polynôme prenant des valeurs données en des points donnés.

On se donne un corps K , $n \geq 1$ un entier, des éléments $a_1, \dots, a_{n+1}$ de K distincts deux à deux, des éléments $b_1, \dots, b_{n+1}$ de K ; on note $K_n[X]$ l'anneau des polynômes de $K[X]$ de degré inférieur ou égal à n . On cherche un polynôme $P \in K_n[X]$ tel que $P(a_k) = b_k$ pour $k = 1, \dots, n+1$.

Les polynômes $X - a_k$, $k = 1, \dots, n+1$, sont deux à deux étrangers entre eux puisque leurs différences sont des constantes non nulles donc inversibles. Notons T leur produit. Le théorème chinois donne un isomorphisme d'anneaux $K[X]/(T) \simeq \prod_{1 \leq k \leq n+1} K[X]/(X - a_k)$.

D'autre part, on sait (voir 19.4), que $K[X]/(X - a_k) \simeq K$, l'isomorphisme étant donné par l'évaluation en a_k .

On en déduit un isomorphisme d'anneaux $K[X]/(T) \simeq K^{n+1}$, défini par $P \bmod T \mapsto (P(a_1), \dots, P(a_{n+1}))$. Enfin, il existe un unique polynôme de $K_n[X]$ définissant la classe $P \bmod T$, d'où $K_n[X] \simeq K^{n+1}$. C'est ce dernier isomorphisme qui montre l'existence du polynôme d'interpolation de Lagrange : c'est le polynôme dont l'image par cet isomorphisme est $(b_1, \dots, b_{n+1})$. Pour trouver explicitement ce polynôme, nous avons donné une démarche dans l'exercice 6.8.

L'énoncé peut-être généralisé au cas de polynômes à coefficients dans un anneau, les $a_1, \dots, a_n$ étant tels que $a_i - a_j$ soit inversible pour $i \neq j$, ce qui permet de conserver la propriété pour les idéaux $(X - a_k)$ d'être étrangers deux à deux.

19.9 ÉLÉMENTS INVERSIBLES

Les éléments inversibles d'un anneau A sont souvent appelés unités de A . Les unités d'un anneau forment un groupe pour la multiplication de l'anneau. Trouver les éléments inversibles d'un anneau et la structure du groupe qu'ils forment n'est pas toujours un problème facile. Nous allons présenter quelques exemples.

La notation utilisée, en général, pour le groupe des éléments inversibles d'un anneau A est $A^\times$ qu'on ne confondra pas avec la notation A^* qui désigne $A - \{0\}$. Cependant, si A est un corps, les deux notations désignent le même ensemble. On trouve aussi la notation $U(A)$ pour $A^\times$.

Par exemple, les éléments inversibles de $\mathbb{Z}$ sont ± 1 , ce qu'on écrit $\mathbb{Z}^\times = \{-1, 1\}$ ou $U(\mathbb{Z}) = \{-1, 1\}$.

Éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$. Dans $\mathbb{Z}/n\mathbb{Z}$, les éléments inversibles sont les classes des entiers a premiers avec n . Nous avons vu en 12.12.3 que, si a et n sont premiers entre eux, il existe, d'après l'identité de Bézout, des entiers u et v tels que

$ua + vn = 1$, d'où $ua = 1 \pmod n$, ce qui prouve que a est inversible et donne son inverse ; si a et n ne sont pas premiers entre eux, on a $d = \text{pgcd}(a, n)$ avec $d > 1$ et il existe des entiers u et v tels que $a = ud$ et $n = vd$; comme $v \neq 0 \pmod n$ et $va = 0 \pmod n$, a n'est pas inversible dans $\mathbb{Z}/n\mathbb{Z}$. Le groupe des unités de $\mathbb{Z}/n\mathbb{Z}$ sera noté $U(n)$ et sa structure sera déterminée pour tout n dans l'exercice 21.5.

Éléments inversibles de $\mathbb{Z}[\sqrt{n}]$, cas $n < 0$. On suppose toujours que n est un entier sans facteur carré. Si $z \in \mathbb{Z}[\sqrt{n}]$ est inversible dans cet anneau, il existe $z' \in \mathbb{Z}[\sqrt{n}]$ tel que $zz' = 1$, donc $N(z)N(z') = 1$; comme $N(z)$ et $N(z')$ sont des entiers, on a donc $N(z) = \pm 1$. Réciproquement, si $N(z) = \pm 1$, alors $z\tilde{z} = \pm 1$ (avec $\tilde{z}$ défini en 19.4), donc z est inversible, d'inverse $\pm\tilde{z}$.

Les unités de $\mathbb{Z}[i]$ sont les nombres de la forme $a + bi$, $a, b \in \mathbb{Z}$ tels que $a^2 + b^2 = 1$, c'est-à-dire $\pm 1, \pm i$. Les unités de $\mathbb{Z}[\sqrt{-2}]$ sont les nombres de la forme $a + b\sqrt{-2}$ tels que $a^2 + 2b^2 = 1$, c'est-à-dire ± 1 ; il en est de même des unités de $\mathbb{Z}[\sqrt{-n}]$ pour $n \geq 2$.

Pour $n > 0$, le problème est plus difficile ; nous allons en dire quelques mots.

Équations de Pell-Fermat. Les unités de $\mathbb{Z}[\sqrt{2}]$ sont les nombres de la forme $a + b\sqrt{2}$ tels que $a^2 - 2b^2 = \pm 1$; on montre dans l'exercice 19.6 comment les déterminer toutes. Elles sont en nombre infini ; il en est de même de celles de $\mathbb{Z}[\sqrt{n}]$ pour tout entier $n \geq 2$ non carré, mais ce résultat n'a rien d'évident : c'est l'un des défis d'arithmétique lancés aux mathématiciens par Fermat en février 1657, sous la forme :

Tout nombre non quarré est de telle nature qu'on peut trouver infinis quarrés par lesquels si vous multipliez le nombre donné et si vous ajoutez l'unité au produit, vienne un quarré.

Autrement dit, si d n'est pas un carré d'entier, il existe une infinité d'entiers b tels que $db^2 + 1$ soit un carré a^2 , ce qu'on peut écrire $a^2 - db^2 = 1$: le problème revient pour nous à celui de montrer qu'il existe une infinité d'unités dans $\mathbb{Z}[\sqrt{n}]$, mais évidemment, ce n'était pas le langage de Fermat.

Les équations en nombres entiers $a^2 - db^2 = 1$ sont connues sous le nom d'équations de Pell, du nom de John Pell (1611-1685), mathématicien anglais auquel Euler les avait attribuées en 1730, mais Pell n'a laissé aucun document permettant de penser qu'il les ait étudiées. Certains disent aujourd'hui pour être un peu moins injustes envers Fermat : équations de Pell-Fermat, en conservant malgré tout le nom de Pell, pour se faire comprendre.

La résolution complète de ces équations, à l'aide des fractions continues, est exposée par Lagrange dans la revue turinoise qu'il avait contribué à créer, les *Mélanges de Turin*, tome IV, 1766-1769.

Les mathématiciens indiens Aryabhata (vers 476-550), Brahmagupta (vers 598-665), Bhaskara (1114-1185) avaient élaboré une méthode pour obtenir des solutions des équations $a^2 - db^2 = 1$, en partant de solutions de $a^2 - db^2 = m$ et en cherchant à diminuer le m jusqu'à 1. Ils l'appelaient le *cakravala* (la méthode de la roue :

cakra en sanscrit) ; le livre, passionnant, d'André Weil : *Number theory, an approach through history from Hammurapi (Hammourabi) to Legendre* (Birkhäuser, 1987), en donne une description.

Éléments inversibles d'un anneau de polynômes. Dans l'anneau $K[X]$ des polynômes sur un corps K , les seuls éléments inversibles sont les constantes non nuls puisque, si P est inversible d'inverse Q , $PQ = 1$ implique $\deg(P) + \deg(Q) = \deg(PQ) = 0$, donc $\deg(P) = 0$. De même, si A est un anneau intègre, les seuls éléments inversibles de l'anneau $A[X]$ sont les éléments inversibles de A .

19.10 DIVISIBILITÉ DANS LES ANNEAUX INTÈGRES

Nous généralisons à un anneau intègre quelconque des notions définies aux chapitres 12 et 13 dans le cadre des anneaux $\mathbb{Z}$ et $K[X]$.

Diviseurs, éléments associés. Soient a et b deux éléments d'un anneau intègre A . On dit que b divise a (ou que b est un diviseur de a , ou que a est multiple de b) s'il existe $c \in A$ tel que $a = bc$. On note $b|a$.

La notion de divisibilité dépend de l'anneau : 5 ne divise pas 7 dans $\mathbb{Z}$ mais le divise dans un anneau plus gros comme $\mathbb{Q}$. On doit préciser : *b divise a dans A* , chaque fois que le contexte l'exige.

Une unité divise tout élément d'un anneau. Tout élément d'un anneau divise 0, mais 0 ne divise aucun élément non nul de A .

Rappelons (voir 12.11) que la notion de divisibilité peut s'exprimer en termes d'idéaux : $b|a$ équivaut à $(b) \supset (a)$.

Deux éléments a et b d'un anneau A sont dits associés s'il existe une unité u de A telle que $a = ub$. Il est immédiat qu'alors $b = va$, avec $v = u^{-1}$ unité de A . On vérifie que la relation d'association est une relation d'équivalence.

Proposition 1. Soient A un anneau intègre, (a) et (b) les idéaux engendrés par deux éléments non nuls de A . On a $(a) = (b)$ si et seulement si a et b sont associés.

Démonstration. S'il existe une unité u tel que $a = ub$, on a $(a) = \{xa, x \in A\} = \{xub, x \in A\} \subset (b)$; comme $b = u^{-1}a$, on a de même $(b) \subset (a)$, donc $(a) = (b)$. Réciproquement, l'égalité $(a) = (b)$ implique l'existence de $c, d \in A$ tels que $a = cb$ et $b = da$; on a donc $a = cda$, donc $(1 - cd)a = 0$; comme $a \neq 0$, on a $cd = 1$; c est donc inversible. $\square$

Dans un anneau A , si a est associé à a' et si b est associé à b' , alors $b|a$ si et seulement si $b'|a'$. Les propriétés de divisibilité seront analogues pour des éléments associés.

Dans $\mathbb{Z}$, les éléments inversibles étant ± 1 , les éléments associés à un entier n non nul sont : $\pm n$. Si $a|b$ et si $b|a$, alors $a = \pm b$.

Dans $\mathbb{Z}[\iota]$, les éléments inversibles étant $\pm 1, \pm \iota$, les éléments associés à un entier n non nul sont : $\pm n, \pm \iota n$. Si $a|b$ et si $b|a$ dans $\mathbb{Z}[\iota]$, alors $a = \pm b$ ou $a = \pm \iota b$.

Plus petit commun multiple. Soient A un anneau intègre, a et b deux éléments de A . On dit que m est un plus petit commun multiple, en abrégé ppcm, de a et b si $(a) \cap (b) = (m)$.

La notion de ppcm n'est définie qu'à un élément inversible près. On écrit souvent $m = \text{ppcm}(a, b)$, mais, en général, les égalités $m = \text{ppcm}(a, b)$ et $m' = \text{ppcm}(a, b)$ ne permettent pas de conclure que $m = m'$. Si on a un moyen de choisir un élément parmi ses associés, par exemple, un élément positif dans $\mathbb{Z}$, un polynôme unitaire dans $K[X]$, on aura tendance à en parler comme *du* ppcm des deux éléments a et b .

La notion de ppcm se généralise à une famille finie $(a_1, \dots, a_n)$, $n \geq 2$, d'éléments de A : m est un ppcm de la famille si $(m) = (a_1) \cap \dots \cap (a_n)$. On écrit souvent $m = \text{ppcm}(a_1, \dots, a_n)$, avec le danger que nous venons de signaler.

Proposition 2. *Soit A un anneau intègre dans lequel les ppcm de deux éléments quelconques existent. Alors toute famille finie $(a_1, \dots, a_n)$, $n \geq 2$, d'éléments de A admet un ppcm.*

Démonstration. On définit par récurrence $\text{ppcm}(a_1, \dots, a_n)$ en remarquant que $(a_1) \cap \dots \cap (a_n) = ((a_1) \cap \dots \cap (a_{n-1})) \cap (a_n)$, donc $\text{ppcm}(a_1, \dots, a_n) = \text{ppcm}(\text{ppcm}(a_1, \dots, a_{n-1}), a_n)$. $\square$

Plus grand commun diviseur. Soit A un anneau intègre. Si $x \in A$, nous noterons $D(x)$ l'ensemble des diviseurs de x .

Soient a et b deux éléments de A . On dit que d est un plus grand commun diviseur, en abrégé pgcd, de a et b si $D(a) \cap D(b) = D(d)$. Autrement dit, d est un diviseur de a et de b et c'est le plus grand, au sens de la relation de divisibilité, des diviseurs communs à a et b .

Le pgcd n'existe pas toujours, voir l'exemple de $\mathbb{Z}[\sqrt{-5}]$ ci-dessous.

La notion de pgcd se généralise à une famille finie $(a_1, \dots, a_n)$, $n \geq 2$, d'éléments de A : d est un pgcd de la famille si $D(d) = D(a_1) \cap \dots \cap D(a_n)$.

La notion de pgcd est définie à élément inversible près. On écrit souvent $d = \text{pgcd}(a, b)$, $d = \text{pgcd}(a_1, \dots, a_n)$, on parle *du* pgcd, avec les mêmes dangers que ceux signalés ci-dessus pour le ppcm.

Proposition 3. *Soit A un anneau intègre dans lequel les pgcd de deux éléments quelconques existent. Alors toute famille finie $(a_1, \dots, a_n)$, $n \geq 2$, d'éléments de A admet un pgcd.*

Démonstration. Il suffit de raisonner comme ci-dessus pour le ppcm. $\square$

Éléments premiers entre eux. Des éléments a et b d'un anneau intègre A sont dits premiers entre eux si 1 est un pgcd de a et b , autrement dit, si les seuls diviseurs communs à a et b sont les unités de A .

Par exemple, dans $\mathbb{Z}[X]$, 2 et X sont premiers entre eux. On notera qu'il n'existe pas de polynômes U et V de $\mathbb{Z}[X]$ tels que $2U + XV = 1$, puisque le terme constant de ces polynômes est pair. Ces éléments ne sont donc pas étrangers.

Éléments irréductibles. Un élément a d'un anneau A est dit irréductible (on dit aussi extrémal) dans A si a n'est pas inversible et si une relation de la forme $a = bc$ implique que b ou c est une unité de A . Autrement dit, a est dit irréductible si a n'est pas une unité de A et si les seuls diviseurs de a sont les éléments associés à a et les unités de A . On dit que a est réductible dans le cas contraire.

On remarquera que 0 n'est pas irréductible ($0 = 0 \times 0$ et aucun des deux facteurs n'est une unité) et qu'un élément associé à un élément irréductible est aussi irréductible.

Dans un corps, tout élément est réductible.

Dans $\mathbb{Z}[\sqrt{n}]$, comme $N(z z') = N(z)N(z')$, tout élément ayant un nombre premier pour norme est irréductible, mais cette condition n'est pas nécessaire (voir l'exemple de $\mathbb{Z}[\sqrt{-5}]$ ci-dessous). On remarquera aussi que la notion d'irréductibilité dépend de l'anneau dans lequel on se place : un élément irréductible dans un anneau peut ne plus l'être dans un anneau plus gros. Par exemple, 2 est irréductible dans $\mathbb{Z}$, mais n'est pas irréductible dans $\mathbb{Z}[\sqrt{2}]$ et est une unité dans $\mathbb{Q}$.

La détermination des éléments irréductibles d'un anneau n'est pas toujours aisée, comme on le verra, par exemple, aux chapitres 20 et 21 pour des anneaux de polynômes. Dans $\mathbb{Z}$, les éléments irréductibles sont les entiers $\pm p$ où p est un nombre premier.

Système de représentants des éléments irréductibles. Soient A un anneau et E l'ensemble de ses éléments irréductibles. La relation d'association est une relation d'équivalence $\mathcal{R}$ sur E . Un sous-ensemble S de E constitué d'un élément irréductible par classe d'équivalence (dans le cas général, l'existence de S résulte de l'axiome de choix) est appelé système de représentants des éléments irréductibles de A .

Dans $\mathbb{Z}$, les éléments irréductibles sont de la forme $\pm p$ où p est un nombre premier et on peut choisir comme système de représentants les nombres premiers positifs ; dans des anneaux de polynômes, on pourra prendre des polynômes unitaires, mais dans d'autres anneaux, le choix des représentants ne peut se faire ainsi.

Exemple de $\mathbb{Z}[\sqrt{-5}]$. Dans cet anneau, la norme est définie par $N(a + b\sqrt{-5}) = a^2 + 5b^2$. Il n'existe donc pas d'éléments de norme 2, puisque $a^2 + 5b^2 = 2$ impose $b = 0$ et que $a^2 = 2$ est impossible dans $\mathbb{Z}$, pas d'éléments de norme 3, puisque $a^2 = 3$ est impossible dans $\mathbb{Z}$, pas d'éléments de norme 12, puisque $a^2 + 5b^2 = 12$ impose $b = 0$ ou $b = 1$, etc.

On en déduit que 2 est irréductible, bien que sa norme ne soit pas un nombre premier ; en effet, $N(2) = 4$ et si $2 = \alpha\beta$, on a $4 = N(\alpha)N(\beta)$; si α et β ne sont pas inversibles, c'est que $N(\alpha) = N(\beta) = 2$, ce qui est impossible. De même, $1 + \sqrt{-5}$, de norme 6, est irréductible.

Dans l'anneau $\mathbb{Z}[\sqrt{-5}]$, le pgcd de $\alpha = 6$ et de $\beta = 2(1 + \sqrt{-5})$ n'existe pas. En effet, les égalités $6 = 2 \times 3$ et $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, montrent que 2 et $1 + \sqrt{-5}$ divisent α et β . Si $d = \text{pgcd}(\alpha, \beta)$ existait, il serait multiple de 2 et $1 + \sqrt{-5}$ et diviserait 6 et

β . Sa norme serait donc un multiple de $N(2) = 4$, de $N(1 + i\sqrt{5}) = 6$, diviserait $N(\alpha) = 36$ et $N(\beta) = 24$; elle serait égale à 12, ce qui est impossible.

19.11 IDÉAUX PREMIERS ET MAXIMAUX

Un idéal I d'un anneau A sera dit strict si $I \subsetneq A$.

Les propriétés des idéaux peuvent servir pour connaître *a priori* des propriétés des quotients, comme on va le voir.

Idéal premier, première définition. Un idéal P d'un anneau A est dit premier si A/P est un anneau intègre.

Idéal premier, seconde définition. Un idéal P d'un anneau A est dit premier si P est un idéal strict de A (soit $P \subsetneq A$) et si, pour tous $x, y \in A$, $xy \in P$ implique $x \in P$ ou $y \in P$.

Équivalence des deux définitions. Notons $\pi : A \rightarrow A/P$ la projection canonique. Si A/P est intègre, A/P n'est pas l'anneau nul (voir 12.8), donc $P \neq A$; de plus, si $xy \in P$, alors $\pi(x)\pi(y) = \pi(xy) = 0$, donc $\pi(x) = 0$ ou $\pi(y) = 0$, ce qui implique $x \in P$ ou $y \in P$. Réciproquement, si P vérifie la seconde définition, A/P n'est pas l'anneau nul et, si $\alpha = \pi(x), \beta = \pi(y)$ ont un produit nul dans A/P , alors $xy \in P$ donc $x \in P$ ou $y \in P$; par conséquent, α ou β est nul et A/P est intègre.

On peut encore dire que, si P est un idéal premier d'un anneau A , le produit d'éléments de A n'appartenant pas à P n'appartient pas à P .

Idéal maximal, première définition. Un idéal M d'un anneau A est dit maximal si A/M est un corps.

Idéal maximal, seconde définition. Un idéal M d'un anneau A est dit maximal s'il est maximal au sens de l'inclusion parmi les idéaux stricts de A , autrement dit si $M \subsetneq A$ et si les seuls idéaux contenant M sont M et A .

Équivalence des deux définitions. Notons $\pi : A \rightarrow A/M$ la projection canonique. Si A/M est un corps, on a $M \neq A$ et, si I est un idéal de A contenant M , $\pi(I)$ est un idéal du corps A/M , donc $\pi(I) = A/M$ ou $\pi(I) = (0)$, soit $I = A$ ou $I = M$, ce qui montre que M est un idéal maximal de A au sens de la seconde définition. Réciproquement, si M est maximal au sens de la seconde définition, A/M est un anneau non nul sans idéal autre que (0) et A/M . Pour tout $x \in A/M$ non nul, l'idéal (x) est égal à A/M , ce qui implique que x est inversible ; donc A/M est un corps.

Idéaux premiers et maximaux de $\mathbb{Z}$. On sait que les idéaux de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$ et qu'on peut supposer $n \geq 0$.

Comme les idéaux premiers et maximaux sont des idéaux stricts, un idéal premier ou maximal de $\mathbb{Z}$ est de la forme $n\mathbb{Z}$ avec $n \neq \pm 1$.

Si $n = 0$, $\mathbb{Z}/(0) = \mathbb{Z}$ est un anneau intègre mais n'est pas un corps, donc l'idéal (0) est un idéal premier non maximal de $\mathbb{Z}$ (alors que 0 n'est pas un nombre premier).

Si $n > 1$ n'est pas premier, on peut écrire $n = ab$ avec $a, b \notin n\mathbb{Z}$, ce qui prouve que l'idéal $n\mathbb{Z}$ n'est ni premier ni maximal.

Si p est un nombre premier, on sait (voir 12.13) que $\mathbb{Z}/p\mathbb{Z}$ est un corps, ce qui prouve que l'idéal $p\mathbb{Z}$ est maximal et premier.

Exemple dans $\mathbb{Z}[X]$. Considérons l'anneau $\mathbb{Z}[X]$. L'idéal (X) est premier car l'anneau quotient $\mathbb{Z}[X]/(X) \simeq \mathbb{Z}$ est intègre. Ce n'est cependant pas un idéal maximal car il est contenu dans l'idéal $(2, X)$. Ce dernier idéal est maximal car $\mathbb{Z}[X]/(2, X) \simeq (\mathbb{Z}[X]/(X))/(2) \simeq \mathbb{Z}/2\mathbb{Z}$ qui est un corps.

Propriétés des idéaux premiers et maximaux

- 1) (0) est un idéal premier si et seulement si A est intègre ;
- 2) (0) est un idéal maximal si et seulement si A est un corps ;
- 3) un idéal maximal est premier ;
- 4) un idéal premier non nul d'un anneau principal (voir 12.9) est maximal ;
- 5) si $f : A \rightarrow A'$ est un homomorphisme d'anneaux et si P est un idéal premier de A' , alors $f^{-1}(P)$ est un idéal premier de A .

Les propriétés 1, 2, 3, 5 sont faciles à vérifier.

Pour montrer la propriété 4, si P est un idéal premier non nul d'un anneau principal A et I un idéal tel que $P \subsetneq I$, il existe des éléments a, b tels que $P = (a)$ avec $a \neq 0$, $I = (b)$ et $b \notin P$. Comme $P \subset I$, il existe $c \in A$ tel que $a = bc$; comme P est premier, $c \in P$; il existe donc $d \in A$ tel que $c = da$; on en déduit $a = bda$; comme a est non nul, on a donc $bd = 1$, b est une unité de A , d'où $I = A$; donc P est maximal.

On notera que l'image réciproque d'un idéal maximal n'est pas nécessairement un idéal maximal, comme le montre l'exemple de l'idéal nul de $\mathbb{Z}$, image réciproque de l'idéal nul de $\mathbb{Q}$ par l'inclusion $\mathbb{Z} \rightarrow \mathbb{Q}$.

Proposition 1 : correspondance entre idéaux premiers et maximaux d'un anneau et d'un quotient. Soient A un anneau et I un idéal de A . On a vu (en 19.5) qu'il existait une correspondance bijective Φ entre idéaux de A contenant I et idéaux de A/I . Cette correspondance induit des correspondances bijectives entre :

- 1) idéaux premiers de A contenant I et idéaux premiers de A/I ;
- 2) idéaux maximaux de A contenant I et idéaux maximaux de A/I .

Démonstration. Notons $\pi : A \rightarrow A/I$ la projection canonique. On a vu que, si $\mathfrak{J}$ est un idéal de A tel que $\mathfrak{J} \supset I$, on a $\Phi(\mathfrak{J}) = \pi(\mathfrak{J})$ et $(A/I)/\pi(\mathfrak{J}) \simeq A/\mathfrak{J}$. Donc, si $\mathfrak{J}$ est un idéal premier de A , $A/\mathfrak{J}$ est intègre, donc $(A/I)/\pi(\mathfrak{J})$ est intègre, donc $\pi(\mathfrak{J})$ est un idéal premier de A/I ; si $\mathfrak{J}$ est un idéal maximal de A , $A/\mathfrak{J}$ est un corps, donc $\pi(\mathfrak{J})$ est un idéal maximal de A/I . Les réciproques sont claires. $\square$

Existence d'idéaux maximaux

Théorème. Tout anneau non nul admet un idéal maximal.

Démonstration. Tout anneau non nul admet des idéaux stricts, comme (0) , par conséquent, l'ensemble ε des idéaux stricts de A est non vide. Considérons une partie $\mathcal{F}$ de ε totalement ordonnée par inclusion. On montre (comme en 19.7) que l'union des idéaux de $\mathcal{F}$ est un idéal de A . Cet idéal est un majorant (au sens de la relation d'ordre définie par l'inclusion) de l'ensemble des idéaux de $\mathcal{F}$. On dit que l'ensemble ε est un ensemble ordonné inductif (ce qui signifie que toute partie totalement ordonnée est majorée). Pour justifier l'existence d'un élément maximal de ε , ce qui termine la démonstration du théorème, il suffit alors d'invoquer une conséquence de l'axiome de choix (voir 18.1), le résultat de théorie des ensembles suivant, établi par Max Zorn (1906-1993) en 1935.

Lemme de Zorn. *Tout ensemble ordonné inductif admet un élément maximal.* $\square$

Ce résultat est nécessaire pour démontrer un certain nombre de résultats pour des ensembles infinis (voir exercice 19.14).

Corollaire. *Dans un anneau A :*

- 1) *tout idéal strict de A est contenu dans un idéal maximal ;*
- 2) *tout élément non inversible de A appartient à un idéal maximal de A .*

Démonstration. 1) Soient I un idéal de A et $\pi : A \rightarrow A/I$ la projection canonique. L'image réciproque par π d'un idéal maximal de A/I répond à la question.
2) Si x est non inversible, l'idéal (x) est distinct de A et on peut appliquer ce qui précède à l'idéal (x) . $\square$

Éléments premiers. Un élément non nul a d'un anneau A est dit premier si l'idéal (a) est premier, autrement dit, pour tout $b, c \in A$, $a \mid bc$ implique $a \mid b$ ou $a \mid c$.

Dans $\mathbb{Z}$, les éléments premiers sont les nombres p premiers.

Proposition 2. *Soient A un anneau intègre et $a \in A$. Si a est un élément premier de A , alors a est un élément irréductible de A .*

Démonstration. Si a est un élément premier de A et si $a = bc$, a divise b ou c , par exemple b . Il existe d tel que $b = ad$, d'où $a = acd$, $(1 - cd)a = 0$; comme a est non nul, c est une unité et a est donc irréductible.

Exemple. La réciproque est fausse. L'exemple suivant se trouve, avec des variantes, dans beaucoup de livres. Dans l'anneau $A = \mathbb{Z}[\sqrt{-5}]$ déjà étudié ci-dessus, les facteurs des deux membres de l'égalité $2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ sont irréductibles, non associés, non premiers (utiliser la norme). La propriété suivante permet de conclure que l'anneau $\mathbb{Z}[\sqrt{-5}]$ n'est pas principal.

Proposition 3. *Dans un anneau principal A , tout élément irréductible engendre un idéal maximal et est premier.*

Démonstration. Soit a un élément irréductible de A et supposons que l'idéal (a) ne soit pas maximal dans A . Il existe alors un idéal maximal de A contenant strictement (a) ; cet idéal est principal ; notons d un élément qui l'engendre. On a $A \supsetneq (d) \supsetneq (a)$, ce qui implique $d|a$, ce qui est impossible (d n'est pas une unité car (d) est distinct de A , ni un élément associé à a car (d) est distinct de (a)). Par conséquent, (a) est un idéal maximal et a est un élément premier. $\square$

Les propositions précédentes montrent que dans un anneau principal, les notions d'élément premier et d'élément irréductible coïncident.

19.12 ANNEAUX EUCLIDIENS

On peut faire des divisions euclidiennes dans les anneaux $\mathbb{Z}$ et $K[X]$ où K est un corps. Cette division permet de montrer que ces anneaux sont principaux. Nous allons généraliser cette situation.

Définition. Un anneau euclidien est un couple (A, j) où A est un anneau intègre et où $j : A^* \rightarrow \mathbb{N}$ est une fonction, appelée jauge euclidienne (certains disent (joliment ?) stathme, du grec $\sigma\tau\alpha\theta\mu\omicron\varsigma$: mesure), telle que, pour tous $a, b \in A$ avec b non nul, il existe des éléments q, r dans A tels que $a = bq + r$ avec $r = 0$ ou $j(r) < j(b)$.

La détermination de q et r est appelée division euclidienne de a par b dans l'anneau euclidien ; q s'appelle un quotient, r un reste de la division euclidienne. Ils ne sont pas forcément uniques.

Par exemple, $\mathbb{Z}$ muni de la fonction valeur absolue $j : \mathbb{Z} \rightarrow \mathbb{N}$, $j(n) = |n|$, est un anneau euclidien, comme on l'a déjà vu. Mais si on ne demande pas que le reste soit positif ou nul, on peut écrire de deux façons la division euclidienne de 18 par 5 :

$$18 = 5 \times 3 + 3 ;$$

$$18 = 5 \times 4 - 2 ;$$

dans les deux cas, la valeur absolue du reste est inférieure à 5.

On pourrait utiliser d'autres jauges euclidiennes sur $\mathbb{Z}$, comme $n \mapsto |n|^k$ pour $k \geq 1$.

Dans ce cadre général, on conserve le résultat important suivant.

Proposition. *Un anneau euclidien est principal.*

Démonstration. Soient (A, j) un anneau euclidien et I un idéal de A . Si $I = (0)$, I est principal. Si $I \neq (0)$, soit b un élément non nul de I de jauge euclidienne minimale. On montre que $I = (b)$ comme pour l'anneau $\mathbb{Z}$ (voir 12.10). $\square$

Un exemple classique d'anneau principal non euclidien est l'anneau $\mathbb{Z}[\sqrt{-19}]$, décrit dans l'exercice 19.9.

Recherche du pgcd et identité de Bézout dans les anneaux euclidiens. Soient a et b deux éléments d'un anneau euclidien (A, j) . La méthode de divisions successives pour obtenir le pgcd vue en 12.12.1 pour $\mathbb{Z}$ et en 13.4 pour un anneau de polynômes se généralise, puisque dans chaque division la jauge euclidienne du reste diminue ; comme ses valeurs sont entières, elle devient donc nécessairement nulle au bout d'un nombre fini de divisions.

Anneau $\mathbb{Z}[\iota]$. On va utiliser la fonction norme $N : \mathbb{Q}[\iota] \rightarrow \mathbb{Q}$ définie par $N(a + \iota b) = a^2 + b^2$, ce qui revient à $N(z) = z\bar{z} = |z|^2$. Soient $\alpha, \beta \in \mathbb{Z}[\iota]$, avec $\beta \neq 0$. Le quotient α / β s'écrit $x + y\iota$, avec $x, y \in \mathbb{Q}$. Notons a, b les entiers les plus proches de x, y respectivement et posons $q = a + b\iota$, $r = \alpha - q\beta = (x - a + (y - b)\iota)\beta$. On a $|x - a| \leq 1/2$, $|y - b| \leq 1/2$, donc $|N(x - a + (y - b)\iota)| \leq 1/2$. D'où $N(r) < N(\beta)$, ce qui prouve que N définit par restriction une jauge euclidienne N' sur $\mathbb{Z}[\iota]$ et que $(\mathbb{Z}[\iota], N')$ est euclidien. On en déduit que $\mathbb{Z}[\iota]$ est principal.

19.13 ANNEAUX FACTORIELS

La proposition 4 de 12.12.3 énonçait l'existence et l'unicité de la décomposition de tout entier de $\mathbb{Z}$ en produit de facteurs premiers. Ce résultat peut être généralisé à certains anneaux appelés factoriels.

Définition : décomposition unique en produit de facteurs irréductibles. Soit A un anneau intègre. On dit qu'un élément non nul a de A admet une décomposition unique en produit de facteurs irréductibles si les deux conditions suivantes sont satisfaites.

- 1) Il existe une unité $u \in A$, un entier $r \geq 0$, des éléments irréductibles $p_1, \dots, p_r$ de A , des entiers $k_1, \dots, k_r \geq 1$ tels que

$$a = up_1^{k_1} \cdots p_r^{k_r}.$$

- 2) Si a admet une seconde décomposition du même type : $a = vq_1^{l_1} \cdots q_s^{l_s}$, alors $r = s$ et il existe une permutation σ des indices telle qu'on ait $p_{\sigma(i)}$ associé à q_i et $k_{\sigma(i)} = l_i$ pour $i = 1, \dots, r$.

Si on a choisi un système S de représentants des éléments irréductibles de A , la définition précédente s'énonce : il existe une unité $u \in A$ et une famille unique $(k_p)_{p \in S}$ d'entiers, nuls sauf un nombre fini d'entre eux, tels que

$$a = u \prod_{p \in S} p^{k_p}.$$

Si une telle décomposition de a existe, k_p est le plus grand des entiers k tels que p^k divise a .

La décomposition en produit de facteurs irréductibles d'une unité u de A est elle-même et alors $r = 0$.

Définition : anneau factoriel. Un anneau A est dit factoriel s'il est intègre et si tout élément non nul admet une décomposition unique en produit de facteurs irréductibles.

L'anneau $\mathbb{Z}$ est factoriel, c'est l'exemple de base.

C'est Nicolas Bourbaki qui choisit le nom de factoriel.

Définition : valuation. Soit A un anneau factoriel et S un système de représentants des éléments irréductibles de A . Tout élément irréductible p de S définit une fonction $v_p : A^* \rightarrow \mathbb{N}$ qui associe à un élément $a \in A^*$ l'exposant $v_p(a)$ de p dans la décomposition de a en produit de puissances d'éléments de S . Cet exposant est appelé valuation de a en p . Pour tout a non nul de A , il existe une unité u de A telle que :

$$a = u \prod_{p \in S} p^{v_p(a)}.$$

Propriétés d'un anneau factoriel. Soient a, b, c des éléments non nuls d'un anneau factoriel A , S un système de représentants des irréductibles de A ; on peut écrire a, b, c comme produits des mêmes facteurs irréductibles de S , quitte à écrire des puissances nulles : $a = up_1^{k_1} \dots p_r^{k_r}, b = vp_1^{l_1} \dots p_r^{l_r}, c = wp_1^{m_1} \dots p_r^{m_r}, k_i, l_i, m_i \geq 0$ pour $i = 1, \dots, r$ avec u, v, w unités de A .

Les propriétés suivantes de $\mathbb{Z}$ se généralisent aux anneaux factoriels.

- 1) a divise b si et seulement si $k_i \leq l_i$ pour $i = 1, \dots, r$.
- 2) $\text{pgcd}(a, b) = p_1^{\inf\{k_1, l_1\}} \dots p_r^{\inf\{k_r, l_r\}}$ (rappelons que le pgcd est défini à une unité près).
- 3) $\text{ppcm}(a, b) = p_1^{\sup\{k_1, l_1\}} \dots p_r^{\sup\{k_r, l_r\}}$ (le ppcm est aussi défini à une unité près).
- 4) Si a et b sont étrangers, alors $k_i > 0$ impose $l_i = 0$ et $l_i > 0$ impose $k_i = 0$.
- 5) Si a et b sont étrangers et si $a|bc$, alors $a|c$.
- 6) Si a et b sont étrangers, si $a|c$ et si $b|c$, alors $ab|c$.
- 7) Si a et b sont étrangers et si ab est une puissance n -ième (carré, cube, etc.) alors a et b sont tous deux des puissances n -ièmes.

Proposition. Dans un anneau factoriel, un élément est irréductible si et seulement si il est premier.

Démonstration. Soient A un anneau factoriel et a un élément irréductible de A . Si a divise bc , il existe d tel que $ad = bc$. En écrivant les décompositions en produit de facteurs irréductibles de ad et de bc , on voit que a divise b ou c .

Pour la réciproque, il suffit d'appliquer la proposition 2 de 19.11 (puisque un anneau factoriel est intègre). $\square$

Exemples d'anneaux factoriels et d'anneaux non factoriels. Nous avons déjà dit que l'anneau $\mathbb{Z}$ est un anneau factoriel ; de même, les anneaux $\mathbb{Z}[\sqrt{-2}]$, $\mathbb{Z}[\iota]$, $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[\sqrt{3}]$ (voir exercice 19.8) et les anneaux de polynômes $A[X]$, $A[X_1, \dots, X_n]$ si l'anneau A est lui-même factoriel (voir 20.3).

L'égalité $2 \times 3 = (1 + \iota\sqrt{5})(1 - \iota\sqrt{5})$ (voir 19.11) où les termes des deux produits sont irréductibles et non associés montre que l'anneau $\mathbb{Z}[\iota\sqrt{5}]$ n'est pas factoriel.

Théorème. *Un anneau principal est factoriel.*

Démonstration. Soit A un anneau principal. Par définition, A est intègre. Montrons d'abord l'existence d'une décomposition en produit de facteurs irréductibles (DPFI) de tout élément de A .

Commençons par montrer le lemme suivant : une suite croissante $(I_k)_{k \geq 0}$ d'idéaux de A est stationnaire (autrement dit constante à partir d'un certain rang). En effet, l'union des idéaux de cette suite est un idéal et cet idéal est principal ; son générateur d est élément de l'un des (I_k) donc $(d) \subset I_k \subset (d)$, ce qui donne $I_l = (d)$ pour $l \geq k$.

Considérons alors le sous-ensemble T de l'ensemble des idéaux stricts de A engendrés par des éléments n'admettant pas de DPFI. Si T est vide, il n'y a rien à montrer. Supposons T non vide. On peut ordonner T par la relation d'inclusion entre idéaux. Le lemme montre que T est un ensemble ordonné inductif. L'axiome de choix implique qu'il admet un élément maximal. C'est un idéal (a) tel que a n'admette pas de DPFI et qui n'est contenu strictement dans aucun autre idéal de T . Comme a n'est pas irréductible, il existe b et c non inversibles tels que $a = bc$ et les idéaux (b) et (c) contiennent strictement (a) . Par conséquent, b et c admettent des DPFI, alors que leur produit $bc = a$ n'en admet pas. Cette contradiction montre l'existence d'une DPFI de tout élément de A .

Montrons maintenant l'unicité de la décomposition. Supposons que $a = up_1^{k_1} \dots p_r^{k_r} = vp_1^{l_1} \dots p_r^{l_r}$ avec des k_i, l_i pas tous égaux. Effectuons toutes les simplifications possibles. S'il reste dans le membre de gauche un facteur p irréductible, il divise le produit de facteurs irréductibles de droite ; comme il est premier, il en divise l'un des termes qui est donc associé à p , contrairement aux simplifications effectuées. Il ne peut donc rester aucun facteur irréductible dans l'un des membres, ce qui montre l'unicité. $\square$

En particulier, un anneau euclidien, étant principal (voir 19.12), est factoriel.

L'importance de l'argument sur les suites croissantes stationnaires remonte à Emmy Noether.

Emmy Noether (1882-1935). Quelles raisons a-t-on pu imaginer, quelles raisons imagine-t-on encore pour faire des discriminations entre hommes et femmes en sciences et dans bien d'autres domaines ! Aujourd'hui, un équilibre naturel n'est pas encore atteint. Nous avons déjà un peu parlé de Sophie Germain qui a dû prendre

une identité masculine pour réussir à se faire entendre par les mathématiciens de son époque. De rares femmes ont réussi à l'imiter au dix-neuvième siècle ; les bancs des facultés des sciences étaient alors réservés aux hommes. Emmy Noether était la fille du mathématicien Max Noether (1844-1921), spécialiste de géométrie algébrique. Ses travaux vont influencer profondément les mathématiques du vingtième siècle. Elle est à l'origine d'une sorte d'immense nettoyage qui met en lumière les idées de base abstraites et essentielles en algèbre : notion d'idéal, d'anneaux noethériens, etc., en topologie : notion de groupe d'homologie, etc., en physique : le théorème de Noether en théorie de la relativité. Dans les années 1920, de nombreux mathématiciens viennent la rencontrer à Göttingen ; son influence est profonde sur Emil Artin, sur Van der Waerden et son grand livre *Moderne Algebra* (voir 5.2), etc., et, plus tard, sur l'organisation des livres de Nicolas Bourbaki. Par contre, ses exposés étaient, dit-on, extrêmement embrouillés. Bien que ses qualités aient été reconnues, la carrière d'Emmy Noether ne correspond pas à son génie à cause de son sexe. Les littéraires de l'Université de Göttingen lui refusent un poste en 1917. Hilbert prend sa défense, soulignant que l'université n'est pas un établissement de bains, puis lui donne la possibilité de donner des cours en la prenant comme assistante, sans qu'elle puisse bénéficier jamais d'une position stable ni de salaires décents. Nous avons déjà raconté (voir 9.6) comment elle est chassée de Göttingen par les nazis en 1933. À son arrivée aux États-Unis, on ne lui propose qu'un poste dans un collège féminin ! Elle meurt brutalement après une opération apparemment sans danger.



Emmy Noether (1882-1935)

19.14 THÉORÈME DE FERMAT POUR $n = 3$

Histoire du théorème de Fermat. Nous avons déjà étudié l'équation $a^2 + b^2 = c^2$ (exercice 12.6) et déterminé l'ensemble de ses solutions en nombres entiers. Fermat propose en 1636 le problème de partager un cube en somme de deux cubes ; plus tard, il affirme que ce problème est impossible, sans jamais donner sa démonstration. Vers 1640, il énoncera en latin dans la marge de son « Diophante » une généralisation de ce problème connue sous le nom de *grand théorème de Fermat* ou *Fermat last theorem* : pour tout entier $n \geq 3$, il n'existe pas d'entiers non nuls a, b, c tels que $a^n + b^n = c^n$:

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duas ejusdem nominis fas est dividere : cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

En français :

Au contraire, il est impossible de partager soit un cube en deux cubes, soit un bicarré en deux bicarrés, soit en général une puissance quelconque supérieure au carré en deux puissances de même degré ; j'en ai découvert une démonstration véritablement merveilleuse que cette marge est trop étroite pour contenir.

Voilà un texte hypercélèbre. Quelle était cette preuve merveilleuse ? On pense aujourd'hui que Fermat a commis une erreur et on a même imaginé une erreur, géniale tout de même, sur l'anneau $\mathbb{Z}[e^{2i\pi/p}]$ pour p premier. Jamais Fermat ne fit part à quiconque de cette note marginale publiée seulement en 1670 dans l'édition de ses œuvres. Est-il possible qu'il l'ait oubliée ? S'était-il aperçu que ses arguments étaient insuffisants ? Impossible de le dire.

Pour le cas $n = 4$, Fermat a donné une démonstration complète. Pour le cas $n = 3$, Euler a donné une démonstration incomplète dans les années 1760 ; plus tard, Gauss trouve la démonstration que nous allons suivre, mais ne la publie pas.

L'histoire un peu détaillée des efforts des mathématiciens pour trouver une démonstration du théorème de Fermat est impossible dans le cadre de ce livre. Le problème initial, qui semble totalement anecdotique, s'est révélé d'une très grande richesse, et lié aux développements des mathématiques. La théorie des idéaux est née des travaux de Eduard Kummer (1810-1893) autour de 1840 sur le théorème de Fermat. En 1969, Yves Hellegouarch (né en 1936) eut l'idée de relier le théorème à un problème sur les courbes elliptiques. Ce qu'on appelait toujours le théorème de Fermat, bien qu'il n'ait pas alors été démontré, devint, après des travaux de Jean-Pierre Serre (né en 1926) et de Ken Ribet (né en 1948), une conséquence d'une conjecture sur les courbes elliptiques, la conjecture de Shimura-Taniyama-Weil¹ (conjecture de Weil principalement, selon Jean-Pierre Serre). Cette conjecture superbe fut résolue en plusieurs étapes entre 1994 et 1999. La partie démontrée en

1. Goro Shimura (1930-2019), Yukata Taniyama (1927-1958)

1994-1995 par Andrew Wiles (né en 1953) seul, puis avec Richard Taylor (né en 1962), pour reprendre un pan défectueux de sa première démonstration, suffisait pour vaincre l'énoncé de Fermat. C'est une des rares fois où des journaux quotidiens ont parlé de mathématiques. Le livre de Yves Hellegouarch *Invitation aux mathématiques de Fermat-Wiles*, publié chez Masson-Dunod explique en détail les chemins pris par la démonstration à un niveau de second cycle.

Théorème : cas $n = 3$ du théorème de Fermat. *L'équation $x^3 + y^3 = z^3$ n'a pas de solution en nombres entiers autres que les triviales, celles où l'un des trois nombres est nul.*

Démonstration. Comme vous allez le constater, même dans ce cas, la démonstration est longue et délicate. Elle utilise les propriétés du sous-anneau $A = \mathbb{Z}[j]$ de $\mathbb{C}$, où

$$j = e^{2\pi i/3} = \frac{-1 + i\sqrt{3}}{2}.$$

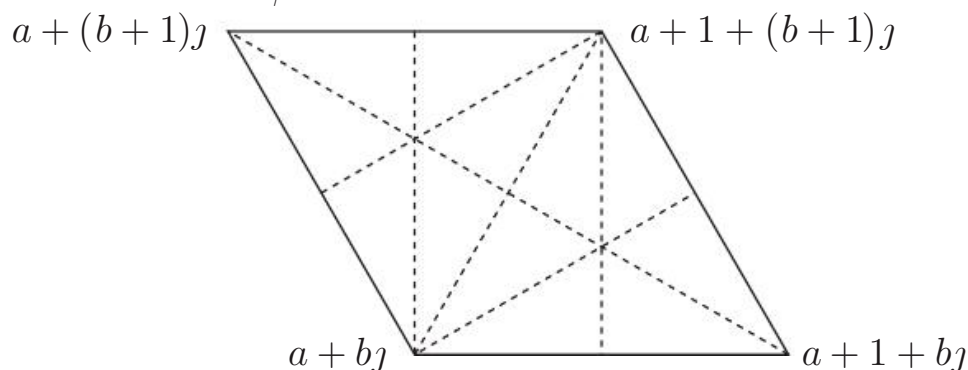
Tout élément de A s'écrit $a + bj$ avec $a, b \in \mathbb{Z}$. On vérifie avec les mêmes méthodes qu'en 19.4 que $A \simeq \frac{\mathbb{Z}[X]}{(X^2 + X + 1)}$.

Le carré du module de $a + bj$ est $N(a + bj) = (a + bj)(a + bj^2) = a^2 - ab + b^2$; la fonction N est appelée norme; elle possède les propriétés données en 19.4 pour la norme sur $\mathbb{Z}[\sqrt{n}]$. Les unités de A sont de norme 1, ce sont $\pm 1, \pm j, \pm j^2$. Une notation u désignera une unité de A .

On ne confondra pas l'anneau A avec l'anneau $\mathbb{Z}[\iota\sqrt{3}]$ (contenu dans A puisque $2j + 1 = \iota\sqrt{3}$ mais non égal à A car ne contenant pas j).

A est un anneau euclidien. Soient $\alpha, \beta \in A$ et supposons $\beta \neq 0$. Le quotient $\frac{\alpha}{\beta}$ est de la forme $x + yj$, avec $x, y \in \mathbb{Q}$. Notons a, b les parties entières de x et y respectivement. Le point $x + yj$ est dans le losange de côté 1 formé par les points $a + s + (b + t)j$ avec $0 \leq s, t \leq 1$.

Dans ce losange formé de deux triangles équilatéraux de côté 1, tout point est à une distance inférieure ou égale à $\sqrt{3}/3$ de l'un des sommets. Il existe donc des entiers a_0 et b_0 tels que $z = \frac{\alpha}{\beta} - (a_0 + b_0j)$ vérifie $N(z) \leq 1/3$. On a



$\alpha = \beta(a_0 + b_0j) + z\beta$, $z\beta \in A$ et $N(z\beta) \leq N(\beta)/3$. Ceci définit une division euclidienne dans A et montre que A est un anneau euclidien pour la jauge définie par la norme.

On en déduit que A est un anneau factoriel.

Propriétés de $\lambda = 1 - j$. L'élément $\lambda = 1 - j$ joue un rôle essentiel dans la démonstration, ce qui est lié à la décomposition $x^3 + y^3 = (x + y)(x + jy)(x + j^2y)$ comme on va le constater. Deux propriétés vont nous être utiles.

- 1) Comme $N(\lambda) = 3$, λ est irréductible donc premier dans A .
- 2) Tout élément de A est égal modulo (λ) à l'un des trois éléments $-1, 0, 1$, donc est de l'une des trois formes $-1 + \lambda\alpha$, $\lambda\alpha$, $1 + \lambda\alpha$ avec $\alpha \in A$.

En effet, si $a + bj \in A$, on a $a + bj = a + b - b\lambda$, ce qui montre que $a + bj$ est égal modulo λ à un entier ; comme $3 = \lambda\bar{\lambda}$, tout entier est bien égal modulo λ à $0, 1, -1$.

On peut aussi raisonner en utilisant les isomorphismes :

$$\frac{A}{(\lambda)} \simeq \frac{\mathbb{Z}[X]}{(X^2 + X + 1, 1 - X)} \simeq \frac{\mathbb{Z}[X]/(1 - X)}{(X^2 + X + 1)} \simeq \mathbb{Z}/3\mathbb{Z}.$$

Calcul de cubes modulo λ^4 dans A . Si $x = 1 \pmod{\lambda}$, il existe $\alpha \in A$ tel que $x = 1 + \lambda\alpha$. Nous allons montrer que $x^3 = 1 \pmod{\lambda^4}$.

D'abord, $x^3 = 1 + 3\alpha\lambda + 3\alpha^2\lambda^2 + \alpha^3\lambda^3$.

D'une part, $3\alpha^2\lambda^2 = 0 \pmod{\lambda^4}$ car $3 = -j^2\lambda^2$.

D'autre part, $3\alpha\lambda + \alpha^3\lambda^3 = -j^2\lambda^2\alpha\lambda + \alpha^3\lambda^3 = \alpha\lambda^3(\alpha^2 - j^2)$. On a $\alpha - j = \alpha - 1 \pmod{\lambda}$, $\alpha + j = \alpha + 1 \pmod{\lambda}$, donc $\alpha(\alpha^2 - j^2) = \alpha(\alpha + 1)(\alpha - 1) = 0 \pmod{\lambda}$ et on conclut.

De même, si $x = -1 \pmod{\lambda}$, alors $x^3 = -1 \pmod{\lambda^4}$.

Impossibilité de $x^3 + y^3 = uz^3$ dans A . On va montrer que l'équation

$$x^3 + y^3 = uz^3,$$

où u est une unité de $A^\times$, n'a pas de solution dans A telle que x, y et z ne soient pas nuls. Cette impossibilité entraîne celle qu'on veut démontrer.

On suppose qu'il existe une solution dans A de cette équation avec x, y, z non nuls ; on peut supposer que x, y, z sont premiers entre eux deux à deux.

L'idée de la démonstration est d'utiliser la relation $z^3 = (x + y)(x + jy)(x + j^2y)$ et d'en déduire que des facteurs du membre de droite sont des cubes. Pour cela, on va utiliser la factorialité de A et on va trouver des termes premiers entre eux deux à deux dans le membre de droite.

- 1) Si λ divise x ou y dans A , alors $u = \pm 1$.

Supposons que λ divise x ; on peut écrire $x = \alpha \lambda$ avec $\alpha \in A$; on a $x^3 = \alpha^3 \lambda^3 = 0 \pmod{\lambda^3}$. Comme λ ne divise ni y ni z , y et z sont égaux à $\pm 1 \pmod{\lambda}$, donc y^3, z^3 sont égaux à $\pm 1 \pmod{\lambda^3}$, car λ^2 divise 3. On en déduit $u = \pm 1 \pmod{\lambda^3}$. Comme le module de λ^3 est $3\sqrt{3}$, les seules possibilités sont $u = \pm 1$.

2) λ divise x, y ou z dans A .

Supposons le contraire. Alors $x, y, z = \pm 1 \pmod{\lambda}$ et $x^3, y^3, z^3 = \pm 1 \pmod{\lambda^4}$. Si $x^3 = 1 \pmod{\lambda^4}$, alors $y^3 = 1 \pmod{\lambda^4}$ ($y^3 = -1 \pmod{\lambda^4}$ donne z multiple de λ). De même, si $x^3 = -1 \pmod{\lambda^4}$, alors $y^3 = -1 \pmod{\lambda^4}$. Donc $u = \pm 2 \pmod{\lambda^4}$, ce qui est impossible puisque le module de λ^4 est 9 et que les unités de A sont de module 1.

3) On peut se ramener à λ divise z dans A .

Si λ divise x (ou y), on a vu que $u = \pm 1$; on peut, en changeant éventuellement leurs signes, échanger les rôles de x (ou y) et z dans la relation.

On suppose désormais que λ divise z .

4) λ^2 divise z .

Si λ divise z , λ ne divise ni x ni y , donc $x, y = \pm 1 \pmod{\lambda}$ et $x^3 + y^3 = 0 \pmod{\lambda}$. Les seules possibilités sont $x = -y = \pm 1 \pmod{\lambda}$. On a alors $x^3 + y^3 = 0 \pmod{\lambda^4}$, λ^4 divise z^3 , donc λ^2 divise z .

5) On peut se ramener à λ^2 divise $x + y$ dans A .

Le 4) montre que λ^6 divise $x^3 + y^3 = (x + y)(x + jy)(x + j^2y)$. L'un des trois facteurs au moins est divisible par λ^2 . On peut échanger les rôles des facteurs en changeant y en jy ou j^2y . On peut donc supposer que λ^2 divise $x + y$ dans A .

6) $v_\lambda(x + y) = 3v_\lambda(z) - 2$, $v_\lambda(x + jy) = 1$, $v_\lambda(x + j^2y) = 1$.

Comme $j = 1 \pmod{\lambda}$, on a $x + jy = x + j^2y = x + y \pmod{\lambda}$, ce qui montre que les trois valuations sont supérieures ou égales à 1. Si λ^2 divise $x + jy$, $x + jy = x + y \pmod{\lambda^2}$, soit $(1 - j)y = 0 \pmod{\lambda^2}$, d'où $y = 0 \pmod{\lambda}$, ce qui est impossible. De même, λ^2 divise $x + j^2y$ est impossible. D'où les valuations annoncées.

7) Obtention d'un nouveau triplet et contradiction.

On vérifie alors que $(x + jy)/\lambda$, $(x + j^2y)/\lambda$, $(x + y)/\lambda^{3v_\lambda(z)-2}$, sont premiers entre eux deux à deux, car si un élément irréductible divise deux de ces termes, il divise x et y . La factorialité de A permet de déduire de la relation $z^3 = x^3 + y^3 = (x + y)(x + jy)(x + j^2y)$ qu'on peut écrire dans A :

$$x + jy = u_1 \lambda x_1^3$$

$$x + j^2y = u_2 \lambda y_1^3$$

$$x + y = u_3 \lambda^{3v_\lambda(z)-2} s^3$$

En prenant la somme de la première égalité, de la seconde multipliée par j , de la troisième multipliée par j^2 , on obtient :

$$u_1 \lambda x_1^3 + u_2 j \lambda y_1^3 + u_3 j^2 \lambda^{3v_\lambda(z)-2} s^3 = 0.$$

En posant $z_1 = \lambda^{v_\lambda(z)-1}s$ et en divisant par λ et u_1 , on obtient :

$$x_1^3 + u_2' y_1^3 = u_3' z_1^3.$$

Modulo λ^3 , cette relation donne $\pm 1 \pm u_2' = 0$, soit $u_2' = \pm 1$ et, en changeant éventuellement y_1 en $-y_1$, on se ramène à la relation :

$$x_1^3 + y_1^3 = u' z_1^3.$$

On a obtenu une nouvelle relation avec un z dont la valuation en λ a diminué de 1. La méthode de descente infinie de Fermat montre que ce n'est pas possible indéfiniment. $\square$

19.15 CORPS DES FRACTIONS D'UN ANNEAU INTÈGRE

Voici maintenant la généralisation de la construction du corps des rationnels $\mathbb{Q}$ à partir de l'anneau $\mathbb{Z}$. Si vous ne connaissez pas cette construction, lisez ce qui suit en pensant que l'anneau intègre A est $\mathbb{Z}$ et que le corps $\text{Cf}(A)$ qu'on construit est $\mathbb{Q}$: la démonstration dans le cas général est exactement copiée de la démonstration dans le cas particulier (ou inversement). L'idée de cette construction est que tout rationnel se présente comme un quotient de deux entiers (le second non nul) et qu'il faut introduire une relation d'équivalence pour identifier les couples d'entiers proportionnels, car ils définissent la même fraction.

Soit A un anneau intègre. On définit une relation d'équivalence sur $A \times A^*$ en posant $(r, s) \mathcal{R} (r', s')$ si $rs' = r's$ (vérification facile).

On note r/s ou $\frac{r}{s}$ la classe d'équivalence de (r, s) et on l'appelle une fraction.

Étant donné un représentant (r, s) d'une fraction, on appelle r le numérateur, s le dénominateur du représentant. On peut simplifier par un facteur commun numérateur et dénominateur sans changer la fraction car $ar/as = r/s$ pour tout $a \in A^*$ puisque $(ar, as) \mathcal{R} (r, s)$; ce qui permet d'obtenir un représentant d'une fraction où les numérateur et dénominateur sont premiers entre eux ; on dit que la fraction est mise sous forme irréductible.

Nous allons munir l'ensemble quotient de $A \times A^*$ par $\mathcal{R}$ d'une structure de corps, le corps des fractions de A , que nous noterons $\text{Cf}(A)$, possédant une propriété universelle naturelle.

Les lois de $\text{Cf}(A)$ sont définies par :

$$\begin{aligned} \frac{r}{s} + \frac{r'}{s'} &= \frac{rs' + r's}{ss'} ; \\ \frac{r}{s} \times \frac{r'}{s'} &= \frac{rr'}{ss'}. \end{aligned}$$

On vérifie facilement que ces définitions ne dépendent pas du choix des représentants et on retrouve toutes les propriétés des fractions de notre enfance : tout se passe comme avec les fractions ordinaires d'entiers. Par exemple, pour vérifier que l'addition ne dépend pas des représentants choisis pour les fractions, on prend

$(r, s)\mathcal{R}(r_1, s_1)$ et $(r', s')\mathcal{R}(r'_1, s'_1)$ et on vérifie que $\frac{rs' + r's}{ss'} = \frac{r_1s'_1 + r'_1s_1}{s_1s'_1}$ ce qui résulte de $(rs' + r's)s_1s'_1 = (r_1s'_1 + r'_1s_1)ss'$.

L'élément neutre pour l'addition est $0/1$, noté 0 , l'élément neutre pour la multiplication est $1/1$, noté 1 . L'inverse de r/s pour r, s non nuls est s/r .

L'application $i : A \rightarrow \text{Cf}(A)$ définie par $i(a) = a/1$ est un homomorphisme injectif tel que $r/s = i(r)i(s)^{-1}$.

Proposition : Propriété universelle du corps des fractions. Soit A un anneau intègre et $i : A \rightarrow \text{Cf}(A)$ l'homomorphisme défini ci-dessus.

Pour tout corps K et tout homomorphisme $f : A \rightarrow K$ injectif, il existe un unique homomorphisme d'anneaux $f' : \text{Cf}(A) \rightarrow K$ tel que le diagramme

$$\begin{array}{ccc} A & \xrightarrow{f} & K \\ \downarrow i & \nearrow f' & \\ \text{Cf}(A) & & \end{array}$$

commute.

Démonstration. Soit $\alpha \in \text{Cf}(A)$ une fraction représentée par le couple (r, s) d'éléments de $A \times A^*$. Comme $s \neq 0, f(s) \neq 0$. On pose $f'(\alpha) = f(r)f(s)^{-1}$. Il est facile de vérifier que $f'(\alpha)$ ne dépend pas du représentant de α choisi et que f' est un homomorphisme d'anneaux vérifiant les conditions demandées. $\square$

Dans le cas où A est un corps, on peut prendre $\text{Cf}(A) = A$ et $i = \text{id}_A$. Dans le cas de l'anneau $\mathbb{Z}$, la construction précédente donne le corps $\mathbb{Q}$ et l'inclusion $i : \mathbb{Z} \rightarrow \mathbb{Q}$; la propriété universelle précédente signifie que tout homomorphisme d'anneaux $\mathbb{Z} \rightarrow K$, où K est un corps de caractéristique 0, se prolonge à $\mathbb{Q}$. Par conséquent, un corps de caractéristique 0 est une extension de $\mathbb{Q}$ (ou plutôt admet un sous-corps isomorphe à $\mathbb{Q}$).

Par contre, la factorisation précédente ne marche pas si l'homomorphisme $A \rightarrow K$ n'est pas injectif. Par exemple, la projection canonique $\mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z}$ ne s'étend pas en un homomorphisme $\mathbb{Q} \rightarrow \mathbb{Z}/2\mathbb{Z}$.

Nous venons de construire un corps où tous les éléments non nuls de l'anneau A sont inversibles. Dans un anneau quelconque, intègre ou non, il est possible de faire une construction analogue en rendant inversibles tous les éléments d'une partie S

stable par multiplication de A (avec $1 \in S$) ; l'anneau obtenu s'appelle anneau des fractions de A par la partie (dite *multiplicative*) choisie. Nous ne pouvons décrire ici cette construction dont les usages sont nombreux.

Soient A et B des anneaux intègres et $f: A \rightarrow B$ un homomorphisme injectif. La propriété universelle permet de construire à partir du composé $j \circ f: A \rightarrow B \rightarrow \text{Cf}(B)$, où j est l'inclusion analogue à i , un homomorphisme $g: \text{Cf}(A) \rightarrow \text{Cf}(B)$ prolongeant

$$f \text{ aux fractions : } g\left(\frac{r}{s}\right) = \frac{f(r)}{f(s)}.$$

Corps des fractions d'un anneau factoriel. Soit A un anneau factoriel. On peut étendre la définition de divisibilité à $\text{Cf}(A)$: si $x, y \in \text{Cf}(A)$, on dira que x divise y si

$\frac{y}{x} \in A$. Si on se donne un système $(p_i)_{i \in I}$ de représentants d'irréductibles de A , on

peut écrire de manière unique tout $x \in \text{Cf}(A)$ sous la forme $x = \prod_{i \in I} p_i^{k_i}$ avec les

$k_i \in \mathbb{Z}$ presque tous nuls et les propriétés de divisibilité, de calcul des pgcd et ppcm énoncées en 19.13 s'étendent aux éléments de $\text{Cf}(A)$. Par exemple, dans

$\mathbb{Q} = \text{Cf}(\mathbb{Z})$, on dira que $3/22$ divise $6/11$ puisque le quotient est 4, on écrira

$$\text{pgcd}\left(\frac{3}{5}, \frac{7}{20}\right) = \pm \frac{1}{20}, \text{ppcm}\left(\frac{3}{5}, \frac{7}{20}\right) = \pm \frac{21}{5}.$$

Corps des fractions rationnelles à une indéterminée. Soit K un corps. Le corps des fractions de l'anneau intègre $K[X]$ est appelé corps des fractions rationnelles à coefficients dans K et noté $K(X)$. Un élément de $K(X)$ est appelé fraction rationnelle ;

il s'écrit $\frac{P}{Q}$ où P et Q sont des polynômes de $K[X]$ avec Q non nul. Cela fait bien

longtemps que vous calculez avec des fractions rationnelles et cette construction enfin rigoureuse ne doit rien changer dans vos habitudes de calcul.

Soit $a \in K$. Si une fraction de $K(X)$ peut s'écrire sous la forme P/Q avec $Q(a) \neq 0$,

on dit que sa valeur en a est $P(a)/Q(a)$. Par exemple, comme $\frac{X^3-1}{X^2-1} = \frac{X^2+X+1}{X+1}$,

cette fraction a pour valeur $3/2$ en 1 . La situation formelle est donc différente de la

situation en analyse où la fonction $x \mapsto \frac{x^3-1}{x^2-1}$ n'est pas définie en 1 .

Si $A = \mathbb{Z}$ et $K = \mathbb{Q}$, toute fraction $\frac{P}{Q}$ de polynômes à coefficients rationnels est égale à une fraction formée de deux polynômes à coefficients entiers : il suffit de multiplier P et Q par le produit des dénominateurs de leurs coefficients.

Plus généralement, si A est un anneau intègre de corps des fractions K , $K(X)$ est isomorphe au corps des fractions de $A[X]$.

Le corps des fractions rationnelles à plusieurs indéterminées sera construit en 20.8.

Sous-corps engendré par un élément. Soient K un corps, L une extension de K , $x \in L$. L'extension de K engendrée par x est l'ensemble des $P(x)/Q(x)$ où P/Q est une fraction rationnelle à coefficients dans K dont le dénominateur ne s'annule pas en x . On la note $K(x)$ avec des parenthèses.

Parfois la description de $K(x)$ est simplifiable. C'est le cas des extensions comme $\mathbb{Q}(\sqrt{2})$: on a vu que $\mathbb{Q}[\sqrt{2}]$ est un corps ; par conséquent, $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}]$. Plus généralement, si n est un entier sans facteurs carrés, $\mathbb{Q}(\sqrt{n}) = \mathbb{Q}[\sqrt{n}]$ est le corps des fractions de $\mathbb{Z}[\sqrt{n}]$.

De $\mathbb{N}$ à $\mathbb{C}$. Ce chapitre permet de contempler les étapes de la construction des différents corps dans lesquels on effectue des calculs depuis le secondaire.

On part de l'ensemble des entiers $\mathbb{N}$, défini par des axiomes dans le cadre de la théorie des ensembles.

On construit $\mathbb{Z}$ en symétrisant l'addition de $\mathbb{N}$; pour cela, on considère le quotient $\mathbb{N} \times \mathbb{N}$ par la relation d'équivalence $(m, n) \mathcal{R} (m', n')$ si et seulement si $m + n' = m' + n$; la classe de (m, n) correspond à l'entier relatif $m - n$; les nombres positifs sont les classes des couples $(m, 0)$, les nombres négatifs sont les classes des couples $(0, n)$. Les vérifications sont fastidieuses, mais, quand on en a compris le principe, tout devient simple.

On construit ensuite le corps des fractions de $\mathbb{Z}$ comme ci-dessus : c'est le corps des nombres rationnels $\mathbb{Q}$.

On construit ensuite $\mathbb{R}$ en complétant $\mathbb{Q}$ par les limites des suites de Cauchy de $\mathbb{Q}$, puis $\mathbb{C}$ en quotientant $\mathbb{R}$ pour avoir un nombre de carré -1 comme c'est indiqué en 19.4.

On a aussi les moyens d'intercaler tout plein de corps entre $\mathbb{Q}$ et $\mathbb{C}$, certains contenus dans $\mathbb{R}$, d'autres non : ce sont, par exemple, les extensions de $\mathbb{Q}$ (comme $\mathbb{Q}[\sqrt{2}]$, $\mathbb{Q}[\iota]$, etc.), engendrées par des nombres algébriques sur $\mathbb{Q}$, c'est-à-dire des nombres racines de polynômes à coefficients dans $\mathbb{Q}$.

Des nombres nouveaux : les périodes. On pourrait s'arrêter là et contempler avec satisfaction le travail accompli. Cependant, en 2001, Maxim Kontsevitch, médaille Fields en 1996 et professeur à l'Institut des Hautes Etudes Scientifiques de Bures-sur-Yvette, et Don Zagier (né en 1951, professeur au Collège de France de 2001 à 2014), ont distingué une nouvelle catégorie particulièrement intéressante de nombres appelés *périodes*. Une période est un nombre complexe dont les parties réelles et imaginaires sont les valeurs d'intégrales convergentes de fonctions rationnelles à coefficients rationnels sur des domaines délimités par des inégalités polynomiales à coefficients rationnels. Il est clair que π , $\ln(k)$ pour k entier et bien d'autres nombres transcendants sont des périodes ; Maxim Kontsevitch et Don Zagier ont formulé de nombreuses conjectures à résoudre dans ce nouveau domaine, comme de déterminer si la base des logarithmes népériens e ou le nombre $1/\pi$ sont ou non des périodes. Joseph Ayoub a généralisé le problème et montré, en 2015, en s'appuyant sur les travaux de Grothendieck, que les relations entre périodes se déduisent du calcul intégral tel que l'a fondé Leibniz, sans introduction de relations nouvelles.

Exercices

19.1 Exercices généraux sur les anneaux

1) Soient E un ensemble et $(A, +, \cdot)$ un anneau. Montrer que l'ensemble $\mathcal{F}$ des applications de E dans A peut-être muni d'une structure d'anneau en posant, pour $f, f' \in A, f + f' : x \mapsto f(x) + f'(x)$ et $f \cdot f' : x \mapsto f(x) \cdot f'(x)$.

2) Montrer que l'injection $i : \mathbb{Z} \rightarrow \mathbb{Q}$ est un épimorphisme d'anneaux, c'est-à-dire que, si $u, v : \mathbb{Q} \rightarrow A$ sont des homomorphismes d'anneaux tels que $u \circ i = v \circ i$ alors $u = v$.

3) Montrer qu'un anneau intègre fini est un corps.

4) a) Quelles sont les caractéristiques des anneaux

$$\mathbb{Z}/11\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, \mathbb{Z}/33\mathbb{Z} \times \mathbb{Z}/55\mathbb{Z}, \mathbb{Z}/52\mathbb{Z} \times \mathbb{Z}/55\mathbb{Z} ?$$

b) Quelle est la caractéristique du produit d'anneaux $A = \prod_{n \geq 1} \mathbb{Z}/n\mathbb{Z}$?

19.2 Exercices généraux sur les idéaux

1) a) Soit K un corps. Déterminer tous les idéaux de $K \times K$.

b) Déterminer tous les idéaux de $\mathbb{Z} \times \mathbb{Z}$.

c) Décrire l'ensemble ordonné par inclusion des idéaux de $\mathbb{Z}/360\mathbb{Z}$.

2) Union d'idéaux

a) Montrer que, dans un anneau, un idéal contenu dans l'union de deux idéaux est contenu dans l'un d'entre eux.

b) A-t-on le même résultat pour trois idéaux ?

3) Opérations sur les idéaux

a) Montrer que la multiplication des idéaux est associative, commutative, possède un élément neutre.

b) Distributivité : si I_1, I_2, I_3 sont des idéaux d'un anneau A , montrer que $I_1(I_2 + I_3) = I_1I_2 + I_1I_3$.

c) Loi modulaire : si I_1, I_2, I_3 sont des idéaux d'un anneau A , montrer que, si $I_2 \subset I_1$, on a :

$$I_1 \cap (I_2 + I_3) = (I_1 \cap I_2) + (I_1 \cap I_3) = I_2 + (I_1 \cap I_3).$$

4) Idéaux maximaux d'un espace de fonctions

Soit E un espace métrique compact et A l'anneau des fonctions continues de E dans $\mathbb{R}$. Pour tout $x \in E$, on note I_x l'ensemble des fonctions de A nulles en x .

- a) Montrer que I_x est un idéal maximal de A .
- b) Soit M un idéal maximal de A . Montrer qu'il existe x dans E tel que $M = I_x$.

5) Anneau des suites de nombres réels

Soit A l'anneau des suites de nombres réels.

- a) Montrer que, pour tout entier n , les suites de A dont le terme de rang n est nul forment un idéal maximal M_n de A .
- b) Donner un exemple d'idéal de A contenu dans l'union des M_n mais dans aucun d'entre eux.

19.3 Exercices sur les quotients

- 1) Soit K un corps. À quel anneau « plus simple » est isomorphe l'anneau $K[X]/(aX + b)$, avec $a, b \in K$, $a \neq 0$?
- 2) Soit K un corps. À quel anneau « plus simple » est isomorphe l'anneau $K[X]/(X^2 - 1)$?
- 3) À quel anneau « plus simple » est isomorphe l'anneau $\mathbb{C}[X]/(X^2 + 1)$?
- 4) Soit K un corps. Montrer que les anneaux $K[X]/(X^2 + 1)$ et $K[X]/(X^2 + 2X + 2)$ sont isomorphes.
- 5) Soit K un corps. Les anneaux $K[X]/(X^2)$ et $K[X]/(X^2 + 1)$ sont-ils isomorphes ?
- 6) Même question pour les anneaux $\mathbb{Q}[X]/(X^2 - 3)$ et $\mathbb{Q}[X]/(X^2 - 5)$.
- 7) À quels anneaux « plus simples » sont isomorphes les anneaux $\mathbb{Z}[t]/(2 + t)$, $\mathbb{Z}[t]/(7 + t)$?
- 8) a) Déterminer les anneaux isomorphes de la liste suivante : $\mathbb{Z}[X]/(3X + 1)$, $\mathbb{Z}[X]/(3X - 1)$, $\mathbb{Z}[X]/(3X)$.
- b) Déterminer les couples d'anneaux isomorphes de la famille $A_k = \mathbb{Z}[X]/(3X - k)$ avec $k \in \mathbb{Z}$.
- c) Déterminer un anneau « plus simple » isomorphe à l'anneau $A = \mathbb{Z}[X]/(4X - 2)$.

9) Lemme du losange

Soient A un anneau, I et J deux idéaux de A . On note $i : I \rightarrow A$ l'injection, $\pi : A \rightarrow A/J$ la projection canonique.

- a) Déterminer l'image et le noyau de $\pi \circ i$.
- b) Montrer que les anneaux (non unitaires) $I/(I \cap J)$ et $(I + J)/J$ sont isomorphes, puis que les anneaux $J/(I \cap J)$ et $(I + J)/I$ sont isomorphes, résultat analogue à celui de l'exercice 18.6.4) pour les groupes.

19.4 Idéaux étrangers

- 1) Montrer que $4X$ et $2X - 1$ sont étrangers dans $\mathbb{Z}[X]$.
- 2) Soient A un anneau et $a, b \in A$. À quelle condition nécessaire et suffisante $X - a$ et $X - b$ sont-ils étrangers dans $A[X]$?
- 3) Les polynômes $X^2 + 1$ et $(X + 1)^2$ sont-ils étrangers dans $\mathbb{Z}[X]$, dans $\mathbb{Q}[X]$?
- 4) Montrer que, si I et J sont des idéaux étrangers d'un anneau A , alors I^r et J^s sont, pour $r, s \geq 1$ des idéaux étrangers de A .

19.5 Théorème chinois

- 1) Résoudre dans $\mathbb{Z}$ le système

$$\begin{cases} x = 1 \pmod{3} \\ x = 5 \pmod{7} \\ x = 9 \pmod{13} \end{cases}$$

- 2) Construire l'inverse de l'isomorphisme $\varphi : \mathbb{Z}/273\mathbb{Z} \rightarrow \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/13\mathbb{Z}$ donné par le théorème chinois, puis retrouver la solution de la première question.
- 3) Trouver l'ensemble des $P \in (\mathbb{Z}/17\mathbb{Z})[X]$ tel que $P(0) = 3$, $P(1) = 11$ et $P(4) = 2$.

19.6 Inversibles

- 1) Soient K un corps et $n \geq 2$ un entier. On note x la classe de X dans $A = K[X]/(X^n)$. Montrer que $1 - x$ est inversible dans A et donner son inverse.
- 2) Unités de $A = \mathbb{Z}[\sqrt{2}]$
 - a) Caractériser les unités de A par leur norme.
 - b) On pose $w = 1 + \sqrt{2}$. Vérifier que les W^n avec $n \in \mathbb{Z}$ sont des unités de A .
 - c) Soit $u = a + b\sqrt{2}$ une unité de A . Déterminer les signes de a et b pour la plus grande (en considérant A comme inclus dans $\mathbb{R}$) des quatre unités $\pm u$ et $\pm u^{-1}$.
 - d) Montrer que si u est une unité de A , on ne peut avoir $1 < u < W$. En déduire l'ensemble des unités de A .
 - e) À quel groupe est isomorphe $U(A)$?

Commentaire. Il existe un théorème général décrivant le groupe des unités des corps de nombres. Il date de 1837 et est dû à Gustav Lejeune Dirichlet dont la famille était originaire de Belgique ; son nom est formé à partir de Le Jeune de Richellette, près de Liège. Dirichlet a occupé la chaire de Gauss à Göttingen, de 1855 à sa mort, en 1859.

19.7 Idéaux premiers et maximaux

1) Soit A un anneau. Montrer les propriétés énoncées en 19.11. Par exemple, montrer que (0) est un idéal premier si et seulement si A est intègre, que (0) est un idéal maximal si et seulement si A est un corps, etc.

2) a) Soit A un anneau. Montrer que, si un idéal premier P de A contient un produit fini d'idéaux de A , il contient l'un d'entre eux.

b) Montrer que, si M est un idéal maximal de A , c'est le seul idéal maximal de A contenant M^n .

3) Lemme d'évitement des idéaux premiers

Soient A un anneau, X une partie non vide de A stable par addition et multiplication, $(P_i)_{1 \leq i \leq n}$ une famille finie non vide d'idéaux de A tous premiers sauf deux au plus. On suppose que X est contenue dans $\cup_{1 \leq i \leq n} P_i$. Montrer qu'il existe i tel que $X \subset P_i$ (raisonner par récurrence et, pour $n \geq 3$, considérer séparément le cas où il existe un j tel que $X \cap P_j \subset \cup_{1 \leq i \leq n, i \neq j} (X \cap P_i)$ et le cas où il n'existe pas de tel j).

19.8 Anneaux euclidiens

1) Soit A un anneau. On suppose que $A[X]$ est euclidien pour la jauge euclidienne définie par le degré. Montrer que A est un corps.

2) Montrer que les anneaux $\mathbb{Z}[\sqrt{n}]$ sont euclidiens pour $n = -2, 2, 3$ (prendre comme jauge euclidienne la valeur absolue de la norme).

3) L'anneau $A = \mathbb{Z}[\sqrt{3}]$ est-il factoriel ? est-il euclidien ?

19.9 Exemple d'anneau principal non euclidien

Posons $\rho = \frac{1+\sqrt{-19}}{2}$ et soit $A = \mathbb{Z}[\rho]$ le sous-anneau de $\mathbb{C}$ engendré par ρ . On utilise la norme N définie par le carré du module.

1) Montrer que A est isomorphe à un quotient de $\mathbb{Z}[X]$ et que les éléments de A peuvent s'écrire sous la forme $a + b\rho$, avec $a, b \in \mathbb{Z}$.

2) Montrer que $N(a + b\rho)$ est un entier $\geq 4b^2$. Quels sont les éléments inversibles de A ?

3) Montrer que, si un anneau (B, j) est euclidien, il existe un élément non inversible $t \in B$ tel que la restriction de la projection canonique $\pi : B \rightarrow B/(t)$ à $B^\times \cup \{0\}$ soit surjective (choisir, si B n'est pas un corps, un élément non inversible t tel que $j(t)$ soit minimal).

4) En déduire que A n'est pas euclidien.

5) Montrer que l'idéal (2) est maximal dans A (montrer que l'anneau quotient $A/(2)$ est un corps).

6) Montrer que A admet une pseudodivision au sens suivant : si $\alpha, \beta \in A$ sont non nuls, il existe $q, r \in A$ tels que $\alpha = q\beta + r$ ou $2\alpha = q\beta + r$ avec $N(r) < N(\beta)$ dans les deux cas (écrire α/β sous la forme $x + y\rho$ avec $x, y \in \mathbb{Q}$ et traiter d'abord le cas où y est à moins de $1/3$ de l'entier le plus proche).

7) Montrer que A est un anneau principal.

19.10 Anneaux principaux

1) Montrer qu'un produit d'anneaux principaux est principal.

2) Montrer que $(2, X)$ n'est pas un idéal principal de $\mathbb{Z}[X]$, que (X, Y) n'est pas un idéal principal de $K[X, Y]$ quand K est un corps.

19.11 ppcm et pgcd

Soient A un anneau intègre, a et b deux éléments de A .

1) On suppose que m est un ppcm de a et b . Montrer que $cm = \text{ppcm}(ac, bc)$ pour tout $c \in A$.

2) On a vu que la définition de pgcd impose que a et b soient des multiples de d , autrement dit que $a, b \in (d)$, mais ne demande pas que l'idéal $(a) + (b)$ engendré par a et b soit principal. Montrer que, si c'est le cas, et si $(c) = (a) + (b)$, alors c est un pgcd de a et b .

19.12 Anneaux factoriels

1) Montrer que $(\mathbb{Z}/6\mathbb{Z})[X]$ n'est pas factoriel.

2) Montrer que $\mathbb{Z}[\sqrt{-3}]$ n'est pas factoriel.

3) Donner un idéal non principal de $A = \mathbb{Z}[\sqrt{-3}]$.

4) Montrer que le sous-anneau A de $\mathbb{R}[X]$ engendré par X^2 et X^3 n'est pas factoriel.

5) Vérifier l'égalité $17 = (5 - 2\sqrt{2})(5 + 2\sqrt{2}) = (7 + 4\sqrt{2})(7 - 4\sqrt{2})$. Peut-on en déduire que $A = \mathbb{Z}[\sqrt{2}]$ n'est pas factoriel ?

6) Soit A un anneau principal. Peut-on supposer que A ne possède pas d'éléments irréductibles ?

19.13 Corps des fractions

Déterminer les corps des fractions de $\mathbb{Z}[t]$, de $\mathbb{Z}[\sqrt{n}]$ pour n sans facteur carré, de $\mathbb{Z}[X]$, de $\mathbb{Q}[X]$.

19.14 Applications du lemme de Zorn

- 1) Montrer qu'étant donnés deux ensembles, il existe toujours une injection de l'un dans l'autre.
- 2) Montrer que, dans un espace vectoriel, tout sous-espace admet un supplémentaire.
- 3) Soient E un espace vectoriel, A un système générateur de E , $C \subset A$ un système libre. Montrer qu'il existe une base B de E telle que $C \subset B \subset A$.
- 4) Montrer qu'il existe une application $f: \mathbb{R} \rightarrow \mathbb{R}$ qui est un endomorphisme du groupe additif $(\mathbb{R}, +)$ mais n'est pas $\mathbb{R}$ -linéaire.
- 5) Montrer que, dans un anneau, tout idéal premier contient un idéal premier minimal.

Solutions

19.1 1) a) Vérifications sans problèmes.

2) Comme $u(1) = v(1)$, on a $u(n) = v(n)$ pour tout entier $n \in \mathbb{Z}$ (récurrence évidente). On a alors, pour tout entier $n \in \mathbb{Z}$ non nul, $u(1/n) = u(1/n)v(n)v(1/n) = u(1/n)u(n)v(1/n) = v(1/n)$. Enfin, $u(p/q) = u(p)u(1/q) = v(p/q)$.

On notera qu'une application ensembliste $f: E \rightarrow F$ possède la propriété d'être un épimorphisme (c'est-à-dire que, pour tout $u, v: F \rightarrow G$ tels que $u \circ f = v \circ f$ alors $u = v$) si et seulement si elle est surjective.

3) Soient A un anneau intègre fini et $a \in A$ non nul. L'application de A dans A définie par $x \mapsto ax$ est injective car A est intègre. Comme A est fini, elle est surjective, donc il existe $x \in A$ tel que $ax = 1$. Par conséquent, a est inversible et la conclusion.

4) a) Les caractéristiques sont des ppcm ; on trouve respectivement : 33, 165 et 2860.

b) Pour tout entier n non nul, n fois l'unité de $\mathbb{Z}/2n\mathbb{Z}$ n'est pas nul, donc n fois l'unité de A a une composante non nulle. La caractéristique de A est 0.

19.2 1) a) Les idéaux de $K \times K$ sont $K \times K$, $K \times \{0\}$, $\{0\} \times K$, $\{0\} \times \{0\}$.

b) Les idéaux de $\mathbb{Z} \times \mathbb{Z}$ sont les idéaux de la forme $m\mathbb{Z} \times n\mathbb{Z}$ avec $m, n \in \mathbb{N}$.

c) La projection canonique $\pi: \mathbb{Z} \rightarrow \mathbb{Z}/360\mathbb{Z}$ établit une bijection entre les idéaux de $\mathbb{Z}/360\mathbb{Z}$ et les idéaux de $\mathbb{Z}$ contenant $360\mathbb{Z}$, c'est-à-dire les idéaux $d\mathbb{Z}$ avec d divisant 360. L'ensemble des idéaux de $\mathbb{Z}/360\mathbb{Z}$ est donc en bijection avec l'ensemble des diviseurs de 360, mais il est muni de l'ordre opposé. On peut représenter cet ensemble ordonné en choisissant des directions différentes pour indiquer les divisibilités par 2, 3, 5.

2) a) Si $I \subset I_1 \cup I_2$ sans que $I \subset I_1$ ni $I \subset I_2$, il existe $x_1 \in I \setminus I_1$, $x_2 \in I \setminus I_2$; posons $x = x_1 + x_2$. On a $x \in I$, mais $x \in I_1$, $x \in I_2$ sont impossibles. Par conséquent, I est contenu dans l'un des deux idéaux I_1 ou I_2 .

b) Voici un exemple dans $A = (\mathbb{Z}/2\mathbb{Z})[X, Y]$. Notons I_1, I_2, I_3 les idéaux engendrés par les polynômes de degré 2 de A et $X, Y, X + Y$ respectivement. Considérons l'idéal $M = (X, Y)$ engendré par X et Y . Les polynômes du premier degré de M sont X, Y et $X + Y$; donc M est contenu dans $I_1 \cup I_2 \cup I_3$ sans être contenu dans l'un des trois idéaux. Comme M^2 est l'idéal des polynômes sans termes de degré 0 ou 1, on notera qu'on peut écrire $I_1 = (X, M^2)$, $I_2 = (Y, M^2)$, $I_3 = (X + Y, M^2)$.

3) a) L'élément neutre est A .

b) Les inclusions $I_1 I_2 \subset I_1(I_2 + I_3)$, $I_1 I_3 \subset I_1(I_2 + I_3)$ entraînent $I_1 I_2 + I_1 I_3 \subset I_1(I_2 + I_3)$. D'autre part, $I_1(I_2 + I_3)$ est engendré par les produits de la forme $x(y + z)$ avec $x \in I_1$, $y \in I_2$, $z \in I_3$; comme $xy \in I_1 I_2$, $xz \in I_1 I_3$, on a $x(y + z) \in I_1 I_2 + I_1 I_3$, ce qui permet de conclure.

c) L'inclusion $I_2 + (I_1 \cap I_3) \subset I_1 \cap (I_2 + I_3)$ résulte de ce que I_2 et $I_1 \cap I_3$ sont contenus dans I_1 et dans $I_2 + I_3$.

D'autre part, $x \in I_1 \cap (I_2 + I_3)$, x s'écrit $x_2 + x_3$ avec $x_2 \in I_2$ et $x_3 \in I_3$, donc $x_3 = x - x_2 \in I_1$ puis l'inclusion réciproque de la précédente.

4) a) Il est clair que I_x est un idéal. S'il n'est pas maximal, il existe un idéal maximal M tel que $I_x \subsetneq M \subsetneq A$. Soit $f \in M - I_x$; on a $f(x) \neq 0$. L'égalité $1 = \left(1 - \frac{f}{f(x)}\right) + \frac{f}{f(x)}$ montre que 1 est la somme d'un élément de I_x et d'un élément de M , donc appartient à M , ce qui contredit le caractère maximal de M .

b) Si M n'est égal à aucun des I_x , pour tout $x \in E$, il existe $f_x \in M$ tel que $f_x(x) \neq 0$, donc il existe un voisinage U_x de x sur lequel f_x ne s'annule pas. Les U_x forment un recouvrement de E . Comme E est compact, on peut extraire de ce recouvrement un recouvrement fini $U_{x_1}, \dots, U_{x_n}$. La fonction $f = \sum_{1 \leq i \leq n} f_{x_i}^2$ est un élément de M jamais nul, donc inversible dans A , ce qui est impossible, d'où la conclusion.

5) a) Raisonner comme pour les idéaux maximaux d'un espace de fonctions (voir question précédente).

b) Prendre l'idéal des suites ayant un nombre fini de termes non nuls.

19.3 1) La division par $aX + b$ montre que tout polynôme P de $K[X]$ est équivalent à sa valeur en $-b/a$, mais cela ne donne pas immédiatement un isomorphisme de corps comme l'un des deux arguments suivants. L'isomorphisme $\varphi : K[X] \rightarrow K[X]$ défini par $\varphi(X) = aX + b$ montre que $K[X]/(aX + b)$ est isomorphe à $K[X]/(X)$ (voir

corollaire 1 de 19.3) donc à K . On peut également considérer l'homomorphisme d'anneaux $K[X] \rightarrow K$ d'évaluation en $-b/a$ et montrer que son noyau est l'idéal $(aX + b)$ et son image est K (voir corollaire 2 de 19.3).

2) Les idéaux $(X - 1)$ et $(X + 1)$ sont étrangers puisque leur différence est inversible. Le théorème chinois montre que

$$K[X]/(X^2 - 1) \simeq K[X]/(X - 1) \times K[X]/(X + 1) ;$$

chacun des deux facteurs est isomorphe à K d'après le 1), donc $K[X]/(X^2 - 1) \simeq K \times K$. L'isomorphisme est donné par

$$P \bmod X^2 - 1 \mapsto (P \bmod X - 1, P \bmod X + 1) \mapsto (P(1), P(-1)).$$

3) La même démarche qu'au **2)** montre que $\mathbb{C}[X]/(X^2 + 1) \simeq \mathbb{C} \times \mathbb{C}$.

4) Il suffit d'utiliser l'isomorphisme $K[X] \rightarrow K[X]$ défini par $X \mapsto X + 1$ et le corollaire 1 de 19.3.

5) Supposons d'abord $\text{car}(K) \neq 2$. Le premier anneau possède un élément nilpotent non nul : la classe de X dont le carré est nul. Montrons que le second anneau ne possède pas d'élément nilpotent non nul : si on note x la classe de X , un élément $ax + b$ de carré nul vérifie $a^2 x^2 + 2abx + b^2 = 0$, d'où $ab = 0$ et $b^2 - a^2 = 0$, ce qui implique $a = b = 0$. Les deux anneaux ne sont donc pas isomorphes. Supposons maintenant $\text{car}(K) = 2$; on a

$$\frac{K[X]}{(X^2 + 1)} = \frac{K[X]}{(X + 1)^2} \simeq \frac{K[X]}{(X^2)}.$$

6) Dans $\mathbb{Q}[X]/(X^2 - 3)$, il existe un élément de carré 3, la classe de X . Dans $\mathbb{Q}[X]/(X^2 - 5)$, notons x la classe de X et cherchons un élément de carré 3. On doit trouver $a, b \in \mathbb{Q}$ tels que $a^2 x^2 + 2abx + b^2 = 3$, c'est-à-dire $2abx + 5a^2 + b^2 = 3$, ce qui impose $ab = 0$ et $5a^2 + b^2 = 3$; si $a = 0$, on doit avoir $b^2 = 3$, si $b = 0$, on doit avoir $5a^2 = 3$. Ces deux égalités sont impossibles dans $\mathbb{Q}$: il suffit de regarder, si un tel a ou un tel b existait, la parité de la valuation en 3 de leurs deux membres. Les deux quotients ne sont donc pas isomorphes.

7) Les résultats de 19.5, 4) de la proposition 2, permettent d'écrire :

$$\frac{\mathbb{Z}[t]}{(2+t)} \simeq \frac{\mathbb{Z}[X]/(X^2 + 1)}{(2 + X)} \simeq \frac{\mathbb{Z}[X]}{(X^2 + 1, 2 + X)}.$$

d'où

$$\frac{\mathbb{Z}[t]}{(2+t)} \simeq \frac{\mathbb{Z}[X]/(X + 2)}{(X^2 + 1)} \simeq \mathbb{Z}/5\mathbb{Z}.$$

Dans cet isomorphisme, la classe de $a + bt$ donne successivement la classe de $a + bX$ puis la classe de $a - 2b \bmod 5$; on vérifie que

$$a + bt = a - 2b + (2 + t)b \text{ et que } 5 = (2 + t)(2 - t).$$

De même, $\mathbb{Z}[t]/(7 + t) \simeq \mathbb{Z}/50\mathbb{Z}$.

8) a) Comme 3 est inversible dans les deux premiers anneaux et pas dans le troisième (c'est un diviseur de 0), l'anneau $\mathbb{Z}[X]/(3X)$ n'est pas isomorphe aux deux autres.

On montre que les deux premiers anneaux sont isomorphes en considérant l'automorphisme $X \mapsto -X$ et en remarquant que $3(-X) + 1 = (-1)(3X - 1)$.

b) Si $k = 0 \bmod 3$, on a $k = 3l$, $A_k = \mathbb{Z}[X]/(3(X - l))$. L'isomorphisme $X \mapsto X + l$ montre que ce dernier anneau est isomorphe à $\mathbb{Z}[X]/(3X)$.

Si $k = 3l \pm 1$, l'isomorphisme $X \mapsto X + l$ montre que A_k est isomorphe à A_1 ou A_{-1} ; ces deux anneaux sont isomorphes, d'après le **a)**, donc les A_k pour $k \neq 0 \bmod 3$ sont isomorphes.

c) Les idéaux (2) et $(2X - 1)$ étant étrangers dans $\mathbb{Z}[X]$, puisque $1 = (X \times 2) - (2X - 1)$, le théorème chinois donne

$$A \simeq \frac{\mathbb{Z}[X]}{(2)} \times \frac{\mathbb{Z}[X]}{(2X - 1)} \simeq (\mathbb{Z}/2\mathbb{Z})[X] \times \mathbb{Z}[1/2].$$

9) a) L'image de $\pi \circ i$ est $\pi(I) = (I + J)/J$. Son noyau est $I \cap J$.

b) Le corollaire 2 de 19.3 donne le premier isomorphisme ; le second s'obtient en échangeant les rôles de I et de J .

19.4 1) On trouve par exemple : $1 = (2X - 1)^2 - 4X(X - 1)$.

2) Si $b - a$ est inversible dans A , $X - a$ et $X - b$ sont étrangers car $(X - a) - (X - b) = b - a$. Réciproquement, si $X - a$ et $X - b$ sont étrangers, il existe des polynômes U et V de $A[X]$ tels que $U(X - a) + V(X - b) = 1$. On a $U(b)(b - a) = 1$, ce qui prouve que $b - a$ est inversible dans A .

3) Les polynômes $X^2 + 1$ et $(X + 1)^2$ ne sont pas étrangers dans $\mathbb{Z}[X]$, car s'il existe des polynômes $U, V \in \mathbb{Z}[X]$ tels que $U(X^2 + 1) + V(X + 1)^2 = 1$, on trouve, en évaluant cette égalité en -1 , $2U(-1) = 1$, ce qui est impossible.

Les polynômes $X^2 + 1$ et $(X + 1)^2$ sont étrangers dans $\mathbb{Q}[X]$, car, en utilisant l'algorithme d'Euclide, on trouve $\left(1 + \frac{X}{2}\right)(X^2 + 1) - \frac{X}{2}(X + 1)^2 = 1$.

4) Il suffit de montrer que I et J^s sont étrangers, ce résultat prouvant qu'alors I^r et J^s sont étrangers. Si I et J sont étrangers, il existe $x \in I$ et $y \in J$ tels que $x + y = 1$. On a $(x + y)^s = 1$, égalité qui peut s'écrire sous la forme $ux + y^s = 1$ et montre que I et J^s sont étrangers.

19.5 1) On cherche x sous la forme $x = s_0 + 3s_1 + 21s_2$; on trouve successivement : $s_0 = 1, s_1 = 6, s_2 = 2$ (les calculs d'inverses modulo 7 et 13 sont simples). Finalement, $x = 61 \bmod 273$. Les entiers satisfaisant le système sont de la forme $61 + 273k$ avec $k \in \mathbb{Z}$.

2) Pour construire φ^{-1} , on résout les systèmes :

$$\begin{cases} x = 1 \bmod 3 \\ x = 0 \bmod 7 \\ x = 0 \bmod 13 \end{cases} \quad \begin{cases} x = 0 \bmod 3 \\ x = 1 \bmod 7 \\ x = 0 \bmod 13 \end{cases}$$

comme ci-dessus ou en recherchant des identités de Bézout. On trouve $\varphi^{-1}(1, 0, 0) = 91$, $\varphi^{-1}(0, 1, 0) = 78$; d'où $\varphi^{-1}(0, 0, 1) = 274 - 91 - 78 = 105$, puis $\varphi^{-1}(x_1, x_2, x_3) = 91x_1 + 78x_2 + 105x_3$. Pour le système de la première question, cette formule donne $x = 91 + 78 \times 5 + 105 \times 9 = 1426 = 61 \bmod 273$.

3) On peut chercher P sous la forme $P = s_0 + s_1 X + s_2 X(X - 1)$. On trouve successivement : $P(0) = s_0 = 3$, $P(1) = 3 + s_1 = 11$, d'où $s_1 = 8$, $P(4) = 3 + 32 + 12s_2 = 2$, d'où $s_2 = 10$ (on peut calculer l'inverse de 12 modulo 17 comme le produit des inverses de 4 et 3 ou appliquer la méthode générale de 12.12.3) ; d'où $P = 10X^2 - 2X + 3$. L'homomorphisme $\varphi : K[X] \rightarrow K \times K \times K$ définie par $P \mapsto (P(0), P(1), P(4))$ est linéaire, de noyau l'idéal engendré par le produit $X(X - 1)(X - 4)$; les polynômes P de $(\mathbb{Z}/17\mathbb{Z})[X]$ vérifiant les conditions du texte sont donc les polynômes de la forme $10X^2 - 2X + 3 + X(X - 1)(X - 4)Q(X)$ où Q est un polynôme quelconque de $(\mathbb{Z}/17\mathbb{Z})[X]$.

19.6 1) L'égalité $(1 - X)(1 + X + \dots + X^{n-1}) = 1 - X^n$ donne, par passage au quotient $(1 - x)(1 + x + \dots + x^{n-1}) = 1$, ce qui prouve que $1 - x$ est inversible dans A et que son inverse est $1 + x + \dots + x^{n-1}$.

2) Unités de $A = \mathbb{Z}[\sqrt{2}]$

a) On sait que $u = a + b\sqrt{2}$ est une unité de A si et seulement si $a^2 - 2b^2 = N(u) = \pm 1$; dans ce cas, l'inverse de u est $u^{-1} = N(u)(a - b\sqrt{2})$.

b) Facile.

c) Si $u = a + b\sqrt{2}$, $-u = -a - b\sqrt{2}$, $u^{-1} = N(u)(a - b\sqrt{2})$, $-u^{-1} = N(u)(-a + b\sqrt{2})$. Pour le plus grand $a_0 + b_0\sqrt{2}$ de ces quatre nombres, on a $a_0, b_0 \geq 0$.

d) Soit u une unité de A telle que $1 < u < w$; posons $u = a + b\sqrt{2}$. On a $-u, -u^{-1} < 0$ et $0 < u^{-1} < 1$ donc le plus grand des quatre nombres $\pm u$ et $\pm u^{-1}$ est u , donc $a, b > 0$, $u \geq 1 + \sqrt{2}$, ce qui contredit $1 < u < w$.

Si $u > 1$, il existe donc n tel que $w^n \leq u < w^{n+1}$, donc $\frac{u}{w^n}$ est une unité de A telle que $1 \leq \frac{u}{w^n} < w$; la seule possibilité est $u = w^n$. Si $0 < u < 1$, $u^{-1} > 1$ et on est

ramené au cas précédent. Enfin, si $u < 0$, on se ramène aux cas précédents en considérant $-u$.

Les éléments de $U(A)$ sont donc de la forme $\pm w^n$ avec $n \in \mathbb{Z}$.

e) Notons H le groupe formé par ± 1 , isomorphe à $\mathbb{Z}/2\mathbb{Z}$, et $K = G(w) \simeq \mathbb{Z}$. Les conditions de l'exercice 18.9 sont réunies, donc $U(A) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}$.

19.7 1) Il suffit d'appliquer les définitions.

2) a) Si P contient le produit $I_1 \dots I_n$ sans contenir aucun des idéaux $I_1, \dots, I_n$, il existe des éléments $x_k \in I_k \setminus P$ pour $k = 1, \dots, n$. Le produit des x_k est un élément du produit des I_k qui n'est pas dans P , puisque P est premier. Contradiction.

b) Si M' est un idéal maximal, donc premier, contenant M^n , le résultat précédent montre que $M' \supset M$; d'où $M' = M$.

3) Pour $n = 1$, il n'y a rien à démontrer. Pour $n = 2$, c'est l'exercice 19.2 2) a). Supposons $n \geq 3$.

On a $X = (X \cap P_1) \cup \dots \cup (X \cap P_n)$. S'il existe j tel que $X \cap P_j \subset \bigcup_{1 \leq i \leq n, i \neq j} (X \cap P_i)$, alors $X \subset \bigcup_{1 \leq i \leq n, i \neq j} (X \cap P_i)$ et l'hypothèse de récurrence s'applique. S'il n'existe pas de tel j , on choisit k tel que P_k soit premier et, pour $i = 1, \dots, n$, $i \neq k$, on choisit $y_i \in X \cap P_i$ n'appartenant à aucun P_j pour $j \neq i$. On pose $z = y_k + \prod_{1 \leq i \leq n, i \neq k} y_i$.

Les hypothèses montrent que $z \in X$ donc il devrait exister j tel que $z \in P_j$. Mais $j = k$

est impossible car, P_k étant premier, le produit $\prod_{1 \leq i \leq n, i \neq k} y_i$ n'est pas dans P_k alors

que $y_k y$ est et $j \neq k$ est aussi impossible car le produit $\prod_{1 \leq i \leq n, i \neq k} y_i$ est dans P_j alors que y_k n'y est pas.

19.8 1) Soit a un élément non nul de A ; la division euclidienne de 1 par a dans $A[X]$ donne $1 = aq + r$ avec $\deg(r) < \deg(a)$ ou $r = 0$; par conséquent, $r = 0$, a est inversible dans A et A est un corps.

2) On raisonne comme dans le cas de l'anneau $\mathbb{Z}[t]$ (voir 19.12). Soient $\alpha, \beta \in \mathbb{Z}[\sqrt{n}]$, avec $\beta \neq 0$. Le quotient α/β s'écrit $x + y\sqrt{n}$, avec $x, y \in \mathbb{Q}$. Notons a, b les entiers les plus proches de x, y respectivement et posons $q = a + b\sqrt{n}$, $r = \alpha - q\beta = (x - a + (y - b)\sqrt{n})\beta$. On a $|x - a| \leq 1/2$, $|y - b| \leq 1/2$, donc $|N(x - a + (y - b)\sqrt{n})| \leq 3/4$ pour les valeurs de n indiquées. On a $\alpha = q\beta + r$ avec $|N(r)| < |N(\beta)|$, ce qui prouve que $\mathbb{Z}[\sqrt{n}]$ est euclidien.

3) La norme d'un élément $a + \iota b\sqrt{3}$ est $a^2 + 3b^2$, ce qui montre que les éléments unités de A sont ± 1 et qu'il n'y a pas d'éléments de norme 2. Les éléments de norme 4 sont donc irréductibles. L'égalité $(1 + \iota\sqrt{3})(1 - \iota\sqrt{3}) = 2 \times 2$ où les termes des deux produits sont de norme 4 et non associés montre que A n'est pas factoriel, donc n'est pas euclidien.

19.9 1) On vérifie d'abord que ρ est une des deux racines de l'équation $x^2 - x + 5 = 0$, puis on considère l'homomorphisme $\varphi : \mathbb{Z}[X] \rightarrow \mathbb{C}$ défini par $\varphi(X) = \rho$. Son image est A ; il reste à déterminer son noyau. La division euclidienne d'un polynôme $P \in \mathbb{Z}[X]$ par $X^2 - X + 5$ s'écrit :

$$P = (X^2 - X + 5)Q + a + bX,$$

avec $a, b \in \mathbb{Z}$ et $Q \in \mathbb{Q}[X]$. On a $P \in \ker(\varphi)$ si et seulement si $\varphi(P) = a + b\rho = 0$, ce qui équivaut à $a = b = 0$, soit $P \in (X^2 - X + 5)$. On en déduit que $A \simeq \mathbb{Z}[X]/(X^2 - X + 5)$ et que les éléments de A s'écrivent sous la forme $a + b\rho$, avec $a, b \in \mathbb{Z}$.

$$\begin{aligned} \mathbf{2)} \text{ Voir que } N(a + b\rho) &= \left(a + b/2 + \frac{ib\sqrt{19}}{2} \right) \left(a + b/2 - \frac{ib\sqrt{19}}{2} \right) \\ &= (a + b/2)^2 + \frac{19b^2}{4} \geq 4b^2. \end{aligned}$$

Les éléments inversibles de A sont les éléments de norme 1, ce qui impose $b = 0$ et $a = \pm 1$.

3) Si B est un corps, il suffit de prendre $t = 0$. Sinon, avec le choix de t indiqué, tout $b \in B/(t)$ est de la forme $\pi(a)$; si la division euclidienne de a par t s'écrit $a = qt + r$, on a $b = \pi(r)$; comme $j(r) < j(t)$ ou $r = 0$, r est inversible ou nul.

4) Si A est euclidien, il existe donc un élément non inversible $t \in A$ pour lequel la restriction de la projection canonique $\pi : A \rightarrow A/(t)$ à $A^\times \cup \{0\}$ soit surjective. Mais $A^\times \cup \{0\} = \{-1, 0, 1\}$ d'après le **2)** ; comme t n'est pas inversible, $A/(t)$ n'est pas l'anneau nul. Les deux seules possibilités sont donc $A/(t) \simeq \mathbb{Z}/2\mathbb{Z}$ et $A/(t) \simeq \mathbb{Z}/3\mathbb{Z}$. Posons $\sigma = \pi(\rho)$. On doit avoir $\sigma^2 - \sigma + 5 = 0$, ce qui conduit dans le premier cas à $\sigma^2 + \sigma + 1 = 0$ qui n'a pas de solution dans $\mathbb{Z}/2\mathbb{Z}$ et dans le second à $\sigma^2 + 2\sigma + 2 = 0$ qui n'a pas de solution dans $\mathbb{Z}/3\mathbb{Z}$. Par conséquent, A n'est pas euclidien.

$$\mathbf{5)} \text{ D'après 19.5, on a les isomorphismes : } \frac{A}{2A} \simeq \frac{\mathbb{Z}[X]}{(X^2 - X + 5, 2)} \simeq \frac{(\mathbb{Z}/2\mathbb{Z})[X]}{(X^2 + X + 1)}.$$

Comme $X^2 + X + 1$ est irréductible dans $(\mathbb{Z}/2\mathbb{Z})[X]$, le quotient est un corps (à quatre éléments). On peut aussi raisonner directement : $a + b\rho$ avec $a, b \in \mathbb{Z}$ est égal, modulo 2, à l'un des quatre éléments $0, 1, \rho, 1 + \rho$. On a $\rho(1 + \rho) = \rho^2 + \rho = 2\rho - 5 = 1 \pmod{2}$, ce qui prouve que les images de ρ et $1 + \rho$ sont inversibles dans le quotient, qui est donc un corps. D'où la conclusion.

6) Le quotient α/β s'écrit $x + y\rho$ avec $x, y \in \mathbb{Q}$. Notons a, b les entiers les plus proches de x et y respectivement. On doit considérer trois cas :

a) $\alpha/\beta \in A$; alors $\alpha = q\beta$ avec $q = \alpha/\beta \in A$;

b) $\alpha/\beta \notin A$ et $|y - b| \leq 1/3$; posons $q = a + b\rho$, $r = \alpha - q\beta$; on a $N((x - a) + (y - b)\rho) \leq (x - a)^2 + |x - a||y - b| + 5(y - b)^2 \leq 1/4 + 1/6 + 5/9 = 35/36$,

donc $\alpha = q\beta + r$ avec $N(r) \leq 35N(\beta)/36 < N(\beta)$;

c) $\alpha/\beta \notin A$ et $|y - b| > 1/3$; on a $2\alpha/\beta = 2x + 2yp$ et $2y$ est à moins de $1/3$ de l'entier qui en est le plus proche ; ce qui précède permet d'affirmer qu'il existe q et r tels que $2\alpha = q\beta + r$ avec $N(r) < N(\beta)$.

7) Soient I un idéal de A , β un élément de norme minimale de $I \setminus \{0\}$. On a $(\beta) \subset I$. Supposons l'inclusion stricte et soit $\alpha \in I \setminus \beta$. On doit considérer deux cas.

a) La pseudodivision de α par β donne $\alpha = q\beta + r$; on a $r \neq 0$, donc $N(r) < N(\beta)$ et $r = \alpha - q\beta \in I$, ce qui contredit le choix de β .

b) La pseudodivision de α par β donne $2\alpha = q\beta + r$; comme $r \neq 0$ conduit encore à une contradiction, $r = 0$ et $2\alpha = q\beta$; comme l'idéal (2) est maximal, il est premier, donc q ou β est multiple de 2 ; si $q = 2q'$, $\alpha = q'\beta$ est dans (β) , contrairement à l'hypothèse ; si $q \notin (2)$, alors $\beta = 2\beta'$, $\alpha = q\beta'$; comme $q \notin (2)$, $(q) + (2) = A$, donc il existe u et v dans A tels que $uq + 2v = 1$, d'où $\beta' = uq\beta' + 2v\beta' = u\alpha + v\beta \in I$, ce qui contredit encore le choix de β .

On conclut que $(\beta) = I$. Par conséquent, A est un anneau principal.

19.10 1) Notons $(A_i)_{i \in I}$ une famille d'anneaux principaux et I un idéal de leur produit A . Notons $j_i : A_i \rightarrow A$ les inclusions. Pour tout $i \in I$, $j_i^{-1}(I \cap j_i(A_i))$ est un idéal (d_i) de A_i . Il est facile de vérifier que I est l'idéal engendré par l'élément $(j_i(d_i))$ de A .

2) D'abord, $(2, X)$ dans $\mathbb{Z}[X]$, (X, Y) dans $K[X, Y]$ sont des idéaux stricts. Si ce sont des idéaux principaux, un générateur ne sera pas inversible dans l'anneau. Si $(2, X)$ est un idéal principal (P) de $\mathbb{Z}[X]$, 2 est un multiple de P , donc $P = \pm 2$; or X n'est pas multiple de 2 , contradiction. Si (X, Y) est un idéal principal (P) de $K[X, Y]$, X est un multiple de P , donc $P = aX$ avec $a \in K^*$; or Y n'est pas multiple de aX , contradiction.

19.11 1) La proposition est vraie pour $c = 0$. Si $c \neq 0$, on voit d'abord que cm est un multiple commun de ca et de cb . Soit d est un multiple commun de ca et de cb . D'abord d est un multiple de c et s'écrit $d = ce$; on voit ensuite que e est multiple de a et b parce que A est intègre, donc un multiple de m ; par conséquent, d est multiple de cm et le résultat.

2) Montrons que $D(c) = D(a) \cap D(b)$. D'abord, il existe $u, v \in A$ tels que $c = ua + vb$. Comme $a, b \in (c)$, c divise a et b . Si c' divise a et b , c' divise $ua + vb = c$. Par conséquent, $D(a) \cap D(b) = D(c)$, ce qui prouve $c = \text{pgcd}(a, b)$.

19.12 1) Ce n'est pas un anneau intègre.

2) Voir l'exercice 19.8 3).

3) La norme est définie par $N(a + b\sqrt{-3}) = a^2 + 3b^2$. On en déduit que les unités de A sont ± 1 et que A ne contient pas d'élément de norme 2 . Montrons que l'idéal $I = (2, 1 + \sqrt{-3})$ n'est pas principal. Les deux éléments 2 et $1 + \sqrt{-3}$ sont de norme

4, non associés et premiers entre eux. Si I était principal, la seule possibilité est donc $I = A$. S'il existe des éléments $a + b\sqrt{-3}$, $a' + b'\sqrt{-3} \in A$ tels que $1 = 2(a + b\sqrt{-3}) + (a' + b'\sqrt{-3})(1 + \sqrt{-3})$, on trouve, en additionnant les deux relations qui en résultent, que 1 est un multiple de 2 : contradiction !

4) L'anneau A a pour éléments les polynômes de la forme $X^2P + X^3Q$ avec $P, Q \in \mathbb{R}[X]$. Il ne contient pas de polynôme de degré 1. Les égalités $X^6 = X^3 \times X^3$ et $X^6 = X^2 \times X^2 \times X^2$ donnent deux décompositions de X^6 en produits de facteurs irréductibles et non associés dans A .

5) Non. Les nombres $5 - 2\sqrt{2}$ et $7 + 4\sqrt{2}$ ont tous les deux pour norme 17 et sont associés : $7 + 4\sqrt{2} = (5 - 2\sqrt{2})(1 + \sqrt{2})^2$.

6) A est un anneau factoriel dont tout élément non nul admet une décomposition en produit de facteurs irréductibles. Comme il n'existe pas d'éléments irréductibles dans A , tout élément non nul de A est inversible, donc A est un corps.

19.13 Le corps des fractions de $\mathbb{Z}[\iota]$ est (à isomorphisme près) $\mathbb{Q}[\iota]$, celui de $\mathbb{Z}[\sqrt{n}]$ est $\mathbb{Q}[\sqrt{n}]$, ceux de $\mathbb{Z}[X]$ et de $\mathbb{Q}[X]$ sont égaux à $\mathbb{Q}(X)$.

19.14 1) Soient E et F deux ensembles. Notons $\mathcal{A}$ l'ensemble des couples (A, i) formés d'une partie A de E et d'une injection $i : A \rightarrow F$.

L'ensemble $\mathcal{A}$ est non vide, car il contient le couple formé de la partie vide de E et de l'application vide $\emptyset \rightarrow F$.

On définit un ordre sur $\mathcal{A}$ en posant que $(A, i) \preceq (A', i')$ si $A \subset A'$ et i' prolonge i . Pour cet ordre, l'ensemble $\mathcal{A}$ est inductif, car si $\{(A_k, i_k), k \in K\}$ est un sous-ensemble totalement ordonné de $\mathcal{A}$, on peut définir un couple (A, i) de $\mathcal{A}$ majorant tous les (A_k, i_k) en posant $A = \cup_k A_k$ et en définissant une injection $i : A \rightarrow F$ par $i(x) = i_k(x)$ si $x \in A_k$.

Le lemme de Zorn affirme alors qu'il existe un élément maximal dans ; notons-le (A, i) . Si $A = E$, le résultat est démontré, de même que si $i(A) = F$, en considérant i^{-1} qui donne une injection de F dans E . Aucun autre cas n'est possible. En effet, si $A \subsetneq E$ et $i(A) \subsetneq F$, il existe $x \in E - A$, $y \in F - i(A)$; on construit une injection $j : A \cup \{x\} \rightarrow F$ prolongeant i en posant $j(x) = y$; le caractère maximal de (A, i) est contredit.

2) Soient E un K -espace vectoriel, F un sous-espace de E . Notons $\mathcal{G}$ l'ensemble des sous-espaces G de E tels que $G \cap F = \{0\}$. Cet ensemble n'est pas vide, contenant l'espace réduit à 0. On peut l'ordonner par inclusion ; tout sous-ensemble $\{G_k, k \in K\}$ de $\mathcal{G}$ totalement ordonné par cet ordre est majoré par $\cup G_k$, sous-espace vectoriel de E ayant une intersection réduite à 0 avec F . Par conséquent, $\mathcal{G}$ admet un élément maximal. Notons-le G . Si G n'est pas un supplémentaire de F , il existe x dans E tel que $x \notin F + G$. Alors $G + Kx$ est un majorant de G dans $\mathcal{G}$, ce qui contredit le caractère maximal de G .

3) Notons $\mathcal{S}$ l'ensemble des systèmes libres de E contenant C et contenus dans A . Cet ensemble est non vide (il contient C) et ordonné par inclusion. Il est facile de vérifier que tout sous-ensemble totalement ordonné $\{S_k, k \in K\}$ d'éléments de $\mathcal{S}$ admet comme majorant $\cup S_k$. Par conséquent, $\mathcal{S}$ admet un élément maximal B .

D'une part, B est un système libre ; d'autre part, B est un système générateur de E , car, sinon, il existerait un vecteur $u \in A$ n'appartenant pas à l'espace engendré par B et $B \cup \{u\}$ appartiendrait à $\mathcal{S}$, ce qui contredirait le caractère maximal de B . Par conséquent, B est une base de E vérifiant les propriétés demandées.

4) Les seuls endomorphismes du $\mathbb{R}$ -espace vectoriel $\mathbb{R}$ sont les homothéties. Comme $\mathbb{Q}$ est un sous-espace (de dimension 1) du $\mathbb{Q}$ -espace vectoriel $\mathbb{R}$, il admet un supplémentaire F dans cet espace. On a donc une décomposition en somme directe de $\mathbb{Q}$ -sous-espaces : $\mathbb{R} = \mathbb{Q} \oplus F$. La projection sur $\mathbb{Q}$ répond à la question.

5) Soit A un anneau, P un idéal premier de A et $\mathcal{P}$ l'ensemble des idéaux premiers de A contenus dans P . Cet ensemble n'est pas vide, contenant P . On peut l'ordonner par l'ordre inverse de l'inclusion. Tout sous-ensemble totalement ordonné $\{P_k, k \in K\}$ d'éléments de $\mathcal{P}$ admet comme majorant (au sens de l'ordre inverse de l'inclusion) $Q = \cap_k P_k$. En effet, il est clair que Q est un idéal de A ; il reste à vérifier qu'il est premier. Si $xy \in Q$ et si $x \notin Q$, il existe k tel que $x \notin P_k$; pour tout P_l tel que $P_l \subset P_k$, on a $x \notin P_l$ et, comme les P_l sont premiers, on a $y \in P_l$ donc $y \in Q$ et on conclut.

OUVERTURES SUR LES POLYNÔMES

20

20.1 LA A -ALGÈBRE $A[X]$

Rappel sur la construction de l'anneau des polynômes à une indéterminée. Dans tout ce chapitre, nous continuerons à dire *anneau* pour *anneau commutatif unitaire* sauf mention du contraire.

Dans le chapitre 13, nous avons construit en détail l'algèbre des polynômes à une indéterminée sur un corps. Nous avons indiqué (13.2.2) que l'algèbre des polynômes à une indéterminée sur un anneau commutatif et unitaire A se construit exactement de la même façon : les polynômes sont des suites $(a_0, a_1, \dots)$ d'éléments de A , nulles à partir d'un certain rang. On définit l'addition et la multiplication de ces suites et, une fois la notation polynomiale $\sum_{0 \leq k \leq n} a_k X^k$ adoptée, tout se passe comme on en a l'habitude.

Proposition 1. *Soit A un anneau intègre. Alors $A[X]$ est un anneau intègre.*

Démonstration. Soient P et Q deux polynômes non nuls de $A[X]$ de termes dominants $a_m X^m$ et $b_n X^n$. Comme $a_m \neq 0$ et $b_n \neq 0$, on a $a_m b_n \neq 0$. Le produit PQ , dont le terme dominant est $a_m b_n X^{m+n}$, est donc non nul, ce qui prouve que $A[X]$ est un anneau intègre. $\square$

Par conséquent, si $PQ = 0$ dans $A[X]$, on a $P = 0$ ou $Q = 0$.

A-algèbre. Soit A un anneau (commutatif et unitaire). On appelle A -algèbre un couple (B, ρ) formé d'un anneau B (non nécessairement commutatif) et d'un homomorphisme d'anneaux $\rho : A \rightarrow B$.

Cette définition correspond bien à la définition 1 de 13.8. Elle équivaut à se donner sur B , d'une part une structure d'anneau, d'autre part une loi de multiplication par les éléments de A , en posant $a.b = \rho(a)b$ (le point sera le plus souvent omis), vérifiant les propriétés suivantes.

- a) Pour tout $b \in B$, on a $1_A.b = b$.
- b) Pour tous $a \in A, b, b' \in B$, on a $a.(b + b') = a.b + a.b'$.
- c) Pour tous $a, a' \in A, b \in B$, on a $(a + a').b = a.b + a'.b$.
- d) Pour tous $a, a' \in A, b \in B$, on a $a.(a'.b) = (aa').b$.
- e) Pour tous $a \in A, b, b' \in B$, on a $a.(bb') = (a.b)b' = b(a.b')$.

Le groupe $(B, +)$ muni de la multiplication par les éléments de l'anneau A vérifie des axiomes (ceux donnés en a, b, c, d) analogues à ceux des espaces vectoriels. On dit que B est ainsi muni d'une structure de A -module. Cette structure joue un rôle central en algèbre, mais elle ne sera pas abordée dans ce livre.

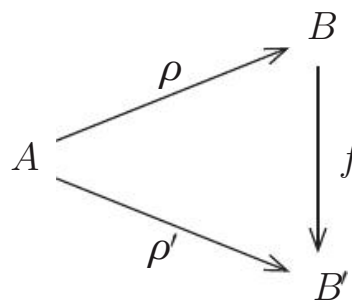
Le couple formé par $A[X]$ et la multiplication par les éléments de A est une A -algèbre.

Tout anneau A est une A -algèbre et possède une unique structure de $\mathbb{Z}$ -algèbre.

Si B est une A -algèbre et C une B -algèbre, alors C est une A -algèbre.

On peut aussi définir un produit de A -algèbres.

Homomorphisme de A -algèbres. Si (B, ρ) et (B', ρ') sont des A -algèbres, un homomorphisme de A -algèbres $f : (B, \rho) \rightarrow (B', \rho')$ est un homomorphisme d'anneaux $f : B \rightarrow B'$ tel que $f \circ \rho = \rho'$, autrement dit, tel que le diagramme



commute, ce qui revient à $f(a.b) = a.f(b)$ pour tous $a \in A$ et $b \in B$.

Sous- A -algèbre. Si B est une A -algèbre, une sous-algèbre de B est un sous-anneau C de B tel que l'homomorphisme $\rho : A \rightarrow B$ se factorise par l'inclusion $C \rightarrow B$, autrement dit tel que $a.c \in C$ pour tous $a \in A$ et $c \in C$.

Idéaux de A -algèbres. Soient B une A -algèbre. Un idéal I de l'anneau B est appelé idéal de l'algèbre B ; I est stable par la multiplication par tout élément de A .

Quotient de A -algèbres. Soient B une A -algèbre, I un idéal de B . En notant $\pi : B \rightarrow B/I$ la projection canonique, l'anneau quotient B/I est muni d'une structure de A -algèbre par l'homomorphisme composé : $\rho' = \pi \circ \rho : A \rightarrow B \rightarrow B/I$: la multiplication d'un élément $\pi(b)$ par un $a \in A$ donne

$$a \cdot \pi(b) = \rho'(a)\pi(b) = \pi(\rho(a)b) = \pi(a \cdot b).$$

Les propriétés de ce quotient sont analogues à celles vues pour les quotients de groupes ou d'anneaux. Les démonstrations sont analogues à celles pour les groupes ou les anneaux.

Proposition 2 : propriété universelle du quotient. Soient B une A -algèbre, I un idéal de B , $\pi : B \rightarrow B/I$ la projection canonique.

Pour toute A -algèbre C et tout homomorphisme de A -algèbres $f : B \rightarrow C$ tel que $f(I) = 0$, il existe un unique homomorphisme de A -algèbres $f' : B/I \rightarrow C$ tel que $f' \circ \pi = f$, autrement dit, tel qu'on ait un diagramme commutatif :

$$\begin{array}{ccc} B & \xrightarrow{f} & C \\ \pi \downarrow & \nearrow f' & \\ B/I & & \end{array}$$

□

Corollaire. Soient $f : B \rightarrow C$ un homomorphisme de A -algèbres et $\pi : B \rightarrow B/\ker(f)$ la projection canonique. La factorisation $f' : B/\ker(f) \rightarrow C$ de f par π est injective et définit un isomorphisme de A -algèbres φ de $B/\ker(f)$ sur $\text{im}(f)$. □

Proposition 3 : propriété universelle de l'algèbre $A[X]$. Soient A un anneau et B une A -algèbre. Pour tout élément $x \in B$, il existe un unique homomorphisme de A -algèbres $f : A[X] \rightarrow B$ tel que $f(X) = x$, défini par $f\left(\sum_{0 \leq k \leq n} a_k X^k\right) = \sum_{0 \leq k \leq n} a_k x^k$. □

Si $P = \sum_{0 \leq k \leq n} a_k X^k \in A[X]$, on écrira $f(P) = P(x) = \sum_{0 \leq k \leq n} a_k x^k$.

Sous-algèbre engendrée par un élément. Avec les notations précédentes, l'image de l'homomorphisme f est appelée sous-algèbre de B engendrée par x et notée $A[x]$. On ne confondra pas les sous-groupe, sous-anneau, sous-algèbre engendrés par x .

La propriété universelle est très utile. Comme en 13.8, elle permet de construire de nombreux homomorphismes : évaluation, fonction polynomiale associée, substitution d'un polynôme à une indéterminée...

Proposition 4 : changement d'anneau de coefficients. Soient A, B des anneaux, $\rho : A \rightarrow B$ un homomorphisme munissant B d'une structure de A -algèbre. L'homomorphisme ρ s'étend de façon unique en un homomorphisme de A -algèbres $\sigma : A[X] \rightarrow B[X]$ tel que $\sigma(X) = X$ et tel que le diagramme suivant soit commutatif.

$$\begin{array}{ccc} A & \xrightarrow{\rho} & B \\ i \downarrow & & \downarrow j \\ A[X] & \xrightarrow{\sigma} & B[X] \end{array}$$

Les flèches verticales i, j sont les homomorphismes associant à un élément de A ou de B le polynôme constant qu'il définit.

Démonstration. La proposition est une application de la propriété universelle de $A[X]$ (proposition 3), en considérant $B[X]$ comme une A -algèbre à l'aide du composé $j \circ \rho : A \rightarrow B[X]$. $\square$

Concrètement, l'application σ ne change que les coefficients des polynômes :

$$\sigma\left(\sum_{0 \leq k \leq n} a_k X^k\right) = \sum_{0 \leq k \leq n} \rho(a_k) X^k.$$

Exemple. Si on applique cette proposition avec $A = \mathbb{Z}$, $B = \mathbb{Z}/p\mathbb{Z}$, p premier, ρ égal à la projection canonique $\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$, on obtient un homomorphisme d'algèbres $\mathbb{Z}[X] \rightarrow \mathbb{Z}/p\mathbb{Z}[X]$ appelé réduction modulo p et qui va nous être souvent utile.

La réduction modulo 3 de $18X^4 - 13X^2 + 7$ est $2X^2 + 1$; dans cet exemple, la réduction des coefficients change le degré.

Proposition 5. Soient A un anneau et I un idéal de A . On note $I[X]$ l'ensemble des polynômes de $A[X]$ dont les coefficients sont dans I . Alors $I[X]$ est un idéal de $A[X]$

et il existe un isomorphisme de A -algèbres $\left(\frac{A}{I}\right)[X] \simeq \frac{A[X]}{I[X]}.$

Démonstration. L'homomorphisme $f : A[X] \rightarrow \left(\frac{A}{I}\right)[X]$ de changement d'anneau de la proposition précédente, défini par la projection canonique $\pi : A \rightarrow A/I$, est surjectif. Son noyau $I[X]$ est un idéal de $A[X]$. Le corollaire de la propriété universelle du quotient permet de conclure. $\square$

20.2 CORPS DE RUPTURE ET DE DÉCOMPOSITION

Soit K un corps, P un polynôme de $K[X]$ de degré $n > 0$, (P) l'idéal engendré par P dans $K[X]$, $A = K[X]/(P)$ l'anneau quotient. On note $\pi : K[X] \rightarrow A$ la projection

canonique et on pose $x = \pi(X)$. Comme P n'est pas inversible, l'idéal (P) est un idéal strict.

Proposition 1.

- 1) L'anneau A est muni d'une structure de K -algèbre, en particulier A est un K -espace vectoriel de dimension n . Une base de cet espace vectoriel est la famille $(x^k)_{0 \leq k \leq n-1}$.
- 2) Si (P) est irréductible, A est un corps, extension de degré n de K .
- 3) Si (P) est un produit $P = P_1 \cdot \dots \cdot P_r$ de polynômes irréductibles, deux à deux premiers entre eux, A est isomorphe à un produit de r corps.

Démonstration. 1) La structure de K -espace vectoriel de A se déduit de la structure de K -espace vectoriel de $K[X]$. Dans cet espace, la famille $(x^k)_{0 \leq k \leq n-1}$ est libre puisque si $\sum_{0 \leq k \leq n-1} a_k x^k = 0$, le polynôme $\sum_{0 \leq k \leq n-1} a_k X^k$ doit appartenir à (P) et ne peut donc qu'être nul pour des raisons de degré. Cette famille est également génératrice. En effet, tout élément de A est de la forme $\pi(S)$ avec $S \in K[X]$; la division euclidienne de S par P donne une égalité $S = PQ + R$ avec $\deg(R) < n$, d'où $\pi(S) = \pi(R)$ avec $\pi(R)$ qui appartient à l'espace engendré par la famille $(x^k)_{0 \leq k \leq n-1}$.

2) On sait que l'anneau $K[X]$ est principal (voir 13.4, proposition 1). On en déduit facilement (voir 19.11, propriété 4) que (P) est un idéal maximal de $K[X]$; donc A est un corps. L'homomorphisme de corps défini par le composé $K \rightarrow K[X] \rightarrow A = K[X]/(P)$ est injectif et permet de considérer A comme une extension de degré n de K (voir 19.1).

3) Comme les idéaux $(P_1), \dots, (P_r)$ sont deux à deux étrangers, le théorème chinois montre que A est isomorphe au produit $K[X]/(P_1) \times \dots \times K[X]/(P_r)$ dont chacun des facteurs est un corps d'après le 2). □

La proposition précédente permet de construire de nombreux corps, en variant le corps K et le polynôme irréductible P choisi. C'est ce que nous allons développer maintenant.

Corps de rupture. Soient K un corps et P un polynôme irréductible de $K[X]$. Une extension L de K est appelée corps de rupture de P si P a une racine a dans L et si $L = K(a)$.

Proposition 2 : Existence et unicité du corps de rupture. Soit K un corps et P un polynôme irréductible de $K[X]$.

- 1) Il existe un corps de rupture de P .
- 2) Deux corps de rupture de P sont isomorphes.

Démonstration. 1) L'existence d'un corps de rupture est facile à montrer : on prend $L = K[X]/(P)$. Si x est l'image de X par la projection $\pi : K[X] \rightarrow K[X]/(P)$, on a $P(x) = P(\pi(X)) = \pi(P(X)) = 0$. D'autre part, la proposition précédente montre que $L = K[X]$.

2) Soit L' un autre corps de rupture de K , $L' = K(a')$ avec $P(a') = 0$. L'homomorphisme de K -algèbres $f : K[X] \rightarrow L'$ défini par $f(X) = a'$ vérifie $f(P) = 0$ donc il existe un homomorphisme de K -algèbres $g : K[X]/(P) \rightarrow L'$ tel que $g \circ \pi = f$. Comme g est un homomorphisme de corps, il est injectif ; comme f est surjectif, g est donc un isomorphisme. $\square$

Le corps de rupture dépend du corps auquel appartiennent les coefficients du polynôme. Par exemple, $\mathbb{C}$ est un corps de rupture de $X^2 + 1$ considéré comme polynôme de $\mathbb{R}[X]$, $\mathbb{Q}[i]$ est un corps de rupture de $X^2 + 1$ considéré comme polynôme de $\mathbb{Q}[X]$. Un corps de rupture de $X^2 - 2$, considéré comme polynôme de $\mathbb{Q}[X]$, est $\mathbb{Q}[\sqrt{2}]$. Le polynôme $X^3 - 2$, considéré comme polynôme de $\mathbb{Q}[X]$, a plusieurs corps de rupture contenus dans $\mathbb{C}$: $\mathbb{Q}[\sqrt[3]{2}]$, $\mathbb{Q}[\zeta\sqrt[3]{2}]$, $\mathbb{Q}[\zeta^2\sqrt[3]{2}]$, isomorphes entre eux. On remarquera sur ce dernier exemple que le polynôme peut ne pas être complètement factorisé dans un corps de rupture ; par exemple, $X^3 - 2$ est un produit de deux polynômes irréductibles sur $\mathbb{Q}[\sqrt[3]{2}]$:

$$X^3 - 2 = (X - \sqrt[3]{2})(X^2 + \sqrt[3]{2}X + \sqrt[3]{4}).$$

Corps de décomposition. Soient K un corps et P un polynôme de $K[X]$. Une extension L de K est appelée corps de décomposition de P si P se factorise en produit de facteurs du premier degré dans $L[X]$ et si $L = K(E)$ où E note l'ensemble des racines de P dans L .

Un corps de décomposition d'un polynôme de $\mathbb{Q}[X]$ est facile à construire : on prend le sous-corps de $\mathbb{C}$ engendré par les racines du polynôme. Pour un corps K qui n'est pas contenu dans $\mathbb{C}$, on peut commencer par construire l'analogue de $\mathbb{C}$ pour K (clôture algébrique de K , voir mon livre *Théorie de Galois* Ch. 14) ou construire directement le corps de décomposition, ce que nous présentons maintenant.

Proposition 3 : Existence et unicité du corps de décomposition. Soit K un corps et P un polynôme de $K[X]$.

- 1) Il existe un corps de décomposition de P .
- 2) Deux corps de décomposition de P sont isomorphes.

Démonstration. On raisonne par récurrence sur le degré n de P .

- 1) Si $n \leq 1$, on prend $L = K$, de même que si toutes les racines de P sont dans K . Si $n > 1$, P a un facteur irréductible Q ; on construit le corps de rupture $L_1 = K[X]/(Q)$; on a $L_1 = K[a]$ (il est possible que $a \in K$ et $L_1 = K$). Dans L_1 , Q a une racine a et $P = (X - a)R$ avec $R \in L_1[X]$. L'hypothèse de récurrence montre l'existence d'un corps de décomposition L de R . On voit d'abord que, dans le

corps L , le polynôme P se factorise en produit de facteurs du premier degré. Si on note E l'ensemble des racines de R dans L , on a $L = L_1[E]$, donc $L = K[a][E] = K[\{a\} \cup E]$; L est un corps de décomposition de P .

- 2) Pour montrer l'unicité du corps de décomposition à isomorphisme près, on suppose que P admet deux corps de décomposition L et L' . Si $n = 1$, alors $L = L' = K$. Si $n > 1$, le facteur irréductible Q choisit ci-dessus a une racine a dans L et une racine a' dans L' et il existe un isomorphisme de corps $\varphi : K[a] \rightarrow K[a']$. Comme L et L' sont des corps de décomposition de R sur $K[a]$, on peut appliquer l'hypothèse de récurrence et conclure. $\square$

20.3 SI A FACTORIEL, ALORS $A[X]$ FACTORIEL

Le but de cette section est de démontrer le théorème donné dans le titre de la section : si A est un anneau factoriel, alors l'anneau des polynômes $A[X]$ est également factoriel.

La démonstration est assez délicate. Pour mieux la comprendre, nos lectrices et lecteurs peuvent se placer dans le cas de $A = \mathbb{Z}$, car la démarche est exactement la même que dans le cas général.

Contenu d'un polynôme. Soit A un anneau factoriel. On appelle contenu d'un polynôme non nul P de $A[X]$, le pgcd, noté $c(P)$, des coefficients de P . Rappelons qu'on dit le pgcd par abus de langage, car, pour toute unité $u \in A^\times$, $uc(P)$ est aussi pgcd des coefficients de P . Les égalités de contenu que nous écrirons seront toujours à une unité près.

Un polynôme de $A[X]$ dont le contenu est 1 est appelé primitif. Pour tout polynôme non nul P de $A[X]$, le polynôme $\frac{1}{c(P)}P$ est primitif.

Par exemple, le contenu du polynôme $3X^3 - 18X + 15$ de $\mathbb{Z}[X]$ est 3 (ou -3) et, en divisant par 3, on obtient le polynôme primitif $X^3 - 6X + 5$ (ou $-X^3 + 6X - 5$).

Si P est un polynôme constant non nul de $A[X]$, on peut prendre $c(P) = P$. Si $P = aP_1$ avec $a \in A^*$, alors $c(P) = ac(P_1)$.

Dans l'article 42 des *Recherches arithmétiques*, Gauss montre que le produit de deux polynômes à coefficients rationnels mais pas tous entiers ne peut être à coefficients entiers, ce qui résulte de la proposition suivante.

Proposition 1. *Soit A un anneau factoriel.*

- 1) *Le produit de deux polynômes primitifs de $A[X]$ est primitif.*
- 2) *Pour tous polynômes P, Q non nuls de $A[X]$, on a $c(PQ) = c(P)c(Q)$.*

Démonstration. 1) Soient P et Q des polynômes primitifs. Supposons que PQ ne soit pas un polynôme primitif, autrement dit que $c(PQ)$ soit non inversible dans A . Soit $p \in A$ un facteur irréductible de $c(PQ)$. L'idéal $I = pA$ est un idéal premier et

l'anneau A/I est intègre. Notons $\pi : A[X] \rightarrow (A/I)[X]$ l'homomorphisme de changement d'anneau induit par la projection canonique $A \rightarrow A/I$. Comme $\pi(PQ) = \pi(P)\pi(Q)$ est nul, $\pi(P)$ ou $\pi(Q)$ est nul, donc P ou Q est divisible par p . Le contenu de l'un des deux polynômes, divisible par p , n'est pas inversible, ce qui donne une contradiction et montre que PQ est un polynôme primitif.

2) Posons $P_1 = \frac{1}{c(P)}P$ et $Q_1 = \frac{1}{c(Q)}Q$. Les polynômes P_1 et Q_1 sont primitifs.

D'après le 1), $P_1 Q_1$ est primitif, d'où $PQ = c(P)c(Q)P_1 Q_1$ a pour contenu $c(PQ) = c(P)c(Q)$. $\square$

Proposition 2. Soient A un anneau factoriel, $K = \text{Cf}(A)$ son corps de fraction. Les polynômes $P \in A[X]$ irréductibles dans $A[X]$ sont :

- les polynômes constants $a \in A$ où a est un élément irréductible de A ;
- les polynômes primitifs $P \in A[X]$ non constants et tels que P soit irréductible dans $K[X]$.

Démonstration. 1) Montrons d'abord que les polynômes indiqués sont bien irréductibles dans $A[X]$.

Si $a \in A$ est un élément irréductible de A , c'est un polynôme irréductible dans $A[X]$, car si $a = PQ$ avec $P, Q \in A[X]$, on a $\deg(P) = \deg(Q) = 0$, donc $P, Q \in A$ et l'un des deux est inversible dans A , donc dans $A[X]$.

Si $P \in A[X]$ vérifie la seconde condition et si $P = QR$ avec $Q, R \in A[X]$, la même égalité a lieu dans $K[X]$. Alors Q ou R est une unité de $K[X]$, par exemple R , ce qui impose $\deg(R) = 0$; R est donc un élément a non nul de A . On a $P = aQ$ dans $A[X]$, et a divise le contenu de P ; comme P est primitif, a est inversible dans A , donc P est irréductible dans $A[X]$.

2) Montrons maintenant la réciproque. Soit $P \in A[X]$ un polynôme irréductible. Si $\deg(P) = 0$, c'est un élément irréductible de A . Si $\deg(P) \geq 1$, P est un polynôme primitif. Montrons que P est irréductible dans $K[X]$. Si $P = QR$ avec $Q, R \in K[X]$, on peut écrire, en réduisant les coefficients de Q au même dénominateur et en faisant

de même pour ceux de R , $Q = \frac{a}{b}Q_1, R = \frac{c}{d}R_1$, avec $a, b, c, d \in A$ et $Q_1, R_1 \in A[X]$

primitifs. On a alors $bdP = acQ_1R_1$; en prenant les contenus des deux membres, on voit que ac et bd sont associés, d'où $P = uQ_1R_1$ avec u unité de A ; comme P est irréductible dans $A[X]$, l'un des polynômes Q_1 ou R_1 est une unité de $A[X]$, ce qui montre l'irréductibilité de P dans $K[X]$. $\square$

Théorème. Soit A un anneau factoriel. Alors $A[X]$ est factoriel.

Démonstration. Notons K le corps de fraction de A .

Il faut montrer que tout polynôme P de $A[X]$ admet une décomposition unique en produit de facteurs irréductibles. Nous allons montrer l'existence de la décomposition au 1) et 2), l'unicité au 4) à l'aide d'un lemme donné au 3).

- 1) Supposons P primitif. Comme $K[X]$ est euclidien, donc factoriel, il existe une décomposition de P en produit de facteurs irréductibles dans $K[X]$: $P = \prod_{1 \leq i \leq r} P_i$. En réduisant les coefficients de chacun des P_i au même dénominateur, on peut écrire $P_i = \frac{a_i}{b_i} Q_i$, avec Q_i primitif et irréductible dans $A[X]$. On a $a \left(\prod_{1 \leq i \leq r} b_i \right) P = \prod_{1 \leq i \leq r} a_i Q_i$, d'où, en simplifiant par les contenus des deux membres, $P = v \prod_{1 \leq i \leq r} Q_i$ où v est une unité de A ; ceci montre l'existence d'une décomposition de P dans $A[X]$ en produit de facteurs irréductibles dans le cas où P primitif.
- 2) Si P n'est pas primitif, on écrit $P = c(P)P_1$ et on décompose $c(P)$ dans A et P_1 comme précédemment pour obtenir une décomposition de P .
- 3) Montrons qu'un élément irréductible F de $A[X]$ est premier dans $A[X]$. Si F est un élément irréductible p de A , on a $\frac{A[X]}{pA[X]} = \frac{A}{pA}[X]$ (proposition 4 de 19.1) ; comme ce dernier anneau est intègre, p est premier dans $A[X]$. Si F n'est pas dans A , la proposition 2 montre que $c(f) = 1$ et que F est irréductible dans $K[X]$. Si F divise dans $A[X]$ un produit $G H$ de deux polynômes de $A[X]$, F divise $G H$ dans $K[X]$; comme $K[X]$ est factoriel, F divise G , par exemple, et il existe un polynôme G_1 de $K[X]$ tel que $FG_1 = G$; en réduisant au même dénominateur les coefficients de G_1 et en prenant les contenus, on voit que $G_1 \in A[X]$, ce qui prouve que F est premier dans $A[X]$.
- 4) Supposons que P admettent deux décompositions en produit de facteurs irréductibles dans $A[X]$: $P = u \prod_{1 \leq i \leq r} P_i = v \prod_{1 \leq j \leq s} T_j$. Après avoir simplifié par les facteurs associés des deux membres, s'il reste un facteur P_i , il est premier d'après le 3), donc divise un des facteurs de droite T_j ; ces deux facteurs sont associés, contradiction et fin de la démonstration. $\square$

20.4 RECHERCHE DES FACTEURS IRREDUCTIBLES D'UN POLYNÔME

Dans cette section, K est un corps de caractéristique 0.

Comme $K[X]$ est un anneau factoriel, tout polynôme $P \in K[X]$ s'écrit comme produit de polynômes irréductibles. Cette décomposition est toute théorique : comment faire, devant un polynôme donné, pour le décomposer en produit de polynômes irréductibles de $K[X]$?

Nous n'allons aborder qu'un petit aspect du problème dans cette section.

Soit $P \in K[X]$. En regroupant les facteurs irréductibles égaux ou associés, on peut écrire P sous la forme :

$$P = \prod_{1 \leq i \leq r} p_i^{k_i} \quad (1)$$

où les p_i sont irréductibles et premiers entre eux deux à deux. Posons $P_k = \prod_{i: k_i=k} p_i$; P_k est le produit des polynômes p_i qui apparaissent à la puissance k dans (1). Si la valeur maximale des k_i est s , on a $P = P_1 P_2^2 \dots P_s^s$. Les P_k , $1 \leq k \leq s$, sont premiers entre eux deux à deux ; nous allons montrer qu'on peut tous les calculer (à une unité près).

L'algorithme d'Euclide permet de déterminer $Q = \text{pgcd}(P, P')$ et $R = \frac{P}{Q}$. On va voir que Q et R s'expriment en fonction des P_k .

$$\begin{aligned} \text{On a } P' &= P_1' P_2^2 \dots P_s^s + 2P_1 P_2 P_2' \dots P_s^s + \dots + sP_1 P_2^2 \dots P_s^{s-1} P_s' \\ &= P_2 P_3^2 \dots P_s^{s-1} (P_1' P_2 \dots P_s + 2P_1 P_2' \dots P_s + \dots + sP_1 P_2 \dots P_s'). \end{aligned}$$

Pour trouver Q , on cherche les p irréductibles divisant P Qui divisent P' .

- 1) Si p divise P_1 , p ne divise pas P_1' , donc ne divise pas P' .
- 2) Si p divise P_2 , il ne divise pas P_2' , donc p divise P' et p^2 ne divise pas P' .
- 3) En général, si p divise P_k il ne divise pas P_k' , donc p^{k-1} divise P' et p^k ne divise pas P' .

On conclut que $Q = P_2 P_3^2 \dots P_s^{s-1}$ et $R = P_1 P_2 \dots P_s$.

De même, on trouve $S = \text{pgcd}(Q, Q') = P_3 \dots P_s^{s-2}$, $T = \frac{Q}{S} = P_2 \dots P_s$. On en déduit $\frac{R}{T} = P_1$. Ainsi, le calcul de P_1 est possible et, de proche en proche, on peut calculer les polynômes $P_2, \dots, P_s$. Mais pour le moment, on ne sait rien faire de plus ; il faut attendre 21.5.

20.5 DÉCOMPOSITION EN ÉLÉMENTS SIMPLES DANS $\mathbb{C}(X)$ ET $\mathbb{R}(X)$

Introduction. En 1702, Leibniz et Jean Bernoulli publièrent des méthodes pour intégrer les fractions rationnelles. Elles s'appuyaient sur un résultat dont ils n'avaient pas de démonstration, le théorème que nous appelons souvent en France « de d'Alembert » qui leur servait pour ramener l'intégration des fractions rationnelles au calcul de primitives de fractions rationnelles particulières, les éléments simples.

Bon nombre de nos lectrices et lecteurs ont sans doute déjà effectué des décompositions de fractions rationnelles en éléments simples. Elles, ils ont dû éprouver un petit sentiment de bricolage pour calculer des coefficients en multipliant par des $(x - a)^k$ puis en faisant $x = a$. La décomposition de la fraction rationnelle dans le corps des fractions rationnelles évite ces difficultés. Nous allons faire un peu d'algèbre linéaire, sans insister sur les détails. L'entraînement sportif à ce genre de calculs a été beaucoup trop mis en avant dans l'enseignement. Une calculatrice de poche ou un logiciel de calcul formel peuvent aujourd'hui aider à calculer la plupart de ces décompositions.

Les anneaux $\mathbb{C}[X]$ et $\mathbb{R}[X]$ sont des anneaux factoriels dont les éléments irréductibles sont les polynômes du premier degré dans le cas de $\mathbb{C}[X]$, les polynômes du premier degré et ceux du second degré sans racine réelle dans le cas de $\mathbb{R}[X]$ (voir 13.7).

Décomposition en somme directe. Le K -espace vectoriel $K(X)$ admet une décomposition en somme directe : $K(X) = K[X] \oplus F$, où F est le sous-espace vectoriel des fractions rationnelles $\frac{U}{V}$ telles que $\deg(U) < \deg(V)$. En effet, d'une part, si un polynôme P de $K[X]$ est égal à une fraction rationnelle $\frac{U}{V} \in F$ avec $\deg(U) < \deg(V)$, on a $PV = U$, donc $\deg(P) + \deg(V) = \deg(U)$, égalité impossible si P n'est pas nul, ce qui prouve $K[X] \cap F = \{0\}$. D'autre part, si $\frac{P}{Q} \in K(X)$, la division euclidienne de P par Q s'écrit $P = P_1 Q + R$ avec $\deg(R) < \deg(Q)$, et on a $\frac{P}{Q} = P_1 + \frac{R}{Q}$ (P_1 est appelée partie entière de la fraction rationnelle) ; par conséquent, $K(X) = K[X] + F$ et la conclusion.

20.5.1 Décomposition en éléments simples dans $\mathbb{C}(X)$

On appelle éléments simples de $\mathbb{C}(X)$ les fractions rationnelles de la forme $\frac{A}{(X-a)^k}$, où $a, A \in \mathbb{C}$ avec $A \neq 0$ et k entier strictement positif.

Proposition. Toute fraction rationnelle de $\mathbb{C}(X)$ est de manière unique somme d'un polynôme et d'éléments simples de $\mathbb{C}(X)$.

Démonstration. Comme $\mathbb{C}(X) = \mathbb{C}[X] \oplus F$, il suffit de montrer la proposition pour les éléments de F . Soit $\frac{R}{Q}$, avec $n = \deg(Q) > \deg(R) \geq 0$ une fraction rationnelle de F et notons F_Q le sous-espace des fractions rationnelles de F dont le dénominateur est Q . On peut supposer Q unitaire. Nous allons utiliser un argument d'algèbre linéaire.

Soient $a_i, 1 \leq i \leq s$, les racines complexes de Q , a_i étant de multiplicité r_i . On a $\sum_{1 \leq i \leq s} r_i = n$ et $Q = \prod_{1 \leq i \leq s} (X - a_i)^{r_i}$.

Pour $i = 1, \dots, s$, notons E_i la famille $\left(\frac{1}{X - a_i}, \dots, \frac{1}{(X - a_i)^{r_i}} \right)$ et montrons que la famille de vecteurs E formé par la réunion des E_i est libre dans F_Q .

Considérons une combinaison linéaire nulle de vecteurs de E :

$$0 = \lambda_{1,1} \frac{1}{X - a_1} + \dots + \lambda_{1,r_1} \frac{1}{(X - a_1)^{r_1}} + \lambda_{2,1} \frac{1}{X - a_2} + \dots$$

Multiplions par $(X - a_1)^{r_1}$. La combinaison linéaire s'écrit $0 = \lambda_{1,r_1} + (X - a_1)W$ où W est une fraction rationnelle dont les différents termes peuvent s'évaluer en a_1 . Par conséquent, le produit $(X - a_1)W$ vaut 0 en a_1 et $0 = \lambda_{1,r_1}$. On peut donc mettre en place un raisonnement par récurrence pour montrer que tous les $\lambda_{i,j}$ sont nuls. La famille E est donc libre dans F_Q et possède n éléments.

D'autre part, l'espace F_Q est engendré par la famille $\frac{X^k}{Q}$ avec $0 \leq k < n$ qui possède n éléments. On en déduit que E est une base de F_Q et qu'il existe des suites de nombres complexes $A_{i,1}, \dots, A_{i,r_i}$ pour $1 \leq i \leq s$ telles que :

$$\frac{R}{Q} = \frac{A_{1,1}}{X - a_1} + \dots + \frac{A_{1,r_1}}{(X - a_1)^{r_1}} + \frac{A_{2,1}}{X - a_2} + \dots$$

où les derniers pointillés de la formule indiquent, pour $i = 2, \dots, s$, des sommes de termes analogues à celle écrite pour $i = 1$.

Techniques de décomposition en éléments simples dans $\mathbb{C}(X)$. Une fois calculée la partie entière de la fraction rationnelle, on écrit la forme de la décomposition cherchée :

$$\frac{R}{Q} = \lambda_{1,1} \frac{1}{X - a_1} + \dots + \lambda_{1,r_1} \frac{1}{(X - a_1)^{r_1}} + \lambda_{2,1} \frac{1}{X - a_2} + \dots$$

La multiplication par $(X - a_1)^{r_1}$ et l'évaluation du résultat en a_1 donnent λ_{1,r_1} . On obtient de même les λ_{1,r_i} pour $i = 2, \dots, s$. Pour obtenir les autres coefficients, on peut calculer la différence de $\frac{R}{Q}$ et des termes déjà trouvés pour se ramener à la décomposition d'une fraction rationnelle avec un dénominateur de degré inférieur. Il existe d'autres méthodes : il est usuel, dans les exercices, d'évaluer les deux membres de l'égalité en des valeurs particulières bien choisies pour obtenir des équations entre les coefficients, puis de résoudre le système obtenu.

20.5.2 Décomposition en éléments simples dans $\mathbb{R}(X)$

On appelle éléments simples de $\mathbb{R}(X)$ les fractions rationnelles de la forme $\frac{A}{(X - a)^k}$, où $a, A \in \mathbb{R}$ avec $A \neq 0$ et k entier strictement positif et les fractions rationnelles de la forme $\frac{BX}{(X^2 + aX + b)^k}$ ou $\frac{C}{(X^2 + aX + b)^k}$, où $a, b, B, C \in \mathbb{R}$ avec $B, C \neq 0$, $X^2 + aX + b$ irréductible sur $\mathbb{R}$ et k entier strictement positif.

Proposition. *Toute fraction rationnelle de $\mathbb{R}(X)$ est de manière unique somme d'un polynôme et d'éléments simples de $\mathbb{R}(X)$.*

Démonstration. Elle est analogue à celle donnée pour $\mathbb{C}(X)$; on a encore $\mathbb{R}(X) = \mathbb{R}(X) \oplus F$. Soient $\frac{R}{Q} \in F$ et F_Q comme précédemment. On suppose que la décomposition de Q dans $\mathbb{R}[X]$ s'écrit :

$$Q = \prod_{1 \leq i \leq s} (X - a_i)^{r_i} \prod_{1 \leq j \leq u} (X^2 - b_j X + c_j)^{t_j}$$

avec $r_i, t_j \geq 1$.

On considère la réunion E des familles :

- $E_i = \left(\frac{1}{X - a_i}, \dots, \frac{1}{(X - a_i)^{r_i}} \right), 1 \leq i \leq s,$
- $E'_j = \left(\frac{1}{X^2 + b_j X + c_j}, \dots, \frac{1}{(X^2 + b_j X + c_j)^{t_j}} \right), 1 \leq j \leq u,$
- $E''_j = \left(\frac{X}{X^2 + b_j X + c_j}, \dots, \frac{X}{(X^2 + b_j X + c_j)^{t_j}} \right), 1 \leq j \leq u,$

Pour montrer qu'elle est libre, on écrit une combinaison linéaire nulle à coefficients réels de ces vecteurs. On voit que le coefficient de $\frac{1}{(X - a_1)^{r_1}}$ est nul en multipliant par $(X - a_1)^{r_1}$ et en évaluant en a_1 . Pour montrer que les coefficients de $\frac{1}{(X^2 + b_j X + c_j)^{t_j}}$ et $\frac{X}{(X^2 + b_j X + c_j)^{t_j}}$ sont nuls, on multiplie par $(X^2 + b_j X + c_j)^{t_j}$ et on évalue en une des deux racines non réelles de $X^2 + b_j X + c_j$. Etc.

On termine la démonstration en utilisant un argument de dimension : E est une base de F_Q car $\dim(F_Q) = s + 2u = \deg(Q)$. $\square$

Techniques de décomposition en éléments simples dans $\mathbb{R}(X)$. On suivra les indications données pour la décomposition dans $\mathbb{C}(X)$, en ajoutant de multiplier par $(X^2 + b_j X + c_j)^{t_j}$ et d'évaluer en une des deux racines non réelles de $X^2 + b_j X + c_j$ pour obtenir simultanément les coefficients de $\frac{1}{(X^2 + b_j X + c_j)^{t_j}}$ et $\frac{X}{(X^2 + b_j X + c_j)^{t_j}}$.

Exemple de calcul de décomposition. La décomposition dans $\mathbb{R}(X)$ de la fraction rationnelle $\frac{3X^2 + X + 1}{(X + 1)^2(X^2 + 1)^2}$ est de la forme

$$\frac{A_1}{X + 1} + \frac{A_2}{(X + 1)^2} + \frac{B_1 X + C_1}{X^2 + 1} + \frac{B_2 X + C_2}{(X^2 + 1)^2}.$$

En multipliant par $(X + 1)^2$ et en évaluant en -1 , on trouve $A_2 = -\frac{3}{4}$. En multipliant par $(X^2 + 1)^2$ et en évaluant en i , on trouve $B_2 = -\frac{1}{2}$ et $C_2 = -1$. Tous calculs faits, on est ramené à

$$\frac{3X+11}{4(X+1)(X^2+1)} = \frac{A_1}{X+1} + \frac{B_1X+C_1}{X^2+1}.$$

En multipliant par $X+1$ et en évaluant en -1 , on trouve $A_1 = 1$. En multipliant par X^2+1 et en évaluant en i , on trouve $B_1 = -1$ et $C_1 = \frac{7}{4}$. Finalement :

$$\frac{3X^3+X+1}{(X+1)^2(X^2+1)^2} = \frac{1}{X+1} - \frac{3}{4(X+1)^2} + \frac{-4X+7}{4(X^2+1)} + \frac{-X-2}{2(X^2+1)^2}.$$

Bien sûr, d'autres méthodes sont possibles...

20.6 MÉTHODES POUR PROUVER L'IRRÉDUCTIBILITÉ D'UN POLYNÔME DE $\mathbb{Z}[X]$, DE $\mathbb{Q}[X]$

Les logiciels actuels de calcul formel disposent d'une version d'un algorithme, mis au point en 1967 par Elwyn Berlekamp (né en 1940), permettant de factoriser un polynôme de $(\mathbb{Z}/p\mathbb{Z})[X]$ (avec p premier), $\mathbb{Z}[X]$ ou $\mathbb{Q}[X]$ en produit de facteurs irréductibles. Nous en présenterons certains aspects au chapitre 21.

Nous avons démontré plus haut (voir 20.3) qu'un polynôme de $\mathbb{Z}[X]$, dont le contenu est 1, est irréductible dans $\mathbb{Q}[X]$ si et seulement s'il est irréductible dans $\mathbb{Z}[X]$. L'étude de l'irréductibilité dans $\mathbb{Q}[X]$ peut donc toujours se ramener à l'étude de l'irréductibilité dans $\mathbb{Z}[X]$.

Donnons différentes méthodes pratiques pour étudier *à la main* l'irréductibilité d'un polynôme de $\mathbb{Z}[X]$.

Critère d'Eisenstein. Soit $P(X) = \sum_{0 \leq k \leq n} a_k X^k$ dans $\mathbb{Z}[X]$. Le polynôme P est irréductible dans $\mathbb{Q}[X]$ dans la circonstance suivante : il existe un nombre premier p qui ne divise pas a_n , divise tous les autres coefficients de P et est tel que p^2 ne divise pas a_0 . Si, de plus, $c(P) = 1$, alors P est irréductible dans $\mathbb{Z}[X]$.

Démonstration. Notons $\sigma : \mathbb{Z}[X] \rightarrow (\mathbb{Z}/p\mathbb{Z})[X]$ l'homomorphisme de réduction modulo p des coefficients (voir 20.1 proposition 4). Supposons que P soit réductible et qu'on ait $P = ST$ dans $\mathbb{Z}[X]$. Dans $(\mathbb{Z}/p\mathbb{Z})[X]$, cette égalité donne $(a_n \bmod p)X^n = \sigma(S)\sigma(T)$. Comme $\mathbb{Z}/p\mathbb{Z}$ est un corps, $(\mathbb{Z}/p\mathbb{Z})[X]$ est factoriel et $\sigma(S)$ et $\sigma(T)$ sont, à des inversibles près, des puissances de X ; les termes constants de S et T sont donc des multiples de p ; leur produit, qui est le terme constant de P est multiple de p^2 , contradiction. $\square$

Applications du critère d'Eisenstein. On peut parfois appliquer le critère à $P(X+a)$ pour une valeur de a bien choisie. L'exemple classique est celui du polynôme

cyclotomique $\Phi_p(X) = \frac{X^p - 1}{X - 1} = \sum_{0 \leq k \leq p-1} X^k$ pour p premier. En effet, $\Phi_p(X+1) = \sum_{1 \leq k \leq p} C_p^k X^{k-1}$ et les $C_p^k = \frac{p!}{k!(p-k)!}$ pour $1 \leq k \leq p$ satisfont les conditions d'application du critère d'Eisenstein. Par conséquent, Φ_p est irréductible dans $\mathbb{Z}[X]$ et dans $\mathbb{Q}[X]$.

Le critère d'Eisenstein est encore valable si on remplace $\mathbb{Z}$ par un anneau factoriel A , $\mathbb{Q}$ par le corps des fractions de A et si on suppose l'existence d'un élément p premier (ce qui équivaut à p irréductible) dans A vérifiant les conditions ci-dessus.

Autres méthodes. Les conditions d'application du critère d'Eisenstein sont rarement réunies et il est bon de disposer d'une petite panoplie d'outils divers. Soit $P(X) = \sum_{0 \leq k \leq n} a_k X^k$ un polynôme de degré n de $\mathbb{Z}[X]$.

- 1) Pour voir si P a une racine rationnelle, on suppose que $\frac{p}{q}$ est une racine rationnelle de P mise sous forme de fraction non simplifiable ; $q^n P(\frac{p}{q})$ est un entier et est nul ; on en déduit que q divise a_n et que p divise a_0 , ce qui permet d'obtenir l'ensemble des rationnels candidats à être racine de P . Si cet ensemble n'a pas trop d'éléments, on peut les tester un à un. Si P est unitaire, ses racines rationnelles, s'il en existe, sont entières (voir exercice 13.5).
- 2) Si P est de degré 2 ou 3 et n'a pas de racine rationnelle, il est irréductible dans $\mathbb{Q}[X]$.
- 3) Si P est de degré 4, n'a pas de racine rationnelle ni de décomposition en produit de deux facteurs du second degré dans $\mathbb{Q}[X]$ (pour le voir, on peut essayer une méthode de coefficients indéterminés), il est irréductible dans $\mathbb{Q}[X]$.
- 4) Le passage aux anneaux $\mathbb{Z}/p\mathbb{Z}$ est souvent utile et plusieurs idées différentes sont à mettre en œuvre.

S'il existe un nombre premier p tel que l'image de P dans $(\mathbb{Z}/p\mathbb{Z})[X]$ soit un polynôme irréductible de même degré, P est lui-même irréductible dans $\mathbb{Q}[X]$. La réciproque est fautive, l'exemple classique est le polynôme $X^4 + 1$, dont on montre (voir exercice 21.3) qu'il est réductible modulo p pour tout p premier alors qu'il est irréductible dans $\mathbb{Z}[X]$.

Si l'image de P dans $(\mathbb{Z}/p\mathbb{Z})[X]$ est un polynôme de degré strictement inférieur, on ne peut rien conclure, comme le montre l'exemple de $(pX + 1)(X + 1)$.

L'irréductibilité de P dans $(\mathbb{Z}/p\mathbb{Z})[X]$ peut se prouver avec des méthodes analogues aux précédentes.

Notons enfin, sans vouloir épuiser la liste des méthodes, qu'une factorisation dans $\mathbb{R}[X]$ peut aussi apporter des informations.

20.7 LOCALISATION DES RACINES D'UN POLYNÔME DE $\mathbb{R}[X]$

Cette section propose d'étudier une des étapes pour déterminer toutes les racines réelles d'un polynôme P à coefficients réels. La démarche est la suivante.

- 1) On détermine une borne M , plus ou moins grossière, peu importe, des valeurs absolues de toutes les racines de P . On sait alors que les racines de P sont dans l'intervalle $[-M, M]$.
- 2) Dans l'intervalle $[-M, M]$, on cherche à déterminer des intervalles dans lesquels P possède une unique racine. Cette seconde étape s'appuie sur le théorème de Sturm que nous allons présenter plus loin.
- 3) On est alors prêt pour la troisième étape : appliquer une méthode de calcul approchée comme la méthode de Newton pour déterminer chacune des racines de P .

La méthode de Newton consiste à utiliser la suite $u_{n+1} = u_n - \frac{P(u_n)}{P'(u_n)}$.
Des détails se trouvent dans les livres d'Analyse¹.

La seconde étape donne un algorithme pour déterminer si un polynôme à coefficients réels a une racine réelle (rappelons que le problème de savoir si un polynôme à coefficients complexes a une racine complexe est résolu par le théorème de d'Alembert : il suffit que $\deg(P) \geq 1$).

Les logiciels de calcul formel ont des procédures pour calculer les racines d'une équation polynomiale. Mais ils peuvent donner une expression compliquée d'une racine réelle où la réalité de celle-ci ne soit pas visible. On utilise l'algorithme de Sturm pour savoir si les racines sont ou non réelles en les comptant dans des intervalles bien choisis.

Proposition : borne pour les racines d'un polynôme de $\mathbb{R}[X]$. Soient $P = \sum_{0 \leq k \leq n} a_k X^k$ un polynôme de degré n à coefficients réels et α une racine réelle de P . Posons $S = \sum_{0 \leq k \leq n-1} \frac{|a_k|}{|a_n|}$ et $M = 1 + S$. Alors $|\alpha| \leq M$.

Démonstration. Pour x réel, on peut écrire $P(x)$ sous la forme

$$P(x) = a_n x^n \left(1 + \sum_{0 \leq k \leq n-1} \frac{a_k}{a_n} \frac{1}{x^{n-k}} \right).$$

Pour $|x| > M$, on a $\frac{1}{|x|} < \frac{1}{M}$, donc $\frac{1}{|x|^{n-k}} \leq \frac{1}{M}$ pour $0 \leq k \leq n-1$. On en déduit

$$\left| \sum_{0 \leq k \leq n-1} \frac{a_k}{a_n} \frac{1}{x^{n-k}} \right| < \frac{S}{M} < 1. \text{ Par conséquent, } P(x) \neq 0.$$

Voilà qui nous suffira pour passer à la seconde étape, mais il faut souligner que la formule précédente n'est pas très bonne et qu'il en existe d'autres.

1. Voir mon livre *Toute l'Analyse de la licence*, Dunod, 2014.

Suite de Sturm. Soient P un polynôme de $\mathbb{R}[X]$ sans racines multiples et a, b deux réels non racines de P avec $a < b$.

Comme P n'a pas de racines multiples, $\text{pgcd}(P, P')$ est une constante non nulle. L'algorithme de recherche du pgcd définit une suite de polynômes $R_0 = P, R_1 = P', R_2, \dots, R_s$, le dernier étant une constante non nulle, tels que $R_k = R_{k+1} Q_{k+2} + R_{k+2}$ pour $k = 0, \dots, s-2$ et $\deg(R_{k+2}) < \deg(R_{k+1})$.

Il faut modifier les signes des polynômes de la suite $(R_0, R_1, \dots, R_s)$ pour démontrer le résultat que nous avons en vue. La suite de Sturm associée à P est la suite $(P_0, P_1, \dots, P_s)$ définie pour $0 \leq k \leq s$ de la manière suivante : $P_k = R_k$ si $k = 0, 1 \bmod 4$, $P_k = -R_k$ si $k = 2, 3 \bmod 4$. En posant, pour $2 \leq k \leq s$, $T_k = Q_k$ si k est pair et $T_k = -Q_k$ si k est impair, on a, par exemple :

$$\begin{aligned} P_0 &= P_1 T_2 - P_2, \\ P_1 &= P_2 T_3 - P_3, \\ P_2 &= P_3 T_4 - P_4, \\ P_3 &= P_4 T_5 - P_5, \end{aligned}$$

et on vérifie que $P_k = P_{k+1} T_{k+2} - P_{k+2}$ pour $k = 0, \dots, s-2$. Les détails précédents sont donnés pour montrer la construction de la suite ; on a $P_k = -P_{k+2} \bmod P_{k+1}$ pour $k \geq 0$.

Comme P et P' sont premiers entre eux, on sait que deux termes consécutifs de la suite de Sturm de P sont premiers entre eux.

Nombre de changement de signes. Soient $u = (u_0, \dots, u_s)$ une suite de nombre réels. On appelle changement de signe dans la suite u la donnée de deux indices $j < l$ tels que $u_j u_l < 0$ et $u_k = 0$ pour tout k tel que $j < k < l$ (s'il existe de tels indices, c'est-à-dire si $l > j+1$). Le nombre de changements de signes de la suite u sera noté $v(u)$. Par exemple, si $u = (0, 0, 2, 0, 0, 0, -1, -2, 3, 0)$, on a $v(u) = 2$ puisqu'il y a un changement de signe de 2 à -1, puis un changement de signe de -2 à 3. En fait, on peut supprimer tous les zéros de la suite pour compter les changements de signe.

Remarquons que le nombre de changements de signes reste le même quand on multiplie certains polynômes de la suite par des nombres strictement positifs, ce qui permet de simplifier parfois les calculs à la main.

Théorème de Sturm. Soit P un polynôme de $\mathbb{R}[X]$ sans racines multiples. On note $(P_0, P_1, \dots, P_s)$ la suite de Sturm associée à P et on définit, pour tout réel x , la suite $u_x = (P_0(x), P_1(x), \dots, P_s(x))$. Soient $a < b$ deux réels non racines de P . Le nombre de racines de P dans l'intervalle $[a, b]$ est $v(u_a) - v(u_b)$.

Démonstration. Il va suffire d'étudier $v(u_x)$ quand x varie dans $[a, b]$. D'abord, v est une fonction constante dans tout intervalle ne contenant aucune racine des polynômes de la suite de Sturm.

Chapitre 20 • Ouvertures sur les polynômes

Si α est racine de P , il n'est pas racine de P' , ce qui montre que P' garde un signe constant au voisinage de α ; on a les deux possibilités suivantes pour le tableau de variation de P :

	α		
P	–	0	+
P'	+	+	+

	α		
P	+	0	–
P'	–	–	–

Si α est racine de P_k , $0 < k < s$, il n'est pas racine de P_{k-1} ni de P_{k+1} et $P_{k-1}(\alpha)P_{k+1}(\alpha) < 0$ par construction de la suite de Sturm. On a, par exemple, le tableau de variation :

	α		
P_{k-1}	+	+	+
P_k	+	0	–
P_{k+1}	–	–	–

Les seules variations de $v(x)$ ont donc lieu au passage d'une racine de P , qui provoque une diminution de 1, et on peut conclure. $\square$

Corollaire. Reprenons les notations précédentes et notons u_∞ la suite des coefficients dominants des polynômes de la suite $(P_0, P_1, \dots, P_s)$ et $u_{-\infty}$ la suite des coefficients dominants des polynômes de la suite $P_0(-X), P_1(-X), \dots, P_s(-X)$. Le nombre de racines réelles de P est $v(u_{-\infty}) - v(u_\infty)$.

Démonstration. Soit M un réel tel que toutes les racines réelles de P appartiennent à l'intervalle $] -M, M[$. Le théorème de Sturm montre que la différence $v(u_M) - v(u_{-M})$ donne le nombre de racines réelles de P et que $v(u_x)$ est constant pour $x \geq M$. Comme pour x suffisamment grand, les signes de la suite $u_x = (P_0(x), P_1(x), \dots, P_s(x))$ sont ceux des coefficients dominants des polynômes de la suite, on a $v(u_M) = v(u_\infty)$. De même, $v(u_{-M}) = v(u_{-\infty})$ et la conclusion.

Exemples. Soit $P = X^2 + X + 1$. On trouve $P' = 2X + 1$, $P = \left(\frac{X}{2} + \frac{1}{4}\right)P' - \left(-\frac{3}{4}\right)$. La suite de Sturm de P est $\left(X^2 + X + 1, 2X + 1, -\frac{3}{4}\right)$. On a $u_\infty = \left(1, 2, -\frac{3}{4}\right)$, $u_{-\infty} = \left(1, -2, -\frac{3}{4}\right)$; donc $v(u_{-\infty}) - v(u_\infty) = 0$, ce qui est bien le nombre de racines réelles de P .

Soit $P = X^3 + X + 1$. La suite de Sturm de P est, à des facteurs strictement positifs près, $(X^3 + X + 1, X^2 + 1/3, -X - 3/2, -1)$. On trouve $v(u_{-\infty}) = 2$, $v(u_\infty) = 1$, ce qui signifie que P a une racine réelle. Pour mieux la localiser, on peut calculer $v(u_0) = 1$,

ce qui montre que la racine est négative. La borne M donnée par la proposition est 3 ; la racine de P est donc dans l'intervalle $[-3, 0]$. Comme $v(u_{-1}) = 2$, on peut préciser que la racine de P est dans l'intervalle $[-1, 0]$.

Bien sûr, pour ces deux exemples, l'emploi de la suite de Sturm est inadapté : le discriminant de $X^2 + X + 1$ montre tout de suite que ce polynôme n'a pas de racine réelle, une étude rudimentaire de la fonction $x \mapsto x^3 + x + 1$ montre qu'elle s'annule une seule fois, entre -1 et 0 .

Un peu d'histoire. En 1637, dans sa *Géométrie*, Descartes énonce une règle pour savoir *Combien il peut y avoir de vraies racines en chaque équation. Il y en peut avoir, écrit-il, autant de vraies racines que les signes + et - s'y trouvent de fois être changés, et autant de fausses qu'il s'y trouve de fois deux signes + ou deux signes - qui s'entre-suivent* (une racine réelle est vraie pour Descartes si elle est positive, fausse si elle est négative). Voilà une belle règle, énoncée abruptement après un exemple. Descartes ne précise nullement s'il en a ou non une démonstration, ce n'est pas le style de l'époque. Mais elle va stimuler diverses recherches (Fourier, Cauchy, Gauss, etc.). Le travail de Charles-François Sturm (1803-1855), en 1829, lui donne une forme définitive ; Hermite disait que c'était un *exemple rare de simplicité et d'élégance, qui ouvre l'ère nouvelle de l'algèbre moderne* ; il est à la base des résultats de logique mathématique d'Alfred Tarski (1902-1983) sur l'élimination des quantificateurs dans les années 1930 et de méthodes de géométrie algébrique réelle (travaux de Michel Coste, Marie-Françoise Roy, etc.)



René Descartes (1596-1650)

Gravure extraite de *The Hundred Greatest Men*, D. Appleton & Company, 1885.

20.8 POLYNÔMES À PLUSIEURS INDÉTERMINÉES

Comme pour les polynômes à une indéterminée, on peut se demander à quoi sert une construction rigoureuse puisqu'on sait manipuler des polynômes à plusieurs indéterminées à coefficients réels ou complexes depuis longtemps sans difficulté. On ne peut être tout à fait d'accord avec cette opinion, car ce qui a été manipulé en fait, ce sont des fonctions polynomiales, mais les éléments de l'anneau des polynômes n'ont jamais été vraiment définis. Une fois la construction théorique mise au point, on pourra continuer à calculer plus ou moins comme on en a l'habitude (attention aux problèmes liés à la caractéristique du corps), mais en sachant que c'est justifié.

Étant donné un anneau, on sait construire l'algèbre des polynômes à une indéterminée à coefficients dans cet anneau. À partir de l'anneau A , cette construction donne la A -algèbre $A[X]$; à partir de l'anneau $A[X]$, elle donne la A -algèbre $A[X][Y]$ des polynômes à une indéterminée Y et à coefficients dans $A[X]$. Un tel polynôme s'écrit $\sum_{0 \leq k \leq n} b_k(Y)X^k$ où les b_k sont des polynômes de $A[X]$. Ce qui est gênant dans cette construction, c'est le rôle disymétrique joué par X et Y . Si on prend l'anneau $A[Y]$ comme anneau des coefficients, on peut construire la A -algèbre $A[Y][X]$ dont les éléments sont les polynômes $\sum_{0 \leq k \leq n} b_k(Y)X^k$ où les b_k sont des polynômes de $A[Y]$. On peut concilier ces deux constructions en construisant directement une troisième A -algèbre : $A[X, Y]$, et en remarquant que les trois algèbres vérifient la même propriété universelle et sont donc isomorphes.

La construction de $A[X, Y]$ est analogue à celle de $A[X]$: on prend l'ensemble des suites doubles d'éléments de A ayant un nombre fini de termes non nuls et on définit sur cet ensemble des lois d'addition, de multiplication (c'est le plus délicat) et de multiplication par un scalaire. Quand on note X, Y les suites dont les seuls termes, non nuls et égaux à 1, sont ceux d'indice $(1, 0)$ pour X , $(0, 1)$ pour Y , tout élément de $A[X, Y]$ s'écrit alors $\sum_{0 \leq k, l \leq n} a_{k,l} X^k Y^l$.

Les éléments de $A[X, Y]$ sont appelés polynômes à deux indéterminées à coefficients dans A .

Proposition : Propriété universelle de la A -algèbre $A[X, Y]$. Soit A un anneau.

1) Pour toute A -algèbre B et tout couple (x, y) d'éléments de B , il existe un unique homomorphisme de A -algèbres $f : A[X, Y] \rightarrow B$ tel que $f(X) = x, f(Y) = y$.

Cet homomorphisme est défini par :

$$f\left(\sum_{0 \leq k, l \leq n} a_{k,l} X^k Y^l\right) = \sum_{0 \leq k, l \leq n} a_{k,l} x^k y^l.$$

2) Propriété d'unicité de $A[X, Y]$ à isomorphisme unique près. Soit C une A -algèbre possédant la même propriété que $A[X, Y]$, ce qui signifie que pour toute A -algèbre B et tout couple (x, y) d'éléments de B , il existe un unique homomorphisme de

A-algèbres $f: C \rightarrow B$ tel que $f(X) = x, f(Y) = y$. Alors il existe un unique isomorphisme de *A*-algèbres $\varphi: A[X, Y] \rightarrow C$.

Démonstration. 1) Il suffit de généraliser les démonstrations données en 13.8, 20.1 et 18.1. $\square$

Il est facile de vérifier alors que $A[X][Y]$ et $A[Y][X]$ vérifient aussi la propriété universelle précédente, en construisant en deux temps l'homomorphisme f ; par exemple, à partir de $\rho: A \rightarrow B$, on construit $f_1: A[X] \rightarrow B$ prolongeant ρ et tel que $f_1(X) = x$, puis $f_2: A[X][Y] \rightarrow B$ prolongeant f_1 et tel que $f(Y) = y$.

Le 2) de la propriété universelle (qui sert enfin) montre alors que $A[X][Y]$ et $A[Y][X]$ sont isomorphes à $A[X, Y]$.

On peut généraliser ce qui précède à un nombre quelconque fini r d'indéterminées pour obtenir l'algèbre $A[X_1, \dots, X_r]$, montrer qu'elle vérifie une propriété universelle permettant aux indéterminées de prendre des valeurs fixées, en déduire qu'elle est isomorphe à bien d'autres algèbres, comme $A[X_1, \dots, X_{r-1}][X_r]$, ce qui permet de voir $A[X_1, \dots, X_r]$ comme une *A*-algèbre obtenue par constructions successives d'algèbres de polynômes : $A, A[X_1], A[X_1][X_2], (A[X_1][X_2])[X_3] \dots$

La construction par récurrence précédente permet de démontrer sans peine certaines propriétés de $A[X_1, \dots, X_r]$. Toute propriété de *A* qui passe d'un anneau à l'anneau de polynômes sur cet anneau sera vraie dans $A[X_1, \dots, X_r]$. Par exemple :

- si *A* est intègre, alors $A[X_1, \dots, X_r]$ est intègre (voir 20.1, proposition 1);
- si *A* est factoriel, alors $A[X_1, \dots, X_r]$ est factoriel (voir 20.3).

On peut même construire une algèbre de polynômes dont les indéterminées sont indexées par un ensemble infini I : c'est l'union des algèbres $A[X_i]_{i \in J}$ où J parcourt les sous-ensembles finis de I . Dans un polynôme de cette algèbre n'apparaissent qu'un nombre fini d'indéterminées.

Dans une algèbre de polynômes, un polynôme de la forme $aX_1^{i_1} \dots X_s^{i_s}$ est appelé monôme. Tout polynôme est une somme de monômes.

Notons enfin que la *A*-algèbre des polynômes à 0 indéterminée s'identifie aux éléments de *A*.

Degré d'un polynôme à plusieurs indéterminées. Soit $P \in A[X, Y]$. Si on le considère comme polynôme de $A[X][Y]$ on peut définir son degré en tant que polynôme en Y . Si on le considère comme polynôme de $A[Y][X]$ on peut définir son degré en tant que polynôme en X . Par exemple, $P = X^2Y^2 + Y^2 - XY^3 - XY$ s'écrit $Y^2X^2 - (Y^3 + Y)X + Y^2$ comme polynôme en X ; son degré comme polynôme en X est 2. Il s'écrit aussi $-XY^3 + (X^2 + 1)Y^2 - XY$; son degré comme polynôme en Y est 3.

On définit aussi le degré total d'un monôme $aX_1^{i_1} \dots X_s^{i_s}$ comme la somme $i_1 + \dots + i_s$ et le degré total d'un polynôme comme le maximum des degrés totaux de ses monômes.

Polynôme homogène. Un polynôme homogène est un polynôme dont tous les monômes ont le même degré total. Par exemple, $P(X, Y, Z) = aX^3 + bXYZ + cY^2Z$, avec $a, b, c \in A$ est un polynôme homogène de degré 3 de $A[X, Y, Z]$.

Construction du corps des fractions rationnelles à n indéterminées. Soient A un anneau intègre, K son corps des fractions. La construction donnée en 19.15 permet de construire le corps des fractions rationnelles à n indéterminées à coefficients dans A : c'est le corps des fractions de l'anneau intègre $A[X_1, \dots, X_n]$. On vérifie facilement que c'est aussi le corps des fractions de $K[X_1, \dots, X_n]$, ce qui généralise un résultat de 19.15. On peut évaluer une fraction rationnelle en un n -uplet $(a_1, \dots, a_n)$ de A^n si elle peut s'écrire $\frac{P}{Q}$ avec $Q(a_1, \dots, a_n) \neq 0$.

20.9 POLYNÔMES SYMÉTRIQUES

Soit $n \geq 1$ un entier. Rappelons qu'on note S_n le groupe des permutations de l'ensemble $\{1, \dots, n\}$.

Polynômes symétriques. Soit A un anneau. La propriété universelle de l'algèbre $A[X_1, \dots, X_n]$ montre que, pour toute permutation $\sigma \in S_n$, il existe un unique homomorphisme de A -algèbres $\varphi_\sigma : A[X_1, \dots, X_n] \rightarrow A[X_1, \dots, X_n]$ rendant commutatif le diagramme

$$\begin{array}{ccc} \{1, \dots, n\} & \xrightarrow{\sigma} & \{1, \dots, n\} \\ \downarrow & & \downarrow \\ A[X_1, \dots, X_n] & \xrightarrow{\varphi_\sigma} & A[X_1, \dots, X_n] \end{array}$$

où les deux flèches verticales sont égales à l'application $i \mapsto X_i$.

Autrement dit : $\varphi_\sigma(X_i) = X_{\sigma(i)}$ pour $i = 1, \dots, n$ et, plus généralement :

$$\varphi_\sigma(P(X_1, \dots, X_n)) = P(X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

pour tout polynôme $P \in A[X_1, \dots, X_n]$.

L'application $\sigma \mapsto \varphi_\sigma$ définit une action du groupe S_n sur l'ensemble des polynômes à n indéterminées. Un point fixe pour cette action est appelé polynôme symétrique ; autrement dit, un polynôme P de $A[X_1, \dots, X_n]$ est dit symétrique si, pour tout $\sigma \in S_n$, on a $\varphi_\sigma(P) = P$; P est donc symétrique si et seulement si, pour tout monôme $x(X_1)^{k_1} \dots (X_n)^{k_n}$ de P et pour tout $\sigma \in S_n$, le monôme $xX_{\sigma(1)}^{k_1} \dots X_{\sigma(n)}^{k_n}$ est un monôme de P .

Si A est un anneau intègre de corps de fractions K , l'homomorphisme φ_σ se prolonge au corps $K(X_1, \dots, X_n)$ des fractions rationnelles en $X_1, \dots, X_n$ à coefficients dans K par :

$$\varphi_\sigma \left(\frac{P(X_1, \dots, X_n)}{Q(X_1, \dots, X_n)} \right) = \frac{P(X_{\sigma(1)}, \dots, X_{\sigma(n)})}{Q(X_{\sigma(1)}, \dots, X_{\sigma(n)})}$$

pour tous polynômes $P, Q \in A[X_1, \dots, X_n]$ avec $Q \neq 0$.

Exemples. Soit A un anneau. Dans $A[X_1, X_2, X_3]$ et dans $\mathbb{Z}[X_1, X_2, X_3]$ en particulier, les polynômes suivants sont symétriques :

- $X_1 X_2 X_3$;
- $X_1 + X_2 + X_3$;
- $X_1 X_2 + X_2 X_3 + X_3 X_1$;
- $X_1^2 X_2 + X_2^2 X_3 + X_3^2 X_1 + X_1 X_2^2 + X_2 X_3^2 + X_3 X_1^2$.
- Par contre, le polynôme $P = X_1^3 X_2 + X_2^3 X_3 + X_3^3 X_1$ n'est pas symétrique, puisque, pour $\sigma = (1\ 2)$, le monôme $\varphi_\sigma(X_1^3 X_2) = X_2^3 X_1$ n'est pas un monôme de P .

Polynômes symétriques élémentaires. Soient A un anneau et n un entier ≥ 0 . Pour tout entier k , le polynôme symétrique élémentaire de degré k de $A[X_1, \dots, X_n]$, que nous noterons s_k , est défini par :

$$s_k = \sum_{H \subset \{1, \dots, n\}, |H|=k} \prod_{i \in H} X_i,$$

autrement dit, H décrit l'ensemble des parties à k éléments de $\{1, \dots, n\}$ et s_k est la somme des produits des X_i pour $i \in H$. Pour tout n , $s_0 = 1$ et, pour $k > n$, $s_k = 0$.

Les s_k sont des polynômes homogènes de degré k .

En général, l'entier n et l'anneau A sont indiqués par le contexte et ne seront pas précisés.

Exemples.

- Pour $n = 0$: $s_0 = 1$, $s_k = 0$ pour $k \geq 1$.
- Pour $n = 1$: $s_0 = 1$, $s_1 = X_1$, $s_k = 0$ pour $k \geq 2$.
- Pour $n \geq 2$: $s_0 = 1$; $s_1 = \sum_{1 \leq i \leq n} X_i$; $s_2 = \sum_{1 \leq i < j \leq n} X_i X_j$; c'est la somme des produits deux à deux des X_i ;
- Plus généralement, s_k est la somme des produits de k des X_i ;
- Enfin, pour tout $n \geq 0$, $s_n = \prod_{1 \leq i \leq n} X_i$.

20.9.1 Produit des $X - X_i$ et relations entre coefficients et racines

Proposition. Soit $n \geq 0$ un entier. Dans l'anneau $\mathbb{Z}[X_1, \dots, X_n]$, on a :

$$\prod_{1 \leq i \leq n} (X - X_i) = \sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k}$$

Démonstration. Montrons cette formule par récurrence sur n .

Elle est vraie pour $n = 0, 1$. Supposons-la vraie pour un entier $n \geq 1$; notons s_k les polynômes symétriques élémentaires de $\mathbb{Z}[X_1, \dots, X_n]$ et t_k ceux de $\mathbb{Z}[X_1, \dots, X_{n+1}]$. On a :

$$\begin{aligned} \prod_{1 \leq i \leq n+1} (X - X_i) &= \left(\prod_{1 \leq i \leq n} (X - X_i) \right) (X - X_{n+1}) \\ &= \left(\sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k} \right) (X - X_{n+1}) \\ &= \sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k+1} - \sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k} X_{n+1} \\ &= X^{n+1} + \sum_{1 \leq k \leq n} (-1)^k [s_k + s_{k-1} X_{n+1}] X^{n+1-k} \\ &\quad + (-1)^{n+1} s_n X_{n+1}. \end{aligned}$$

D'une part, $s_n X_{n+1} = t_{n+1}$. D'autre part, pour $1 \leq k \leq n$, on voit, en distinguant les parties à k éléments de $\{1, \dots, n+1\}$ qui contiennent $n+1$ et celles qui ne le contiennent pas, que $s_k + s_{k-1} X_{n+1} = t_k$. C'est gagné.

Corollaire. Soit K un corps et soit $P(X) = \sum_{0 \leq i \leq n} a_i X^i$ un polynôme unitaire de degré n de $K[X]$. Si P s'écrit $P(X) = \prod_{1 \leq i \leq n} (X - x_i)$ où $x_1, \dots, x_n$ sont les racines (distinctes ou non) de P dans K , on a, pour $0 \leq k \leq n$

$$a_{n-k} = (-1)^k s_k(x_1, \dots, x_n).$$

Démonstration. Les propriétés universelles de $\mathbb{Z}[X_1, \dots, X_n]$ et de $\mathbb{Z}[X_1, \dots, X_n][X]$ donnent un homomorphisme $\varphi : \mathbb{Z}[X_1, \dots, X_n][X] \rightarrow K[X]$ tel que $\varphi(X) = X$ et $\varphi(X_i) = x_i$ pour $1 \leq i \leq n$. Comme $P(X) = \varphi\left(\prod_{1 \leq i \leq n} (X - X_i)\right)$ et que les $s_k(X_1, \dots, X_n)$ sont envoyés par φ sur les $s_k(x_1, \dots, x_n)$, on obtient le résultat. $\square$

Exemples. Si on note x_1, x_2 les racines complexes du polynôme $X^2 + aX + b$ de $\mathbb{C}[X]$, on a :

- $x_1 + x_2 = -a$;
- $x_1 x_2 = b$.

Si on note x_1, x_2, x_3 les racines complexes du polynôme $X^3 + aX^2 + bX + c$ de $\mathbb{C}[X]$, on a :

- $x_1 + x_2 + x_3 = -a$;
- $x_1x_2 + x_2x_3 + x_3x_1 = b$;
- $x_1x_2x_3 = -c$.

20.9.2 Polynômes symétriques et polynômes symétriques élémentaires

Soient $n \geq 1$ un entier et $s_1, \dots, s_n$ les polynômes symétriques élémentaires de $A[X_1, \dots, X_n]$. Pour tout polynôme T de $A[X_1, \dots, X_n]$, le polynôme $T(s_1, \dots, s_n)$ est symétrique. L'énoncé réciproque est utilisé librement au XVIII^e siècle ; Lagrange le qualifie d'*évident par soi-même*. Il est démontré indépendamment par Edward Waring (1736-1798) et Vandermonde en 1770.

Théorème. Soient A un anneau (le résultat sera donc vrai en particulier pour les corps) et P un polynôme symétrique de $A[X_1, \dots, X_n]$. Il existe un unique polynôme T de $A[s_1, \dots, s_n]$ tel que $T(s_1, \dots, s_n) = P$.

Exemples.

- $X_1^2 + X_2^2 = (X_1 + X_2)^2 - 2X_1X_2 = s_1^2 - 2s_2$.
- $X^3Y + XY^3 = XY(X^2 + Y^2) = s_1^2s_2 - 2s_2^2$.

Démonstration. 1) Prouvons l'existence du polynôme T , en donnant un algorithme permettant de le calculer.

Remarquons d'abord que, par linéarité, il suffit de montrer le résultat pour des polynômes symétriques $S(M)$ définis par les monômes $M = (X_1)^{k_1} \dots (X_n)^{k_n}$ en permutant les indéterminées de toutes les manières possibles et en faisant la somme des différents monômes obtenus : $S(M)$ est la somme des éléments de l'orbite de M sous l'action de S_n , autrement dit, $S(M) = \sum_{U \in O_M} U$ avec $O_M = \{\varphi_\sigma(M), \sigma \in S_n\}$. Par

exemple, dans $\mathbb{Z}[X, Y]$, $S(X^2Y) = X^2Y + XY^2$ et $S(X^2Y^2) = X^2Y^2$.

Soit donc P un polynôme symétrique défini par un monôme comme ci-dessus. Choisissons un ordre total, qu'on note $>$, sur l'ensemble $\{X_1, \dots, X_n\}$ des indéterminées, par exemple $X_1 > \dots > X_n$, et ordonnons l'ensemble des monômes de P suivant l'ordre lexicographique des n -uplets de puissances :

$$(X_1)^{k_1} \dots (X_n)^{k_n} > (X_1)^{l_1} \dots (X_n)^{l_n}$$

s'il existe $r < n$ tel que, pour $i \leq r$, $k_i = l_i$ et $k_{r+1} > l_{r+1}$.

Par exemple : $(X_1)^3(X_2)^2X_3 > (X_1)^2(X_2)^3X_3 > X_1(X_2)^2(X_3)^3$.

Pour des monômes M, M', N, N' , on voit facilement que $M \geq M'$ implique $MN \geq M'N$; on en déduit que $M \geq M'$ et $N \geq N'$ impliquent $MN \geq M'N \geq M'N'$.

Soit $M_0 = (X_1)^{k_1} \dots (X_n)^{k_n}$ le plus grand monôme de P pour l'ordre adopté. On a $k_1 \geq \dots \geq k_n$ car, sinon, la transposition échangeant deux indices i et j tels que $i < j$ et $k_i < k_j$ transformerait M_0 en un autre monôme de P plus grand que M_0 . Considérons alors un polynôme de la forme $(s_1)^{l_1} \dots (s_n)^{l_n}$. Dans ce produit, d'après la propriété précédente, le monôme le plus grand est le produit des plus grands monômes de $s_1, \dots, s_n$, élevés aux puissances $l_1, \dots, l_n$, c'est-à-dire :

$$(X_1)^{l_1} (X_1 X_2)^{l_2} \dots (X_1 \dots X_n)^{l_n} = (X_1)^{l_1 + \dots + l_n} (X_2)^{l_2 + \dots + l_n} \dots (X_n)^{l_n}.$$

Ce monôme sera égal à M_0 si $k_1 = l_1 + \dots + l_n$, $k_2 = l_2 + \dots + l_n$, etc., c'est-à-dire si :

$$\begin{aligned} l_n &= k_n; \\ l_{n-1} &= k_{n-1} - k_n; \\ &\dots; \\ l_1 &= k_1 - k_2. \end{aligned}$$

Posons $P_1 = P - (s_1)^{l_1} \dots (s_n)^{l_n}$ où les l_i ont les valeurs précédentes ; on a fait ce qu'il fallait pour que P_1 soit un polynôme symétrique nul ou dont le monôme le plus grand soit strictement inférieur à M_0 .

Si P_1 est nul, on a écrit P sous la forme voulue : $P(X_1, \dots, X_n) = T(s_1, \dots, s_n)$ où $T(X_1, \dots, X_n) = (X_1)^{l_1} \dots (X_n)^{l_n}$; sinon, P_1 est une combinaison linéaire de polynômes de la forme $S(M)$, avec M strictement inférieur à M_0 , et on recommence le même procédé pour chacun d'entre eux, jusqu'à ce qu'on obtienne le polynôme nul et l'expression de P cherchée. Voir l'exemple ci-dessous.

2) Prouvons maintenant l'unicité de U : supposons qu'il existe un polynôme non nul U tel que $U(s_1, \dots, s_n) = 0$.

Il existe un monôme unique M de U donnant le monôme le plus grand en $X_1, \dots, X_n$ de $U(s_1, \dots, s_n)$, ce qui contredit l'hypothèse $U(s_1, \dots, s_n) = 0$.

Exemple. Pour exprimer le polynôme symétrique $P(X, Y, Z) = Y^3 Z + YZ^3 + Z^3 X + ZX^3 + X^3 Y + XY^3$ en fonction des polynômes symétriques élémentaires $s_1 = X + Y + Z$, $s_2 = XY + YZ + ZX$, $s_3 = XYZ$, choisissons l'ordre $X > Y > Z$; le monôme le plus grand de P est $X^3 Y$, ce qui conduit à poser $P_1 = P - (s_1)^2 s_2$. Soit, tous calculs faits :

$$P_1 = -5(X^2 YZ + XY^2 Z + XYZ^2) - 2(X^2 Y^2 + Y^2 Z^2 + Z^2 X^2).$$

Le monôme le plus grand de P_1 est $X^2 Y^2$ qui est bien inférieur à $X^3 Y$. Comme P_1 est une combinaison linéaire de deux polynômes symétriques définis par des monômes, on peut travailler séparément sur chacun d'eux. Le premier est visible-

ment égal à $-5s_1s_3$ (ce que l'algorithme donne immédiatement) ; pour le second, on forme $X^2Y^2 + Y^2Z^2 + Z^2X^2 - (s_2)^2$, etc.

Finalement $P = (s_1)^2s_2 - s_1s_3 - 2(s_2)^2$.

Fractions rationnelles symétriques. Soient K un corps et $K(X_1, \dots, X_n)$ le corps des fractions rationnelles à n indéterminées à coefficients dans K . On dit qu'une fraction rationnelle $\frac{P}{Q}$ est symétrique si, pour tout $\sigma \in S_n$, on a $\frac{\varphi_\sigma(P)}{\varphi_\sigma(Q)} = \frac{P}{Q}$.

L'extension du théorème précédent aux fractions rationnelles symétriques est facile et fait l'objet de la proposition suivante.

Proposition. Soient A un anneau intègre, P, Q des polynômes de $A[X_1, \dots, X_n]$, $Q \neq 0$, tels que $\frac{P}{Q}$ soit une fraction rationnelle symétrique. Il existe des polynômes S et T de $A[X_1, \dots, X_n]$ tel que $\frac{S(s_1, \dots, s_n)}{T(s_1, \dots, s_n)} = \frac{P}{Q}$.

Démonstration. Si Q est symétrique, alors P est symétrique et la proposition résulte directement du théorème. Si Q n'est pas symétrique, on forme l'ensemble $E = \{\varphi_\sigma(Q), \sigma \in S_n\}$ (c'est l'orbite de Q sous l'action de S_n). Le produit des polynômes de E est un polynôme symétrique Q_1 dont un facteur est Q : $Q_1 = QQ_2$. Comme $\frac{P}{Q} = \frac{PQ_2}{Q_1}$ où le dénominateur est symétrique, on revient au premier cas. $\square$

20.10 FRACTIONS CONTINUES

On dit en allemand : *der Kettenbruch* au singulier, *die Kettenbrüche* au pluriel ; la dénomination anglaise est plus proche de la nôtre : *continued fraction*, fraction continuée.

Histoire. Christian Huygens (1629-1695) est appelé par Jean-Baptiste Colbert (1619-1683), à faire partie de l'Académie Royale des sciences à sa création en 1666. Il en est le membre le plus prestigieux. Dans les années 1670-1680, il travaille à la construction d'un automate, toujours visible au musée Boerhave de Leyde, représentant le système solaire. Les planètes de son automate se déplacent sur de petits rails circulaires autour du soleil ; leur mouvement est réglé par des engrenages. Huygens écrit : *Toute la question se réduit donc à ceci : étant donnés deux grands nombres ayant entre eux un certain rapport, en trouver d'autres plus petits pour les dents des roues qui ne soient pas incommodes par leurs grandeurs et qui aient entr'eux à peu près le même rapport, de telle façon qu'aucun couple de nombres plus petits ne fournisse un rapport plus approchant de la vraie valeur.* Huygens donne l'exemple de Saturne qui doit parcourir 12 degrés 13 minutes 34 secondes 18 tierces pendant

que la Terre parcourt 359 degrés 45 minutes 40 secondes 31 tierces. Le rapport est $\frac{12 \times 216\,000 + 13 \times 3\,600 + 34 \times 60 + 18}{359 \times 216\,000 + 45 \times 3\,600 + 40 \times 60 + 18} = \frac{2\,640\,858}{77\,708\,431}$. Le calcul qu'il fait est un développement en fraction continue (le nom est donné par Wallis en 1695) de son rapport :

$$\frac{2\,640\,858}{77\,708\,431} \approx 29 + \frac{1}{2 + \frac{1}{3}} = \frac{206}{7}.$$

Huygens conclut : *C'est pourquoi nous avons donné 206 dents à la roue de Saturne et 7 à la roue motrice*. Huygens complète son calcul par une justification de la qualité de l'approximation qu'il a obtenue et l'énonce de plusieurs propriétés que nous allons retrouver. Nous expliquerons cela plus loin.

C'est dans un article en latin : *De fractionibus continuis dissertatio*, présenté en 1737 à l'Académie de Saint-Petersbourg, qu'Euler étudie la formation des numérateurs et dénominateurs des fractions continues, et donne quelques exemples de développements comme ceux de $\sqrt{2}, \sqrt{3}, \sqrt{a^2+1}$, e , $\sqrt{e}$, $(e^2-1)/2$, etc.

Les résultats de Lagrange des années 1768-1774 fondent rigoureusement la théorie des fractions continues. Lagrange les applique pour résoudre l'équation $x^2 - dy^2 = 1$ (équation dite de Pell-Fermat, voir 19.9) et donner un résultat sur les nombres algébriques de degré 2 sur $\mathbb{Q}$ que nous évoquerons plus loin.

Les développements en fractions continues sont toujours d'actualité. Jean-Christophe Yoccoz (1957-2016) en a utilisé pour l'étude de systèmes dynamiques sur le cercle, travaux qui lui ont valu la médaille Fields en 1994.



Jean-Christophe Yoccoz
© Académie des sciences.

Fractions continues. On se place dans le corps K des fractions rationnelles à coefficients dans $\mathbb{Q}$ et à une infinité d'indéterminées X_i avec $i \in \mathbb{N}$; K est le corps des fractions de l'anneau A des polynômes à coefficients dans $\mathbb{Z}$ et à une infinité d'indéterminées X_i avec $i \in \mathbb{N}$.

Pour $n \geq 0$, les fractions rationnelles F_n suivantes sont appelées fractions continues : pour $n = 0$, $F_0 = X_0$;

$$\text{pour } n = 1, F_1 = X_0 + \frac{1}{X_1} ;$$

$$\text{pour } n = 2, F_2 = X_0 + \frac{1}{X_1 + \frac{1}{X_2}} ;$$

$$\text{en général, } F_{n+1} = X_0 + \frac{1}{X_1 + \frac{1}{\ddots X_{n-1} + \frac{1}{X_n + \frac{1}{X_{n+1}}}}} , \text{ autrement dit :}$$

$$F_{n+1}(X_0, \dots, X_{n+1}) = F_n \left(X_0, \dots, X_{n-1}, X_n + \frac{1}{X_{n+1}} \right) \text{ pour } n \geq 0. \text{ FRC 1}$$

Polynômes d'Euler. On trouve facilement, comme le fit Euler :

$$\bullet F_0 = \frac{X_0}{1} ,$$

$$\bullet F_1 = \frac{X_0 X_1 + 1}{X_1} ,$$

$$\bullet F_2 = \frac{X_0 X_1 X_2 + X_0 + X_2}{X_1 X_2 + 1} ,$$

$$\bullet F_3 = \frac{X_0 X_1 X_2 X_3 + X_0 X_3 + X_0 X_1 + X_2 X_3 + 1}{X_1 X_2 X_3 + X_1 + X_3} ,$$

Pour n plus grand, la première impression est que ça se complique ! Mais il y a moyen d'y voir plus clair.

Considérons les suites de polynômes (P_n) et (Q_n) définies par récurrence de la façon suivante :

$$P_0 = X_0, P_1 = X_0 X_1 + 1 \text{ et, pour } n \geq 2, P_n = X_n P_{n-1} + P_{n-2} ; \quad \text{FRC 2}$$

$$Q_0 = 1, Q_1 = X_1 \text{ et, pour } n \geq 2, Q_n = X_n Q_{n-1} + Q_{n-2} ; \quad \text{FRC 3}$$

Ces polynômes sont à coefficients entiers (récurrence évidente) et leurs deux suites sont très semblables : la relation de récurrence est la même et la seconde suite s'obtient à partir de la première en décalant les indices des polynômes et en changeant les X_i en X_{i+1} .

Chapitre 20 • Ouvertures sur les polynômes

Montrons que :

$$F_n = \frac{P_n}{Q_n} \text{ pour tout } n. \quad \text{FRC 4}$$

L'égalité est vraie pour $n = 0, 1$; raisonnons par récurrence et supposons qu'elle soit vraie pour un entier n . L'égalité FRC1 donne :

$$\begin{aligned} F_{n+1} &= \frac{P_n \left(X_0, \dots, X_n + \frac{1}{X_{n+1}} \right)}{Q_n \left(X_0, \dots, X_n + \frac{1}{X_{n+1}} \right)} \\ &= \frac{\left(X_n + \frac{1}{X_{n+1}} \right) P_{n-1} + P_{n-2}}{\left(X_n + \frac{1}{X_{n+1}} \right) Q_{n-1} + Q_{n-2}} \\ &= \frac{X_{n+1} (X_n P_{n-1} + P_{n-2}) + P_{n-1}}{X_{n+1} (X_n Q_{n-1} + Q_{n-2}) + Q_{n-1}} \end{aligned}$$

et il suffit d'utiliser les relations de récurrence (FRC 2) et (FRC 3) pour montrer que

$$F_{n+1} = \frac{P_{n+1}}{Q_{n+1}} \text{ et conclure.}$$

D'autres relations vont nous être utiles. En utilisant (FRC 2) et (FRC 3), on a

$$\begin{aligned} P_{n+1} Q_n - P_n Q_{n+1} &= (X_{n+1} P_n + P_{n-1}) Q_n - P_n (X_{n+1} Q_n + Q_{n-1}) \\ &= -(P_n Q_{n-1} - P_{n-1} Q_n) \text{ pour } n \geq 1. \end{aligned}$$

Comme $P_1 Q_0 - P_0 Q_1 = 1$, on en déduit, par récurrence :

$$P_{n+1} Q_n - P_n Q_{n+1} = (-1)^n, \text{ pour } n \geq 0 ; \quad \text{FRC 5}$$

$$F_{n+1} - F_n = \frac{(-1)^n}{Q_n Q_{n+1}} \text{ pour } n \geq 0. \quad \text{FRC 6}$$

Enfin, pour $n \geq 0$,

$$\begin{aligned} P_{n+2} Q_n - P_n Q_{n+2} &= (X_{n+2} P_{n+1} + P_n) Q_n - P_n (X_{n+2} Q_{n+1} + Q_n) \\ &= X_{n+2} (P_{n+1} Q_n - P_n Q_{n+1}) \end{aligned}$$

d'où

$$P_{n+2} Q_n - P_n Q_{n+2} = (-1)^n X_{n+2}.$$

FRC 7

Relation avec l'algorithme d'Euclide. Tout rationnel $\frac{a}{b}$, $b \neq 0$, admet un développement en fraction continue qui s'obtient à l'aide de l'algorithme d'Euclide pour le calcul du pgcd de deux entiers (voir 12.12.1). Rappelons que, dans le cas où $a > b > 0$, on pose $a_0 = a$, $a_1 = b$ et on définit par récurrence, pour $k \geq 1$ et tant que a_k n'est pas nul, a_{k+1} comme le reste de la division euclidienne de a_{k-1} par a_k : $a_{k-1} = a_k q_k + a_{k+1}$ avec $0 \leq a_{k+1} < a_k$ (où q_k est la partie entière de $\frac{a_{k-1}}{a_k}$). On note N l'indice du premier reste nul ; on a vu que $\text{pgcd}(a, b) = a_{N-1}$, dernier reste non nul.

On vérifie par récurrence sur k que $\frac{a_0}{a_1} = F_k \left(q_1, \dots, q_k, q_{k+1} + \frac{a_{k+2}}{a_{k+1}} \right)$, d'où $\frac{a_0}{a_1} = F_{N-2}(q_1, \dots, q_{N-1})$. On obtient ainsi un développement en fraction continue de $\frac{a}{b}$ avec $a, b > 0$ (attention : la même notation q_k désigne ici les quotients dans l'algorithme d'Euclide et, ailleurs dans cette section, les dénominateurs des fractions continues).

Par exemple, $\frac{21}{13} = F_5(1, 1, 1, 1, 1, 2)$.

Si $b > a$, on a $\frac{a}{b} = \frac{1}{\frac{b}{a}}$ et ce qui précède s'applique ; par exemple, $\frac{3}{7} = \frac{1}{2 + \frac{1}{3}}$.

Si $\frac{a}{b} < 0$, avec $b > 0$, on écrit $\frac{a}{b} = a_0 + \frac{a_1}{b}$ où a_0 est la partie entière de $\frac{a}{b}$ et on

applique ce qui précède à $\frac{a_1}{b}$ qui est positif. Par exemple, $-\frac{3}{7} = -1 + \frac{1}{1 + \frac{3}{4}}$. Des

expressions opposées ont donc des développements en fraction continue qui ne se ressemblent pas.

Développement en fraction continue d'un réel. Si $a_0, \dots, a_n$ sont des entiers, avec $a_1, \dots, a_n > 0$, le nombre $F_n(a_0, \dots, a_n)$ est un rationnel. On va voir qu'on peut décrire tout réel x non rationnel par un développement en fraction continue infini qui donne une suite de rationnels convergeant vers x et ayant de belles propriétés.

On note $[X]$ la partie entière d'un réel x , c'est-à-dire le plus grand entier inférieur ou égal à x .

Soient α un nombre réel, $a_0 = [\alpha]$. Si α n'est pas un entier, on peut écrire $\alpha = a_0 + r_1$, avec $0 < r_1 < 1$. En posant $r_1 = 1/\alpha_1$, on a $\alpha_1 > 1$ et $\alpha = a_0 + \frac{1}{\alpha_1} = F_1(a_0, \alpha_1)$.

Chapitre 20 • Ouvertures sur les polynômes

Si α_1 n'est pas un entier, on peut recommencer le même travail : $\alpha_1 = a_1 + r_2 = a_1 + 1/\alpha_2$, avec $a_1 = [\alpha_1]$, $0 < r_2 < 1$, $\alpha_2 > 1$. Tant que α_k n'est pas entier, on peut continuer à poser $\alpha_k = a_k + r_{k+1} = a_k + \frac{1}{\alpha_{k+1}}$ avec $a_k = [\alpha_k]$, $0 < r_{k+1} < 1$ et $\alpha_{k+1} > 1$. On vérifie l'égalité :

$$\alpha = F_n(a_0, \dots, a_{n-1}, \alpha_n)$$

Si α est rationnel, on a vu que le procédé s'arrêtait au bout d'un certain nombre d'étapes n . On remarquera que la dernière étape donne un entier $a_n \geq 2$. Ainsi, on définit une bijection entre l'ensemble des rationnels et l'ensemble des suites finies $(a_0, \dots, a_n)$ d'entiers, avec $a_1, \dots, a_n > 0$ et $a_n \geq 2$.

Si α n'est pas rationnel, le procédé ne s'arrête pas et définit une suite $(a_n)_{n \in \mathbb{N}}$ d'entiers, strictement positifs à partir du rang 1, appelé développement en fraction continue de α . Quand Euler obtient, en 1737, des développements infinis en fraction continue de e , e^2 , etc., il en déduit l'irrationalité de ces nombres. La même démarche permet à Lambert de prouver l'irrationalité de π en 1761.

Notations. Pour ne pas avoir à écrire une suite de fractions, on note souvent $[a_0, \dots, a_n]$ la fraction continue $F_n(a_0, \dots, a_n)$, $[a_0, \dots, a_n, \dots]$ si le développement est infini, $[a_0, \dots, a_{r-1}, \overline{a_r, \dots, a_{r+s-1}}]$ si le développement est périodique à partir du rang r et de période s : $a_{n+s} = a_n$ si $n \geq r$.

Exemples. Si on pose $\alpha = \frac{1+\sqrt{5}}{2}$, on a $\alpha = 1 + \frac{1}{\alpha}$; on en déduit :

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\ddots}}}$$

Le développement de $\sqrt{2}$ s'obtient en écrivant

$$\sqrt{2} = 1 + (\sqrt{2} - 1) = 1 + \frac{1}{2 + (\sqrt{2} - 1)},$$

ce qui permet de voir que $\alpha = \sqrt{2} - 1$ vérifie $\alpha = \frac{1}{2 + \alpha}$ et donne

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\ddots}}} = [1, \overline{2}].$$

Inversement, si on se propose de connaître le nombre irrationnel $[\overline{1, 3}]$, on écrit que $\alpha = 1 + \frac{1}{3 + \frac{1}{\alpha}}$, d'où $3\alpha^2 - 3\alpha - 1 = 0$; comme $\alpha > 1$, on trouve $\alpha = \frac{3 + \sqrt{21}}{6}$.

Plus généralement, tout nombre dont le développement est périodique à partir d'un certain rang est racine d'une équation du second degré à coefficients entiers. Lagrange montre en 1774 dans une addition aux *Éléments d'algèbre* d'Euler que la réciproque est vraie : *Tout nombre racine d'une équation du second degré à coefficients entiers admet un développement en fraction continue périodique à partir d'un certain rang.*

Pour un nombre comme π , une calculatrice donne le début du développement :

$$\pi = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac{1}{292 + \frac{1}{\ddots}}}}} = [3, 7, 15, 1, 292, \dots].$$

On ne constate dans ce début de développement aucune régularité.

Approximation d'un réel non rationnel. Supposons que α soit un réel non rationnel et étudions les propriétés du développement de α en fraction continue.

On appelle $f_n = F_n(a_0, \dots, a_n)$ la réduite (*convergent* en anglais) d'indice n de α . Cette réduite se présente comme le quotient des entiers $p_n = P_n(a_0, \dots, a_n)$ et $q_n = Q_n(a_1, \dots, a_n)$. La définition des polynômes Q_n permet de voir que $q_n \geq n$ (on peut donner de bien meilleures minoration, par exemple $q_n \geq F_n$, terme de rang n de la suite de Fibonacci, voir 2.5).

Les relations de récurrence démontrées ci-dessus donnent :

$$f_{n+1} = F_n \left(a_0, \dots, a_{n-1}, a_n + \frac{1}{a_{n+1}} \right) \text{ pour } n \geq 0 ; \quad \text{frc 1}$$

$$p_0 = a_0, p_1 = a_0 a_1 + 1 \text{ et, pour } n \geq 2, p_n = a_n p_{n-1} + p_{n-2} ; \quad \text{frc 2}$$

$$q_0 = 1, q_1 = a_1 \text{ et, pour } n \geq 2, q_n = a_n q_{n-1} + q_{n-2} ; \quad \text{frc 3}$$

$$f_n = \frac{p_n}{q_n} \text{ pour tout } n ; \quad \text{frc 4}$$

$$p_{n+1} q_n - p_n q_{n+1} = (-1)^n, \text{ pour } n \geq 0 ; \quad \text{frc 5}$$

$$f_{n+1} - f_n = \frac{(-1)^n}{q_n q_{n+1}} \text{ pour } n \geq 0 ; \quad \text{frc 6}$$

$$p_{n+2} q_n - p_n q_{n+2} = (-1)^n a_{n+2}. \quad \text{frc 7}$$

L'égalité (frc 7) montre que $f_{n+2} - f_n = (-1)^n \frac{a_{n+2}}{q_n q_{n+2}}$ a le signe de $(-1)^n$; par conséquent, la suite (f_{2n}) est croissante et la suite (f_{2n+1}) est décroissante. L'égalité (frc 5) montre que $|f_{n+1} - f_n| = \frac{1}{q_n q_{n+1}}$, donc $|f_{n+1} - f_n| < \frac{1}{n^2}$.

On en conclut que la suite des réduites de rang pair et la suite des réduites de rang impair sont adjacentes, la première croissante, la seconde décroissante, toutes les deux convergeant, comme on va l'expliquer, vers α :

$$f_0 < f_2 < \dots < \alpha < \dots < f_3 < f_1$$

$$\lim_{n \rightarrow \infty} f_n = \alpha$$

Notons que l'égalité (frc 5) montre que p_n et q_n sont des entiers premiers entre eux.

On a :

$$\alpha = F_{n+1}(a_0, \dots, a_n, \alpha_{n+1}) = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}}.$$

D'où :

$$\alpha - f_n = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}} - \frac{p_n}{q_n},$$

soit, d'après frc 5 :

$$\alpha - f_n = \frac{(-1)^n}{q_n (\alpha_{n+1} q_n + q_{n-1})}$$

Comme $\alpha_{n+1} > a_{n+1} \geq 0$, frc 3 donne

$$|\alpha - f_n| = \frac{1}{q_n q_{n+1}}$$

ce qui montre la convergence de la suite (f_n) vers α .

Après avoir montré la convergence des réduites vers le réel α , nous allons étudier les propriétés de cette convergence, propriétés qui font le grand intérêt des fractions continues.

Proposition 1 : meilleures approximations rationnelles d'un réel α non rationnel. Pour tout $n \geq 1$, tout p entier et tout q tel que $0 < q < q_{n+1}$, on a :

$$|\alpha q - p| \geq |\alpha q_n - p_n|.$$

On a l'égalité si et seulement si $p = p_n$ et $q = q_n$.

Démonstration. Il existe des entiers u et v tels que :

$$\begin{cases} p = up_n + vp_{n+1} \\ q = uq_n + vq_{n+1} \end{cases}$$

En effet, le déterminant de ce système est $(-1)^{n+1}$ d'après (frc 5).

On a $u \neq 0$, sinon $q = vq_{n+1}$ et q ne vérifie pas l'inégalité de l'énoncé. Si $v = 0$, on a $p = up_n$, $q = uq_n$, d'où l'inégalité dans ce cas ; l'égalité a lieu si $u = 1$.

Si u et v sont non nuls, on a $0 < q = uq_n + vq_{n+1} < q_{n+1}$, donc $-vq_{n+1} < uq_n < (1 - v)q_{n+1}$. Ces inégalités montrent que si $v < 0$, alors $u > 0$ et si $v \geq 1$, alors $u < 0$. On a $\alpha q - p = \alpha(uq_n + vq_{n+1}) - (up_n + vp_{n+1}) = u(\alpha q_n - p_n) + v(\alpha q_{n+1} - p_{n+1})$. Comme u et v sont de signes contraires ainsi que $\alpha q_n - p_n$ et $\alpha q_{n+1} - p_{n+1}$, on a $|\alpha q - p| = |u(\alpha q_n - p_n)| + |v(\alpha q_{n+1} - p_{n+1})|$. Le second terme n'étant pas nul, l'inégalité est stricte ; on peut alors conclure.

Corollaire. Pour tout $n \geq 1$, tout p entier et tout q tel que $0 < q < q_n$, on a :

$$\left| \alpha - \frac{p}{q} \right| > \left| \alpha - \frac{p_n}{q_n} \right|.$$

Ainsi, une réduite donne-t-elle une meilleure approximation rationnelle de α que tout rationnel dont le dénominateur est strictement inférieur à α . On comprend mieux maintenant pourquoi Huygens avait choisi un développement en fraction continue de ses rapports pour mieux approcher les mouvements réels.

Réduites de π . Les premières réduites de π sont 3, $[3, 7] = \frac{22}{7}$, $[3, 7, 15] = \frac{333}{106}$, $[3, 7, 15, 1] = \frac{355}{113}$. La fraction $\frac{22}{7}$ est l'approximation de π donnée par Archimède 7 (– 287 à – 212) dans son texte sur la mesure du cercle où il calcule des périmètres de polygones réguliers inscrits ou circonscrits à un cercle ; on a $\frac{22}{7} - \pi = 0,0012\dots$

La fraction $\frac{355}{113}$ est l'approximation de π trouvée, en appliquant la méthode d'Archimède, par Adriaan Mélius et publiée par son fils (qui portait le même nom) en 1625 ; on a $\frac{355}{113} - \pi = 0,00000026\dots$. Devant la qualité de cette approximation, il n'y a pas lieu de retenir la fraction $\frac{333}{106}$ qui donne seulement : $\frac{333}{106} - \pi = -0,00008321\dots$

Proposition 2 : reconnaissance de réduite. Soit α un irrationnel positif.

1) Si $\frac{p_n}{q_n}$ et $\frac{p_{n+1}}{q_{n+1}}$ sont deux réduites successives de α , on a :

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2q_n^2} \text{ ou } \left| \alpha - \frac{p_{n+1}}{q_{n+1}} \right| < \frac{1}{2q_{n+1}^2}$$

2) Si p et q sont des entiers strictement positifs tels que $\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q^2}$, alors $\frac{p}{q}$ est une réduite de α .

Démonstration.

1) Supposons que $\left| \alpha - \frac{p_n}{q_n} \right| \geq \frac{1}{2q_n^2}$ et que $\left| \alpha - \frac{p_{n+1}}{q_{n+1}} \right| \geq \frac{1}{2q_{n+1}^2}$

D'une part, frc 5 montre que

$$q_{n+1}(q_n\alpha - p_n) - q_n(q_{n+1}\alpha - p_{n+1}) = p_{n+1}q_n - p_nq_{n+1} = (-1)^n.$$

D'autre part, comme $(q_n\alpha - p_n)$ et $q_{n+1}\alpha - p_{n+1}$ sont de signes opposés :

$$|q_{n+1}(q_n\alpha - p_n) - q_n(q_{n+1}\alpha - p_{n+1})| \geq \frac{q_{n+1}}{2q_n} + \frac{q_n}{2q_{n+1}}.$$

Le second membre de l'inégalité est de la forme $\frac{1}{2}(a + \frac{1}{a})$ avec $a \neq 1$; il est donc strictement supérieur à 1, ce qui donne une contradiction.

2) Il existe un entier n tel que $q_n \leq q < q_{n+1}$. Comme :

$$\left| \frac{p}{q} - \frac{p_n}{q_n} \right| \leq \left| \alpha - \frac{p}{q} \right| + \left| \alpha - \frac{p_n}{q_n} \right|$$

(par l'inégalité triangulaire), la proposition 1 donne :

$$\left| \frac{p}{q} - \frac{p_n}{q_n} \right| \leq \left(\frac{1}{q} + \frac{1}{q_n} \right) |q\alpha - p|, \text{ soit } |q_n p - p_n q| \leq (q + q_n) |q\alpha - p|,$$

donc $|q_n p - p_n q| < 2q \times \frac{1}{2q} = 1$; d'où $|q_n p - p_n q| = 0$ et le résultat.

Solutions des équations de Pell-Fermat. Soit $d \geq 2$ un entier sans facteur carré. Si x et y sont des entiers vérifiant l'équation $x^2 - dy^2 = 1$, on a $\left| \frac{x}{y} - \sqrt{d} \right| < \frac{1}{2y^2}$. En effet, l'équation donne $\frac{x^2}{y^2} - d = \frac{1}{y^2}$, donc $\left(\frac{x}{y} - \sqrt{d} \right) \left(\frac{x}{y} + \sqrt{d} \right) = \frac{1}{y^2}$ et il suffit de remarquer que $\frac{x}{y} > \sqrt{d}$.

Le 2) de la proposition 2 montre alors que $\frac{x}{y}$ est une réduite du développement en fraction continue de $\sqrt{d}$. Il reste à déterminer laquelle, ce que nous n'expliquerons

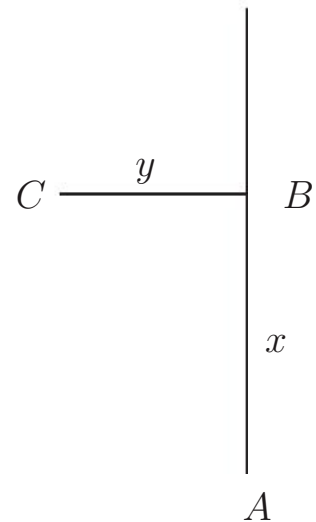
pas. Mentionnons cependant que le résultat est lié à la période du développement de $\sqrt{d}$ et que les calculs peuvent être fort longs et les solutions exprimées avec des entiers très grands : pour $d < 1000$, le record est pour $d = 991$, où il faut calculer 59 termes du développement en fraction continue de $\sqrt{991}$. On trouve

$$\begin{aligned}x &= 379516400906811930638014896080, \\y &= 12055735790331359447442538767.\end{aligned}$$

Des textes de François Jabœuf (IREM de Montpellier, 1985) et Marcel Couchouron (préparation à l'agrégation sur le site de Rennes) m'ont été très utiles pour rédiger cette section.

20.11 GÉOMÉTRIE ALGÈBRE

Dans sa *Géométrie*, Descartes montre comment travailler d'une façon nouvelle sur les courbes : *Je choisis une droite fixe, comme AB, pour rapporter à ses divers points tous ceux de cette ligne courbe EC, et en cette ligne je choisis un point, comme A, pour commencer par luy ce calcul.* Il peut ainsi repérer tout point d'un quart de plan par sa distance y à AB et la distance x de sa projection sur AB à A .



C'était la naissance de la géométrie cartésienne, même si Descartes ne considérait que les coordonnées positives. En même temps, Fermat avait lui aussi commencé à utiliser des axes de coordonnées. Cela ouvrait un nouveau domaine des mathématiques. L'étude des courbes algébriques, courbes définies par une équation de la forme $P(x, y) = 0$ où P est un polynôme.

La géométrie algébrique s'est considérablement développée depuis et elle est actuellement une des grandes branches des mathématiques dites pures.

Elle a subi plusieurs transformations fondamentales, la dernière étant due à Alexander Grothendieck (1928-2014), l'un des mathématiciens les plus extraordinaires du vingtième siècle. Il a entièrement renouvelé la théorie pour unifier la géométrie et l'arithmétique. La construction de la théorie des schémas qu'il commence à édifier dans les années 1955 est absolument monumentale. Elle mobilise autour de Grothendieck une pléiade de mathématiciens brillants à l'Institut des Hautes Études Scientifiques (IHÉS) de Bures-sur-Yvette. Elle atteint *toute la souplesse et la généralité désirable*, comme le dit Grothendieck, et a permis de résoudre des problèmes considérés comme extrêmement difficiles.



Alexander Grothendieck (de dos)

– Cliché CSF – René Bouillot. Le séminaire de géométrie algébrique de l’IHES à Bures-sur-Yvette au début des années 1960. Alexandre Grothendieck (de dos) devant un auditoire de mathématiciens dont cette photo montre une partie : Jean Dieudonné (ici caché par Grothendieck), Jean-Pierre Serre, Claude Chevalley, Armand Borel, Pierre Samuel, François Bruhat, Michel Demazure, Italo Giorgiutti, Jean-Pierre Jouanolou, Michèle et Michel Raynaud, Jean-Louis Verdier, Jean Giraud, Pierre Gabriel, Daniel Ferrand, etc.

Depuis les années 1970, les travaux d’Alexander Grothendieck se sont partagés entre les mathématiques, l’écologie et des textes immenses de réflexion personnels comme *Récoltes et semailles*, écrit dans une langue admirable et où Grothendieck explicite la philosophie profonde de ses recherches... et ses obsessions douloureuses. On peut les lire sur le réseau, mais je ne sais quand ils pourront être, enfin, édités. Depuis le début des années 1990, Grothendieck a vécu seul dans un petit village de l’Ariège, Lasserre, limitant au maximum ses contacts avec le monde, écrivant inlassablement. Il a laissé 100 000 pages qui attendent toujours d’être lues, 28 000 confiées en 1991 à un de ses élèves à Montpellier, le reste récupéré à son décès et entassé (en 2019) dans un sous-sol à Saint-Germain-des-Prés à Paris, la possibilité que tout cela soit accessible au public dépendant d’un accord des cinq héritiers de Grothendieck avec un éventuel acheteur (ce pourrait être la Bibliothèque nationale de France, mais elle ne peut y mettre qu’une petite somme ou une université américaine disposant de moyens financiers bien plus importants).

Exercices

20.1 Calculs de quotients de $\mathbb{Z}[X]$

À quels anneaux « plus simples » sont isomorphes les anneaux $\mathbb{Z}[X]/(X^2 + 3, p)$ pour $p = 3, 5, 7, 11, 13, 17, 19$? Donner pour les trois premiers cas le nombre d'éléments, l'expression des éléments en fonction de la classe x de X , la formule donnant le produit de deux éléments.

20.2 Polynômes irréductibles

- 1) Déterminer les polynômes irréductibles de degré ≤ 4 de $\mathbb{Z}/2\mathbb{Z}[X]$.
- 2) Montrer que $X^5 + 21X^2 - 63$ est irréductible dans $\mathbb{Z}[X]$, que $Y^5 - (1 + 2X^2 + X^4)Y + 1 + X^2$ est irréductible dans $\mathbb{R}[X, Y]$.
- 3) Décomposer $X^n - 1$ en produit de facteurs irréductibles dans $\mathbb{Z}[X]$, $\mathbb{Q}[X]$, $\mathbb{R}[X]$, $\mathbb{C}[X]$ pour $n = 3, 7, 12$.
- 4) Même question pour $X^4 + 64$.

20.3 Isomorphismes

- 1) On pose $A = \frac{\mathbb{R}[X, Y, Z]}{(X - a, Y - b, Z - c)}$, avec $a, b, c \in \mathbb{R}$. Montrer que $A \simeq \mathbb{R}$.
- 2) Soient K un corps et $a, b \in K$. À quels anneaux plus simples sont isomorphes $\frac{K[X, Y]}{(Y - aX)}$, $\frac{K[X, Y, Z]}{(Z - aX - bY)}$.
- 3) On pose $A = \frac{\mathbb{R}[X, Y]}{Y^2 - X^2 - X^3}$. Montrer que A est intègre et que son corps de fractions est isomorphe à $\mathbb{R}(T)$; on pourra utiliser l'homomorphisme $f: \mathbb{R}[X, Y] \rightarrow \mathbb{R}[T]$ défini par $f(X) = T^2 - 1$, $f(Y) = T(T^2 - 1)$.

20.4 Propriétés des anneaux de polynômes

- 1) Montrer que, si $A[X]$ est un anneau factoriel, alors A est un anneau factoriel.
- 2) Montrer que, si $A[X]$ est un anneau principal, alors A est un corps.
- 3) Soient A un anneau factoriel, $K = \text{Cf}(A)$, $P \in K[X]$ non nul. Montrer que $P = u P_1$ avec $u \in K$ et P_1 polynôme primitif de $A[X]$, u et P_1 étant uniquement déterminés à des inversibles de A près.
- 4) Soient A un anneau factoriel, $P, Q \in A[X]$, $P_1 = P/c(P)$, $Q_1 = Q/c(Q)$. On note en indice l'anneau dans lequel on cherche le pgcd.
 - a) Montrer que : $\text{pgcd}_{A[X]}(P, Q) = \text{pgcd}_A(c(P), c(Q)) \text{pgcd}_{A[X]}(P_1, Q_1)$.
 - b) Montrer que $\text{pgcd}_{A[X]}(P_1, Q_1)$ est le polynôme primitif de $A[X]$ associé à $\text{pgcd}_{\text{Cf}(A)[X]}(P_1, Q_1)$.

Ce résultat permet de déterminer le pgcd de deux polynômes primitifs en effectuant les calculs de l'algorithme d'Euclide dans $\text{Cf}(A)[X]$.

20.5 Quotients d'anneau de polynômes à deux indéterminées

On pose $A = \frac{\mathbb{R}[X, Y]}{(X^2 + Y^2 - 1)}$, $B = \frac{\mathbb{C}[X, Y]}{(X^2 + Y^2 - 1)}$. On note y la classe de Y dans A ou B

(le contexte précisera). Cet exercice est sans doute plus difficile que la moyenne ; il montre comment travailler dans un quotient.

1) Montrer que A et B sont des anneaux intègres.

2) Montrer que tout élément de A (ou de B) peut s'écrire $Sy + T$ pour S, T uniques dans $\mathbb{R}[X]$ (ou dans $\mathbb{C}[X]$) et que $\mathbb{R}[X]$ (ou $\mathbb{C}[X]$) s'identifie à un sous-anneau de A (ou de B).

3) On définit $N : A \rightarrow \mathbb{R}[X]$ par $N(Sy + T) = (Sy + T)(-Sy + T)$. Vérifier que N est compatible avec la multiplication dans A (on dit que N est une norme sur A) et que $N(Sy + T) = T^2 + S^2X^2 - S^2$.

4) Déterminer les éléments inversibles de A .

5) Montrer que $X - 1$ est irréductible dans A .

6) Montrer que A n'est pas un anneau factoriel.

7) Montrer que l'idéal $I = (X - 1, y)$ de A n'est pas principal.

8) On pose $C = \frac{\mathbb{C}[U, V]}{(UV - 1)}$. Montrer que $B \simeq C$.

9) Reprendre pour B et C les différentes questions posées pour A .

20.6 Équation du troisième degré

La résolution de l'équation du troisième degré présentée en 1545 par Jérôme Cardan (1501-1576) suit les idées suivantes. Pour résoudre $x^3 + ax^2 + bx + c = 0$, avec $a, b, c \in \mathbb{R}$ (la démarche pour une équation à coefficients dans $\mathbb{C}$ est analogue), on se ramène d'abord au cas $x^3 + px + q = 0$, avec $p, q \in \mathbb{R}$.

1) Expliquer comment.

2) Si x est une solution de $x^3 + px + q = 0$, on définit u et v par les conditions $x = u + v$ et $uv = -p/3$; on pose $U = u^3$, $V = v^3$. Montrer que U et V sont solutions d'une équation du second degré (calculer la somme et le produit de U et de V) et terminer la résolution de $x^3 + px + q = 0$.

20.7 Décomposition en éléments simples

1) Décomposer en éléments simples

$$\frac{1}{(X-1)(X+1)}, \frac{1}{(X-a)(X-b)}, \frac{2X^2 - 5X + 3}{(X^2+1)(X^2+X+1)}$$

sur $\mathbb{R}[X]$ et $\mathbb{C}[X]$.

2) Soit P un polynôme scindé de $\mathbb{R}[X]$ ou de $\mathbb{C}[X]$, $P = \prod_{1 \leq i \leq r} (X - \alpha_i)^{k_i}$. Décomposer $\frac{P'}{P}$ en éléments simples.

20.8 Suites de Sturm

1) Déterminer, en appliquant l'algorithme de Sturm, le nombre de racines réelles de $P = X^2 - 2,8X + a$ pour $a = 1,95$, $a = 1,96$ et $a = 1,97$.

2) Déterminer, en appliquant l'algorithme de Sturm, le nombre de racines réelles de $P = X^4 + 4X^3 + 1$.

3) Soient U et V des polynômes de $\mathbb{R}[X]$ premiers entre eux et tels que $\deg(V) < \deg(U)$; on suppose de plus U sans racines multiples. On définit la suite des restes $R_0 = U'V$, $R_1 = U$, ..., R_s constante non nulle, dans le calcul du pgcd de par divisions successives. On change les signes des polynômes de cette suite comme dans la démonstration du théorème de Sturm (voir 20.7). On note $(P_0, \dots, P_s)$ la suite obtenue; pour tout x réel, on pose $u_x = (P_0(x), \dots, P_s(x))$.

Soient $a < b$ deux réels non racines de U . Montrer que $v(u_a) - v(u_b)$ donne la différence entre le nombre de racines α de U dans $[a, b]$ telles que $V(\alpha) > 0$ et le nombre de racines α de U dans $[a, b]$ telles que $V(\alpha) < 0$.

20.9 Théorème de Fermat dans $\mathbb{C}[X]$

Soient $n \geq 3$ un entier. Le but de l'exercice est de montrer qu'il n'existe pas de polynômes A, B, C de $\mathbb{C}[X]$, non constants et premiers entre eux deux à deux, tels que :

$$A^n + B^n = C^n.$$

On pose $\zeta = e^{2\pi i/n}$ et on note $\mu_n = \{\zeta^k, k = 0, \dots, n-1\}$ l'ensemble des racines n -ièmes de l'unité dans $\mathbb{C}$.

Supposons l'existence de tels polynômes et choisissons-en un, parmi tous les triplets (A, B, C) vérifiant $A^n + B^n = C^n$, tel que le maximum des degrés des trois polynômes soit le plus petit possible.

1) Montrer que $A^n = \prod_{0 \leq k \leq n-1} (C - \zeta^k B)$ et que deux facteurs distincts de ce produit sont premiers entre eux.

2) En déduire que, pour $k = 0, \dots, n-1$, $C - \zeta^k B = F_k^n$ où F_k est un polynôme de $\mathbb{C}[X]$ de degré strictement inférieur aux degrés de A, B, C .

3) Montrer que F_0^n, F_1^n, F_2^n sont linéairement dépendants.

4) En déduire une contradiction.

Commentaire. La méthode de descente infinie de Fermat a encore donné la clef de la démonstration. Curieusement, le théorème de Fermat pour les entiers a résisté 350

ans aux efforts des mathématiciens alors que son analogue dans $\mathbb{C}[X]$ est un résultat abordable.

20.10 Théorème de Mason-Stothers

Ce théorème est présenté dans l'*Algèbre* de Lang (page 201) comme étant une relation entièrement nouvelle découverte indépendamment par Wilson Stothers (en 1981) et R. C. Mason (en 1984). La démonstration élémentaire suivante a été publiée par Noah Snyder en 2000.

On se place dans l'anneau des polynômes à coefficients complexes. Si P est un polynôme, on note $v(P)$ le nombre de racines distinctes de P et on pose $D(P) = \text{pgcd}(P, P')$.

Théorème : Soient A, B, C trois polynômes premiers entre eux deux à deux tels que $A + B + C = 0$. On a l'inégalité

$$\deg(A) \leq v(ABC) - 1 = v(A) + v(B) + v(C) - 1$$

et des inégalités analogues pour $\deg(B), \deg(C)$.

- 1) Montrer que $\deg(D(P)) = \deg(P) - v(P)$ pour tout $P \in \mathbb{C}[X]$.
- 2) Montrer que $v(ABC) = v(A) + v(B) + v(C)$.
- 3) Vérifier que $A'B - AB' = AC' - A'C$.
- 4) Montrer que $D(A)D(B)D(C)$ divise $A'B - AB'$.
- 5) En déduire le théorème.
- 6) Utiliser ce résultat pour donner une autre démonstration du théorème de Fermat pour les polynômes à coefficients complexes (exercice 20.9).

20.11 Version universelle du théorème de Cayley-Hamilton

Soient n un entier, $M = (X_{i,j})$ la matrice carrée d'ordre n en les n^2 indéterminées $X_{i,j}, 1 \leq i, j \leq n$, $\chi(X) = \det(M - XI) = \sum_{0 \leq k \leq n} A_k X^k$ le polynôme caractéristique de M , considéré comme élément de l'anneau $A[X]$ de polynômes à coefficients dans l'anneau $A = \mathbb{Z}[(X_{i,j})_{1 \leq i,j \leq n}]$; les coefficients A_k sont donc des polynômes de A . On note K le corps des fractions de A ($K = \mathbb{Q}((X_{i,j})_{1 \leq i,j \leq n})$) et L une extension de K dans laquelle $\chi(X)$ se factorise en produit de facteurs du premier degré en X .

- 1) Montrer que M a n valeurs propres distinctes dans L .
- 2) Montrer le théorème de Cayley-Hamilton pour les matrices diagonales.
- 3) Montrer le théorème de Cayley-Hamilton pour la matrice M .
- 4) En déduire le théorème de Cayley-Hamilton pour les matrices à coefficients dans un anneau B .

20.12 Formules de Newton

Soit n un entier. Pour tout entier $d \geq 1$, posons $p_d = \sum_{1 \leq i \leq n} (X_i)^d$ dans $\mathbb{Z}[X_1, \dots, X_n]$.

Nous allons montrer que, pour tout entier $d \geq 1$:

$$p_d = \sum_{1 \leq k \leq d-1} (-1)^{k-1} s_k p_{d-k} + (-1)^{d-1} d s_d.$$

1) Vérifier directement les formules pour $n = 3$ et $d = 1, 2, 3$.

2) Montrer les formules si $d \geq n$. On pourra partir de la formule de 20.9.1 et faire successivement $X = X_1, \dots, X = X_n$.

3) On suppose désormais $d < n$ et on pose $P = \prod_{1 \leq i \leq n} (X - X_i)$.

a) Montrer que $P' = \frac{P}{X - X_1} + \dots + \frac{P}{X - X_n}$.

b) Montrer que $\frac{P}{X - X_i} = \sum_{0 \leq k \leq n-1} \left(\sum_{0 \leq l \leq k} (-1)^l s_l X_i^{k-l} \right) X^{n-k-1}$.

c) En déduire les formules de Newton.



Sir Isaac Newton (1643-1727)

Gravure extraite de *History of the World*, Helmut H.F., Dodd, Mead & Company, 1902.

Commentaires. Les formules de Newton sont universelles : si un polynôme $P = \sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k}$ d'un anneau de polynômes $A[X]$ admet pour racines

$x_1, \dots, x_n$ dans A , il existe, d'après la propriété universelle 20.8, un homomorphisme $f: \mathbb{Z}[X_1, \dots, X_n] \rightarrow A$ tel que $f(X_i) = x_i$ pour $i = 1, \dots, n$ et les formules de Newton entre les puissances des x_i et les valeurs des polynômes symétriques élémentaires en les x_i sont valables.

Newton énonce ces formules vers 1666 et les publie sans démonstration en 1707. Elles permettent, de proche en proche, le calcul des polynômes p_d sommes de puissances d -ièmes des X_i . Il faut se rappeler que $s_k = 0$ pour $k > n$, ce qui tronque la formule si $d > n$. La première démonstration apparaît dans le Traité d'algèbre de Colin Mac Laurin (1698-1746), publié deux ans après sa mort. On ne sait s'il la tenait de Newton, avec qui il avait été en relation avant la mort de celui-ci, en 1727.

En 1629, Girard avait déjà donné l'expression de p_d en fonction des s_k pour $d \leq 4$: il appelait joliment *premier meslé*, *second meslé*, etc., les $s_1, s_2 \dots$. C'est sans doute la première fois que des racines, éventuellement *impossibles*, étaient placées dans des formules comme si c'étaient des nombres.

20.13 Fractions continues

1) Déterminer les développements en fraction continue de $29/12$, $\sqrt{34}$ (on pourra utiliser deux méthodes : un calcul algébrique et une calculatrice).

2) Quel nombre a pour développement en fraction continue $[1, 2, 3]$?

3) En – 46, Jules César, imposa le calendrier dit julien : trois années de 365 jours et une année avec un jour supplémentaire : le sixième jour avant les calendes de mars était répété : *bis sexto ante calendas martii*, d'où le nom de *bissextile* (le premier jour de l'année était alors le premier mars). Cela aurait pu être parfait, mais l'année ne fait pas 365,25 jours, l'année est égale à 365 jours 5 heures 48 minutes 46 secondes = 365,2422 jours à quelques poussières de temps près. La différence est de 0,0078 jour. Au bout de 1500 ans, la différence est de 12 jours et le début des saisons se décale. Le pape Grégoire XIII a laissé son nom au calendrier grégorien décidé par une bulle du 24 février 1582 : les années bissextiles des fins de siècle sont supprimées, sauf celles divisibles par 400 comme l'an 2000, ce qui laisse une différence de 26 secondes par an, et le lendemain du jeudi 4 octobre 1582 fut le vendredi 15 octobre 1582. Mais attention, la suppression n'a pas eu d'effet universel. La France, par exemple, a suivi le pape en supprimant les 10, ..., 19 décembre 1582, mais les Anglais ont attendu septembre 1752 pour se recalculer ; les orthodoxes n'ont rien fait et c'est pourquoi on parle de la Révolution d'octobre pour des événements de novembre 1917. Ces décalages font que Cervantès est mort le 23 avril 1616 (calendrier grégorien), Shakespeare à la même date du calendrier julien anglais, autrement dit 10 jours plus tard. Newton n'est pas né l'année de la mort de Galilée : Galilée est mort le 8 janvier 1642, Newton est né le 4 janvier 1643, mais le calendrier julien anglais le fait naître le 25 décembre 1642. D'autre part Newton est mort le 31 mars 1727 (calendrier grégorien), mais le calendrier julien anglais donne la date du 20 mars

1726, car le changement d'année avait alors lieu le 25 mars en Angleterre. Il faut donc être très soigneux quand on compare des dates de cette époque.

Enfin une question mathématique : qu'aurait-on pu conseiller à Grégoire XIII à l'aide des fractions continues ?

4) Un laboratoire médical annonce que son médicament est efficace dans 53,84 % des cas. Est-il sérieux ?

Solutions

20.1 Dans tous les cas, p est premier, $\mathbb{Z}/p\mathbb{Z}$ est un corps et $\mathbb{Z}[X]/(X^2 + 3, p) \simeq \mathbb{Z}[X]/(p\mathbb{Z}[X]/(X^2 + 3))$. Nous notons x la classe de X dans le quotient.

Si $p = 3$, $\mathbb{Z}[X]/(X^2 + 3, p) \simeq (\mathbb{Z}/3\mathbb{Z})[X]/(X^2)$, anneau ayant 9 éléments : $a + bx$ avec $a, b \in \mathbb{Z}/3\mathbb{Z}$, la multiplication étant définie par $(a + bx)(a' + b'x) = aa' + (ab' + a'b)x$ puisque $x^2 = 0$ (dans un anneau, un élément dont une puissance est nulle est appelé nilpotent).

Si $p = 5$, $\mathbb{Z}[X]/(X^2 + 3, p) \simeq (\mathbb{Z}/5\mathbb{Z})[X]/(X^2 + 3)$. Comme -3 n'est pas un carré modulo 5, $X^2 + 3$ est irréductible dans $(\mathbb{Z}/5\mathbb{Z})[X]$; par conséquent, $(\mathbb{Z}/5\mathbb{Z})[X]/(X^2 + 3)$ est un corps, extension de degré 2 de $\mathbb{Z}/5\mathbb{Z}$; ce corps a 25 éléments : $a + bx$ avec $a, b \in \mathbb{Z}/5\mathbb{Z}$, la multiplication étant définie par $(a + bx)(a' + b'x) = aa' - 3bb' + (ab' + a'b)x$ puisque $x^2 = -3$.

Si $p = 7$, $\mathbb{Z}[X]/(X^2 + 3, 7) \simeq (\mathbb{Z}/7\mathbb{Z})[X]/(X^2 + 3)$. Comme $-3 = 4$ est un carré modulo 7, on a $X^2 + 3 = (X - 2)(X + 2)$; les deux facteurs $X - 2$ et $X + 2$ sont étrangers dans $(\mathbb{Z}/7\mathbb{Z})[X]$ et on peut appliquer le théorème chinois. Par conséquent, $\mathbb{Z}[X]/(X^2 + 3, 7) \simeq (\mathbb{Z}/7\mathbb{Z}) \times (\mathbb{Z}/7\mathbb{Z})$, anneau produit ayant 49 éléments (a, b) avec $a, b \in \mathbb{Z}/7\mathbb{Z}$, la multiplication se faisant composante par composante ; cet anneau n'est pas intègre puisque, par exemple $(1, 0)(0, 1) = 0$, mais il ne contient pas d'éléments nilpotents.

Si $p = 11$ ou 17 , -3 n'est pas un carré et ces cas sont analogues à celui de $p = 5$. Si $p = 13$, $-3 = 6^2$, si $p = 19$, $-3 = 4^2$ et ces cas sont analogues à celui de $p = 7$. La loi de réciprocité quadratique (voir 21.3.2) permet de savoir facilement si -3 est ou non un carré modulo p .

20.2 1) Degré 1 : X et $X + 1$ sont irréductibles.

Degrés 2 et 3 : Il faut chercher les polynômes sans racines dans $\mathbb{Z}/2\mathbb{Z}[X]$, c'est-à-dire ne s'annulant pas en 0 et 1. On trouve $X^2 + X + 1$, $X^3 + X + 1$, $X^3 + X^2 + 1$.

Degré 4 : Les polynômes de degré 4 sans racines dans $\mathbb{Z}/2\mathbb{Z}[X]$ ont 3 ou 5 termes ; ce sont $X^4 + X^3 + 1$, $X^4 + X^2 + 1$, $X^4 + X + 1$, $X^4 + X^3 + X^2 + X + 1$. Comme le seul polynôme irréductible de $\mathbb{Z}/2\mathbb{Z}[X]$ de degré 2 est $X^2 + X + 1$, $X^4 + X^2 + 1 = (X^2 + X + 1)^2$ doit être éliminé de la liste précédente.

2) Appliquer le critère d'Eisenstein dans $\mathbb{Z}[X]$ avec $p = 7$ (mais non avec $p = 3$) dans le premier cas, dans $\mathbb{R}[X][Y] = \mathbb{R}[X, Y]$ avec $P = 1 + X^2$ polynôme irréductible de $\mathbb{R}[X]$ dans le second.

3) $X^3 - 1 = (X - 1)(X^2 + X + 1)$ dans $\mathbb{Z}[X]$, $\mathbb{Q}[X]$ et $\mathbb{R}[X]$;

$X^3 - 1 = (X - 1)(X - j)(X - j^2)$ dans $\mathbb{C}[X]$.

$X^7 - 1 = (X - 1)(X^6 + X^5 + X^4 + X^3 + X^2 + X + 1)$ dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$ d'après l'application du critère d'Eisenstein donnée en 20.6 ; en posant $\zeta = e^{2i\pi/7}$,

$$X^7 - 1 = (X - 1)(X - \zeta)(X - \zeta^2)(X - \zeta^3)(X - \zeta^4)(X - \zeta^5)(X - \zeta^6)$$

dans $\mathbb{C}[X]$; en regroupant les facteurs ayant des racines conjuguées, on trouve la décomposition dans $\mathbb{R}[X]$:

$$X^7 - 1 = (X - 1)(X^2 - 2X \cos(2\pi/7) + 1)(X^2 - 2X \cos(4\pi/7) + 1)$$

$$(X^2 - 2X \cos(6\pi/7) + 1).$$

Dans $\mathbb{C}[X]$, on a $X^{12} - 1 = \prod_{0 \leq k \leq 11} (X - \zeta^k)$ avec $\zeta = e^{i\pi/6}$; en regroupant les facteurs ayant des racines conjuguées, on trouve la décomposition dans $\mathbb{R}[X]$:

$$X^{12} - 1 = (X - 1)(X + 1)(X^2 + 1)(X^2 - X\sqrt{3} + 1)(X^2 - X + 1)$$

$$(X^2 + X\sqrt{3} + 1)(X^2 + X + 1) ;$$

enfin, dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$:

$$X^{12} - 1 = (X - 1)(X + 1)(X^2 + 1)(X^4 + X^2 + 1)(X^2 + X + 1)(X^2 + X + 1)$$

4) On a $X^4 + 64 = (X^2 + 4X + 8)(X^2 - 4X + 8)$ dans $\mathbb{Z}[X]$, $\mathbb{Q}[X]$, $\mathbb{R}[X]$ et, en posant $\alpha = 2\sqrt{2}, \zeta = \frac{1+i}{\sqrt{2}} = e^{i\pi/4}$, on a dans $\mathbb{C}[X]$:

$$X^4 + 64 = (X - \alpha\zeta)(X - \alpha\zeta^3)(X - \alpha\zeta^5)(X - \alpha\zeta^7).$$

20.3 1) Notons $\pi : \mathbb{R}[X, Y, Z] \rightarrow A$ la projection canonique. L'homomorphisme $f : \mathbb{R}[X, Y, Z] \rightarrow \mathbb{R}$ défini (grâce à la propriété universelle) par $f(X) = a, f(Y) = b, f(Z) = c$ s'annule sur l'idéal $(X - a, Y - b, Z - c)$. Il existe donc un homomorphisme $f' : A \rightarrow \mathbb{R}$ tel que $f = f' \circ \pi$. Comme f est surjectif, f' l'est aussi. Un polynôme $P \in \ker(f)$ vérifie $P(a, b, c) = 0$. Le polynôme P_1 défini par $P_1(X, Y, Z) = P(X - a, Y - b, Z - c)$ vérifie $P_1(0, 0, 0) = 0$. Donc P_1 n'a pas de terme constant et $P_1 \in (X, Y, Z)$; d'où $P \in (X - a, Y - b, Z - c)$. On en déduit $\ker(f) = (X - a, Y - b, Z - c)$; f' est un isomorphisme.

2) Notons $\pi : K[X, Y] \rightarrow \frac{K[X, Y]}{(Y - aX)}$ la projection canonique. L'homomorphisme $f : K[X, Y] \rightarrow K[X]$ défini (grâce à la propriété universelle) par $f(X) = X, f(Y) = aX$

s'annule sur l'idéal $(Y - aX)$. Il existe donc un homomorphisme $f' : \frac{K[X, Y]}{(Y - aX)} \rightarrow K[X]$

tel que $f = f' \circ \pi$. Comme f est surjectif, f' l'est aussi. Soit $P \in \ker(f)$; la

division euclidienne de P par $Y - aX$ dans $K[X][Y]$ s'écrit $P = (Y - aX)Q(X, Y) + R(X)$. Comme $f(P) = 0$, on a $R = f(R) = 0$. Par conséquent, f' est un isomorphisme.

On montre de même que $\frac{K[X, Y, Z]}{(Z - aX - bY)} \simeq K[X, Y]$.

3) Le polynôme $A = Y^2 - X^2 - X^3$ est irréductible dans $\mathbb{R}[X, Y]$. S'il ne l'était pas, on aurait $A = (Y + U)(Y + V)$ avec U, V polynômes en X . On aurait $U + V = 0$ et $UV = -X^2(X + 1)$, ce qui est impossible. On pouvait aussi appliquer le critère d'Eisenstein dans $\mathbb{R}[X][Y]$ avec l'anneau factoriel $\mathbb{R}[X]$ et l'élément irréductible $X - 1$.

Notons $\pi : \mathbb{R}[X, Y] \rightarrow A$ la projection canonique. L'homomorphisme $f : \mathbb{R}[X, Y] \rightarrow \mathbb{R}[T]$ défini par $f(X) = T^2 - 1$, $f(Y) = T(T^2 - 1)$ s'annule sur l'idéal $(Y^2 - X^2 - X^3)$ (vérification facile). Il existe donc un homomorphisme $g : A \rightarrow \mathbb{R}[T]$ tel que $f = g \circ \pi$.

Soit $P \in \ker(f)$; la division euclidienne de P par $Y^2 - X^2 - X^3$ dans $\mathbb{R}[X][Y]$ s'écrit

$$P = (Y^2 - X^2 - X^3)Q(X, Y) + YR(X) + S(X).$$

Comme $f(P) = 0$, on a : $T(T^2 - 1)R(T^2 - 1) + S(T^2 - 1) = 0$; comme $S(T^2 - 1)$ ne peut avoir que des termes de degré pair et que $T(T^2 - 1)R(T^2 - 1)$ ne peut avoir que des termes de degré impair, on a $R = S = 0$, ce qui montre que g est injectif.

On en déduit un homomorphisme $h : \text{Cf}(A) \rightarrow \mathbb{R}(T)$ rendant commutatif le diagramme ci-après.

Comme un homomorphisme de corps est injectif, il reste à montrer que h est surjectif, ce qui résulte de $h\left(\frac{\pi(Y)}{\pi(X)}\right) = T$.

$$\begin{array}{ccc} \mathbb{R}[X, Y] & & \\ \pi \downarrow & \searrow f & \\ A & \xrightarrow{g} & \mathbb{R}[T] \\ \text{incl} \downarrow & & \downarrow \text{incl} \\ \text{Cf}(A) & \xrightarrow{h} & \mathbb{R}(T) \end{array}$$

20.4 1) Appliquer la définition.

2) Si A n'est pas un corps, il existe dans A un élément non inversible. Notons-le a . On vérifie que l'idéal $I = (a, X)$ n'est pas principal.

3) Soit m le produit (ou le ppcm) des dénominateurs des coefficients de P . Le polynôme $Q = mP$ est dans $A[X]$ et s'écrit $Q = c(Q)Q_1$ avec Q_1 polynôme primitif de $A[X]$. D'où $P = \frac{c(Q)}{m}Q_1$ qui est de la forme demandée. Il reste à montrer la condition

d'unicité. Si $P = \frac{a_1}{b_1} P_1 = \frac{a_2}{b_2} P_2$ avec $a_1, b_1, a_2, b_2 \in A$ et P_1, P_2 primitifs, on a $b_1 b_2 P =$

$a_1 b_2 P_1 = a_2 b_1 P_2$, donc il existe une unité v de A telle que $a_1 b_2 = v a_2 b_1$ et $P_2 = v P_1$, d'où le résultat.

4) a) La décomposition de P et de Q en produit de facteurs irréductibles de $A[X]$ peut se scinder en deux produits avec les facteurs irréductibles de A d'une part, de $A[X]$ de l'autre. Choisissons un système de représentants S des premiers et un système de représentants S' des seconds. On a $P = u \prod_{p \in S} p^{v_p} \prod_{\pi \in S'} \pi^{v_\pi(P)}$, avec u inversible, et une formule analogue pour Q . Le pgcd de P et Q est obtenu en faisant le produit des $p^{\inf(v_p(P), v_p(Q))}$ par le produit des $\pi^{\inf(v_\pi(P), v_\pi(Q))}$ ce qui donne l'égalité demandée.

b) Cela résulte de ce que les polynômes irréductibles non constants de $A[X]$ sont irréductibles dans $K[X]$.

20.5 1) Il suffit de montrer que le polynôme $X^2 + Y^2 - 1$ est irréductible dans $\mathbb{C}[X, Y]$; il le sera alors dans $\mathbb{R}[X, Y]$. Si $X^2 + Y^2 - 1 = P(X, Y)Q(X, Y)$ dans $\mathbb{C}[X, Y]$ avec P et Q non inversibles, on voit d'abord que la seule possibilité est que P et Q soient de degré 1 en X et en Y ; si

$$(aX + bY + c)(a'X + b'Y + c') = X^2 + Y^2 - 1,$$

alors $aa' = 1$ et on peut choisir $a = a' = 1$; on doit alors avoir $b + b' = 0$ et $bb' = 1$, d'où $b = \iota = -b'$, par exemple ; il reste à satisfaire $c + c' = c - c' = 0$ et $cc' = -1$, ce qui n'est pas possible. On pouvait aussi appliquer le critère d'Eisenstein dans $\mathbb{R}[Y][X]$ avec l'anneau factoriel $\mathbb{R}[Y]$ et l'élément irréductible $Y - 1$.

2) Un élément de A est la classe d'un polynôme P de $\mathbb{R}[X, Y]$; en divisant P par $Y^2 + X^2 - 1$ dans $\mathbb{R}[X][Y]$, on obtient la forme demandée. Montrons que l'application linéaire $(S, T) \mapsto Sy + T$ est injective (elle permettra les identifications indiquées). Si $Sy + T = 0$, il existe $W \in \mathbb{R}[X, Y]$ tel que $Sy + T = W(X^2 + Y^2 - 1)$; le degré en Y du membre de droite ne peut être ≤ 1 que si $W = 0$, ce qui impose $S = 0$ et $T = 0$. On peut penser à y comme étant $\sqrt{1 - X^2}$.

3) Si on appelle $-Sy + T$ le conjugué de $Sy + T$, on vérifie que le conjugué de $(Sy + T)(Uy + V)$ est le produit des conjugués $(-Sy + T)(-Uy + V)$. On en déduit

$$N((Sy + T)(Uy + V)) = N(Sy + T)N(Uy + V).$$

Enfin, $N(Sy + T) = T^2 - S^2 y^2 = T^2 + S^2 X^2 - S^2$.

4) Si $Sy + T$ est inversible dans A , sa norme $T^2 + S^2 X^2 - S^2$ est inversible dans $\mathbb{R}[X]$; donc est égale à une constante. En regardant le terme de plus haut degré du membre de gauche, dans cette égalité entre polynômes en X , on voit qu'elle ne peut avoir lieu que si $S = 0$, ce qui entraîne que T est une constante non nulle (l'argument ne marche pas si les polynômes sont à coefficients complexes, les termes de plus haut

degré de $X^2 S^2$ et de T^2 pouvant s'annuler). Les éléments inversibles sont les réels non nuls dans A .

5) Si $X - 1 = (Sy + T)(Uy + V)$ dans A , on a, en prenant la norme comme cidessus,

$$(X - 1)^2 = (T^2 + S^2 X^2 - S^2)(V^2 + U^2 X^2 - U^2) ;$$

des considérations de degré montrent que S ou U est nul ; si $S = 0$ et si T n'est pas une constante, alors $U = 0$, V est une constante non nulle et $T = X - 1$ à un facteur réel non nul près ; de même si $U = 0$. On en déduit l'irréductibilité de $X - 1$ dans A .

6) Comme $X - 1$ est irréductible dans A , $-X - 1$ l'est également, car l'homomorphisme $A \rightarrow A$ défini par $X \mapsto -X$ et $y \mapsto y$ conserve l'irréductibilité. La méthode du **5)** permet de montrer que y est irréductible dans A . Si $X - 1$ et y étaient associés, il existerait une unité non nulle $a \in A$ telle que $y = a(X - 1)$; le **2)** montre que c'est impossible.

Alors l'égalité $y^2 = -(X - 1)(1 + X)$ (conséquence immédiate de $X^2 + y^2 - 1 = 0$) montre que A n'est pas factoriel car il n'y a pas de factorisation unique en produit de facteurs irréductibles dans A .

7) Si I est principal, il existe un élément $d \in A$ engendrant I et des éléments a et b de A tels que $X - 1 = ad$ et $y = bd$. Comme $X - 1$ et y sont irréductibles dans A , soit d est inversible, soit d n'est pas inversible et alors a et b le sont. Si d est inversible, $I = A$, donc il existe $f, g \in A$ tels que $1 = (X - 1)f + yg$, soit

$$1 = (X - 1)F(X, Y) + Y G(X, Y) + W(X, Y)(X^2 + Y^2 - 1)$$

avec $F, G, W \in \mathbb{R}[X, Y]$; cette dernière égalité donne $1 = 0$ pour $X = 1$ et $Y = 0$, contradiction. Si a et b sont inversibles, on aurait $X - 1 = ab^{-1}y$, alors qu'on a montré que $X - 1$ et y ne sont pas associés dans A .

8) Notons $p : \mathbb{C}[U, V] \rightarrow C$ et $q : \mathbb{C}[X, Y] \rightarrow B$ les projections canoniques. On définit l'homomorphisme $f : \mathbb{C}[U, V] \rightarrow \mathbb{C}[X, Y]$ par $f(U) = X + iY$ et $f(V) = X - iY$; comme $f(UV - 1) = X^2 + Y^2 - 1$, $q \circ f$ se factorise par p et il existe f' tel que $q \circ f = f' \circ p$. Pour montrer que f' est un isomorphisme, on construit g' telle que $g' \circ q = p \circ g$ en factorisant $g : \mathbb{C}[X, Y] \rightarrow \mathbb{C}[U, V]$ défini par $g(X) = \frac{U+V}{2}$, $g(Y) = \frac{U-V}{2i}$. Le diagramme suivant donne à voir ces constructions.

$$\begin{array}{ccc} \mathbb{C}[X, Y] & \begin{array}{c} \xrightarrow{g} \\ \xleftarrow{f} \end{array} & \mathbb{C}[U, V] \\ \downarrow q & & \downarrow p \\ B & \begin{array}{c} \xrightarrow{g'} \\ \xleftarrow{f'} \end{array} & C \end{array}$$

On vérifie que $g' \circ f' \circ p = g' \circ q \circ f = p \circ g \circ f = p$ et que $f' \circ g' \circ q = f' \circ p \circ g = q \circ f \circ g = q$ pour montrer que g' est l'inverse de f' .

L'isomorphisme f' permet de voir simplement dans C des propriétés moins évidentes dans B .

9) Posons $u = p(U)$, $v = p(V)$; on a $uv = 1$ donc $u^{-1} = v$. Un élément de C s'écrit de manière unique $S(u) + T(v)$, avec $S, T \in \mathbb{C}[X]$, T sans terme constant. En effet, l'image par p d'un polynôme de $\mathbb{C}[U, V]$ est de cette forme et il y a unicité car si $S(u) + T(v) = S_1(u) + T_1(v)$, on a

$$S(U) + T(V) = S_1(U) + T_1(V) + W(U, V)(UV - 1)$$

dans $\mathbb{C}[U, V]$ et W ne peut être que nul.

On peut écrire tout élément non nul de C de manière unique sous la forme $u^n S(u)$ avec $S \in \mathbb{C}[X]$, $S(0)$ non nul et $n \in \mathbb{Z}$.

Les éléments inversibles de C sont de la forme au^n avec $a \in \mathbb{C}$ non nul et $n \in \mathbb{Z}$.

On peut montrer que C est un anneau factoriel dont les éléments irréductibles sont de la forme $u^n S(u)$ avec S irréductible dans $\mathbb{C}[X]$.

Les propriétés de C que nous venons de donner se traduisent en propriétés de B à l'aide de l'isomorphisme f' ; par exemple, comme C est un anneau factoriel, B est également factoriel.

On a $g'(X - 1) = (u + v - 2)/2 = (u + \frac{1}{u} - 2)/2 = (u - 1)^2/2u$, ce qui prouve que $X - 1$ n'est pas irréductible dans B et que

$$X - 1 = (X + iy - 1)^2/2(X + iy)$$

où $X + iy$ est inversible dans B et $X + iy - 1$ est irréductible dans B . De même, $g'(y) = (u - v)/2i = (u - 1)(u + 1)/2iu$ montre que y n'est pas irréductible dans B .

L'idéal $j = (X - 1, y)$ de B correspond à l'idéal

$$j_1 = ((u + v - 2)/2, (u - v)/2i) = ((u - 1)^2, (u - 1)(u + 1))$$

de C ; on a $j_1 \subset (u - 1)$ et, comme

$$u - 1 = (u + v - 2)/2 + (u - v)/2,$$

$j_1 = (u - 1)$; j_1 est donc principal et $j = f'(j_1) = (X + iy - 1)$ également.

20.6 1) On change x en $x - a/3$.

2) On a successivement :

$$(u + v)^3 + p(u + v) + q = 0 ;$$

$$u^3 + v^3 + (u + v)(3uv + p) + q = 0 ;$$

$$\begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -p^3 / 27 \end{cases}$$

$$\begin{cases} rU + V = -q \\ UV = -p^3 / 27 \end{cases}$$

Puisqu'on connaît leur somme et leur produit, U et V sont solutions de l'équation $X^2 + qX - p^3/27 = 0$ dont le discriminant est égal à $\Delta = q^2 + 4p^3/27$. Soit d un nombre complexe tel que $d^2 = \Delta$. On a, par exemple :

$$\begin{cases} U = \frac{-q+d}{2} \\ V = \frac{-q-d}{2} \end{cases}$$

Si u est solution de $u^3 = U$, il en est de même de ju et de j^2u . Les valeurs de v correspondantes sont $v = -\frac{p}{3u}$, j^2v et jv . Les trois solutions de $x^3 + px + q = 0$ sont donc :

$$\begin{cases} x = u + v \\ x = ju + j^2v \\ x = j^2u + jv \end{cases}$$

La formule donnée par Cardan (à ceci près que Cardan ne connaît pas l'usage des lettres) s'écrit :

$$x = \sqrt[3]{\frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

Si on avait inversé les choix de U et V , une racine cubique de $\frac{-q-d}{2}$ est un des trois nombres v, jv, j^2v précédents, ce qui redonne les trois mêmes racines.

Cardan n'avait rien dit de ces subtilités et sa formule est ambiguë pour le choix des racines cubiques (comment aurait-il pu imaginer de pareilles choses ?). À la fin de son livre, Cardan est le premier à présenter quelques calculs avec les nombres complexes ou plutôt avec des racines de nombres négatifs. Pour l'histoire à rebondissements des équations du troisième degré, voir [ESC1].

$$\text{20.7 1)} \frac{1}{(X-1)(X+1)} = \frac{1}{2} \left(\frac{1}{X-1} - \frac{1}{X+1} \right).$$

$$\frac{1}{(X-a)(X+b)} = \frac{1}{a-b} \left(\frac{1}{X-a} - \frac{1}{X+b} \right).$$

$\frac{2X^2 - 5X + 3}{(X^2 + 1)(X^2 + X + 1)} = \frac{aX + b}{X^2 + 1} + \frac{cX + d}{X^2 + X + 1}$; en multipliant par $X^2 + 1$ et en évaluant en i , on trouve $a = -1$ et $b = -5$; en multipliant par $X^2 + X + 1$ et en évaluant en j , on trouve $c = 1$ et $d = 8$.

La décomposition de $\frac{2X^2 - 5X + 3}{(X^2 + 1)(X^2 + X + 1)}$ dans $\mathbb{C}(X)$ s'obtient en écrivant

$$\frac{2X^2 - 5X + 3}{(X^2 + 1)(X^2 + X + 1)} = \frac{a}{X - i} + \frac{b}{X + i} + \frac{c}{X - j} + \frac{d}{X - j^2}$$

En multipliant par $X - i$ et en évaluant en i , on trouve $a = (-1 + 5i)/2$, d'où $b = (-1 - 5i)/2$; en regroupant les deux termes obtenus, on retrouve $\frac{-X - 5}{X^2 + 1}$. De même, on trouve $c = -2 - 5j$, $d = -2 - 5j^2$, etc.

2) Il vaut mieux faire un calcul direct : P' est la somme des termes ou un seul facteur est dérivé, qui sont de la forme $k_1(X - a_1)^{k_1-1}(X - a_2)^{k_2} \dots (X - a_r)^{k_r}$, etc. D'où $\frac{P'}{P} = \frac{k_1}{X - a_1} + \dots + \frac{k_r}{X - a_r}$.

20.8 1) La suite de Sturm est $R_0 = X^2 - 2,8X + a$, $R_1 = 2X - 2,8$, $R_2 = 1,96 - a$. On voit que P a une racine double pour $a = 1,96$ et que $P = (X - 1,4)^2$ dans ce cas (on pourrait adapter le résultat de Sturm au cas où P a des racines réelles multiples). Pour $a \neq 1,96$, le tableau de signes en $\pm\infty$ est le suivant.

$a = 1,95$	$-\infty$	$+\infty$	$a = 1,97$	$-\infty$	$+\infty$
$R_0 = P$	+	+	$R_0 = P$	+	+
$R_1 = P'$	-	+	$R_1 = P'$	-	+
R_2	+	+	R_2	-	-

Si $a = 1,95$, le nombre de changements de signes en $-\infty$ est 2 et 0 en $+\infty$; le polynôme a deux racines réelles. Si $a = 1,97$, le nombre de changements de signes est 1 en $+\infty$ comme en $-\infty$ et il n'y a pas de racines réelles ; ces résultats sont évidents si on remarque que $X^2 - 2,8X + a = (X - 1,4)^2 + a - 1,96$.

2) En divisant P' par 4 pour faciliter les calculs, on trouve la suite $R_0 = X^4 + 4X^3 + 1$, $R_1 = X^3 + 3X^2$, $R_2 = 3X^2 - 1$, $R_3 = -X - 3$, $R_4 = -26$.

Le tableau de signes est le suivant.

	$-\infty$	-4	-3	-1	0	$+\infty$
$R_0 = P$	+	+	-	-	+	+
$R_1 = P'$	-	-	0	+	0	+
R_2	+	+	+	+	-	+
R_3	+	+	0	-	-	-
R_4	-	-	-	-	-	-

Le nombre de changements de signes en $-\infty$ est 3 et 1 en $+\infty$. Le nombre de racines réelles est donc 2 ; en prenant des valeurs particulières, on voit avec les suites de Sturm (mais le simple théorème des valeurs intermédiaires suffirait) qu'elles sont toutes deux négatives, l'une dans $[-4, -3]$, l'autre dans $[-1, 0]$.

3) La fonction $v(u_x)$ est constante en tout point de $[a, b]$ qui n'est pas une racine de U ; l'argument est le même que dans la démonstration du théorème de Sturm. Il reste à examiner ce qui se passe quand x passe d'une valeur inférieure à une valeur supérieure à une racine α de U dans $[a, b]$; quatre cas se présentent suivant que U' et V sont positifs ou négatifs en α . Quand $V(\alpha) > 0$, on a les deux cas suivants.

	α		
U	−	0	+
U'	+	+	+
V	+	+	+
$P_0 = U' V$	+	+	+
$P_1 = U$	−	0	+

	α		
U	+	0	−
U'	−	−	−
V	+	+	+
$P_0 = U' V$	−	−	−
$P_1 = U$	+	0	−

Dans ces deux cas, on voit que $v(u_x)$ diminue de 1 quand x passe par α en croissant. L'étude quand $V(\alpha) < 0$ est analogue ; on voit que $v(u_x)$ augmente de 1 quand x passe par α en croissant. En regroupant les résultats pour toutes les racines de U situées dans l'intervalle $[a, b]$, on obtient le résultat annoncé.

On peut remarquer qu'en prenant $V = 1$, on retrouve le théorème de Sturm donné en 20.7.

20.9 1) L'égalité résulte de l'identité $Y^n - Z^n = \prod_{1 \leq k \leq n-1} (Y - \zeta^k Z)$. Si un polynôme D divise $C - \zeta^k B$ et $C - \zeta^{k'} B$ pour $k \neq k'$, il divise leur différence, donc B ; par conséquent, il divise C , contradiction.

2) La factorialité de $\mathbb{C}[X]$ montre que, pour tout $k = 0, \dots, n-1$, $C - \zeta^k B$ est une puissance n -ième de polynôme à un inversible près, inversible qu'on peut lui-même, dans $\mathbb{C}$, écrire comme une puissance n -ième, d'où le résultat.

3) Les trois polynômes appartiennent à l'espace vectoriel à deux dimensions engendré par B et C ; ils sont donc linéairement dépendants.

4) La relation de dépendance linéaire du **3)** s'écrit $xF_0^n + yF_1^n + zF_2^n = 0$. Comme $x, y, -z$ sont, dans $\mathbb{C}$, des puissances n -ièmes, l'égalité précédente s'écrit $A_1^n + B_1^n = C_1^n$ avec le maximum des degrés de A_1, B_1, C_1 strictement inférieur à celui de A, B, C , contrairement à l'hypothèse de départ.

20.10 1) Si $P = \prod_{1 \leq i \leq r} (X - a_i)^{k_i}$ avec les a_i distincts deux à deux, alors $\text{pgcd}(P, P') = \prod_{1 \leq i \leq r} (X - a_i)^{k_i - 1}$ à un facteur près, d'où la relation.

2) Les racines de A, B, C forment des ensembles disjoints car les trois polynômes sont premiers entre eux.

3) Le calcul est facile

4) Les polynômes $D(A), D(B), D(C)$ sont premiers entre eux deux à deux ; les deux premiers divisent $A'B - AB'$, le troisième divise

$$AC' - A'C = A'B - AB'.$$

5) Le **4)** donne : $\deg(D(A)) + \deg(D(B)) + \deg(D(C)) \leq \deg(A'B - AB') \leq \deg(A) + \deg(B) - 1$. En exprimant $\deg(D(A)), \deg(D(B)), \deg(D(C))$ à l'aide du **1)**, on obtient le résultat pour $\deg(C)$, puis le résultat général puisque A, B, C jouent des rôles symétriques.

6) Notons que $v(P^n) = v(P) \leq \deg(P)$ pour tout polynôme P . Si $A^n + B^n = C^n$, le théorème de Mason-Stothers donne $\deg(A^n) = n\deg(A) \leq \deg(A) + \deg(B) + \deg(C) - 1$ et des inégalités analogues pour $\deg(B)$ et $\deg(C)$. D'où :

$$n(\deg(A) + \deg(B) + \deg(C)) \leq 3(\deg(A) + \deg(B) + \deg(C)) - 3$$

et une contradiction pour $n \geq 3$.

20.11 1) Le polynôme $\chi(X)$ se factorise dans $L[X]$ en produit de facteurs du premier degré : $\chi(X) = (-1)^n \prod_{1 \leq i \leq n} (X - x_i)$. Pour montrer que les racines x_i sont distinctes, il suffit de le montrer dans le cas particulier où les indéterminées $X_{i,j}, 1 \leq i, j \leq n$ prennent les valeurs $X_{i,i} = i, X_{i,j} = 0$ pour $1 \leq i, j \leq n$ et $j \neq i$, c'est-à-dire quand M est la matrice diagonale définie par $1, 2, \dots, n$.

2) Si D est une matrice diagonale et que $d_1, \dots, d_n$ sont les coefficients diagonaux, on a

$$\chi(X) = (-1)^n \prod_{1 \leq i \leq n} (X - d_i), \text{ donc } \chi(D) = (-1)^n \prod_{1 \leq i \leq n} (D - d_i I) ;$$

pour $i = 1, \dots, n$, la matrice $D - d_i I$ a un 0 comme i -ième coefficient de la diagonale ; le produit de ces matrices est nul.

3) C'est une conséquence immédiate des **1)** et **2)**, puisque le polynôme caractéristique ne change pas par changement de base. On a donc $\sum_{0 \leq k \leq n} A_k M^k = 0$.

4) Le cas du **3)** est universel. Soit $m = (x_{i,j})$ une matrice carrée d'ordre n à coefficients dans B . D'après la propriété universelle de l'algèbre A , il existe un homomorphisme $f: A \rightarrow B$ tel que $f(X_{i,j}) = x_{i,j}$ pour $1 \leq i, j \leq n$. Si on pose $f(A_k) = a_k$, les a_k sont les coefficients du polynôme caractéristique de m et $\sum_{0 \leq k \leq n} a_k m^k = f\left(\sum_{0 \leq k \leq n} A_k M^k\right) = 0$.

20.12 1) Les formules $p_1 = s_1$, $p_2 = s_1 p_1 - 2s_2 = s_1^2 - 2s_2$ sont évidentes. Pour $d = n = 3$, $p_3 = s_1 p_2 - s_2 p_1 + 3s_3$ correspond à l'égalité $X^3 + Y^3 + Z^3 = (X + Y + Z)(X^2 + Y^2 + Z^2) - (XY + YZ + ZX)(X + Y + Z) + 3XYZ$.

2) Pour $i = 1, \dots, n$, la formule $\prod_{1 \leq j \leq n} (X - X_j) = \prod_{0 \leq k \leq n} (-1)^k s_k X^{n-k}$ qui est une formule dans l'anneau $\mathbb{Z}[(X_i)_{1 \leq i \leq n}][X]$ donne, en évaluant en $X = X_i$, $0 = X_i^n + \sum_{1 \leq k \leq n} (-1)^k s_k X_i^{n-k}$, d'où $X_i^n = -\sum_{1 \leq k \leq n} (-1)^{k-1} s_k X_i^{n-k}$; en faisant la somme de ces n égalités multipliées par X_i^{d-n} , on obtient p_d dans le membre de gauche et ce qu'il faut dans le membre de droite (en distinguant les cas $d = n$ où apparaît le terme $(-1)^{d-1} ds_d$ et le cas $d < n$ où il n'apparaît pas ; ne pas oublier que $s_d = 0$ pour $d > n$).

3) a) Voir exercice 20.7 2).

b) En effectuant la division euclidienne de P par les $X - X_i$, on obtient

$$\frac{P}{X - X_i} = X^{n-1} + (-s_1 + X_i)X^{n-2} + (s_2 - s_1 X_i + X_i^2)X^{n-3} + \dots,$$

ce qui donne le résultat.

c) On a :

$$P' = \sum_{0 \leq k \leq n-1} (-1)^k (n-k) s_k X^{n-k-1} = \frac{P}{X - X_1} + \dots + \frac{P}{X - X_n}.$$

Il reste à faire la somme des expressions des $\frac{P}{X - X_i}$ pour $i = 1, \dots, n$ trouvées dans

le **b)**, ce qui fait apparaître des p_{k-1} et des $(-1)^k n s_k$ dans le coefficient de X^{n-k-1} , et à comparer les coefficients des puissances de X dans les deux expressions de P' pour arriver aux formules de Newton.

20.13 1) On trouve $29/12 = [2, 2, 2, 2]$ et $\sqrt{34} = [5, \overline{1, 4, 1, 10}]$. Pour ce dernier calcul, une calculatrice est très rapide : on calcule $\sqrt{34}$, on soustrait la partie entière, on inverse le résultat et on continue jusqu'à une répétition. Le calcul algébrique est plus délicat :

$$\begin{aligned}\sqrt{34} &= 5 + (\sqrt{34} - 5) = 5 + \frac{1}{(\sqrt{34} + 5)/9}; \\ (\sqrt{34} + 5)/9 &= 1 + (\sqrt{34} - 4)/9 = 1 + \frac{1}{(\sqrt{34} + 4)/2}; \\ (\sqrt{34} + 4)/2 &= 4 + (\sqrt{34} - 4)/2 = 4 + \frac{1}{(\sqrt{34} + 4)/9}, \text{ etc.}\end{aligned}$$

2) Posons $\alpha = \overline{[1, 2, 3]}$. On a $\alpha = 1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{\alpha}}}$, d'où $\alpha = \frac{4 + \sqrt{37}}{7}$.

3) Le début du développement en fraction continue de 365,2422 est [365, 4, 7, 1, 3]. Il donne les approximations $365 + 1/4$, $365 + 7/29$, $365 + 8/33$, $365 + 31/128$. Au lieu de concevoir un système peu précis avec les multiples de 100, on aurait pu supprimer une année bissextile tous les 128 ans, ce qui donne un bien meilleur résultat :

$$365 + \frac{31}{128} = 365,24219 \text{ au lieu de } 365 + \frac{97}{400} = 365,2425.$$

4) Comme 0,5384 a pour développement en fraction continue $[0, 1, 1, 6\dots]$, proche de celui de $7/13$, il est possible que ce pourcentage masque une étude sur 13 cas seulement, 7 ayant retiré un bénéfice du traitement. Il est cependant difficile d'évaluer le sérieux de l'étude sans en savoir plus. Les pourcentages sont bien commodes pour des comparaisons, mais s'ils portent sur des nombres faibles, la publication de ces nombres semble relever d'un minimum de rigueur. L'exemple de l'exercice a été inventé.

La notion de corps pose un petit problème de terminologie. Dans ce livre, le mot corps désigne une structure d'anneau où tout élément non nul a un inverse et où la loi multiplicative est commutative. Joseph Wedderburn (1882-1948) et Leonard Dickson (1874-1954) ont montré en 1905 (c'est ce qu'on appelle le théorème de Wedderburn, mais c'est Dickson qui en a donné une démonstration correcte) que l'hypothèse : *la loi multiplicative est commutative* est automatiquement vérifiée pour les structures d'anneau fini où tout élément non nul a un inverse et où la loi de multiplication n'est pas supposée *a priori* commutative, ce qu'on trouve souvent énoncé sous la forme : *Tout corps fini est commutatif*. Quand la loi multiplicative n'est pas commutative, on parle de corps non commutatif, corps gauche ou anneau à division ; l'exemple classique est celui des quaternions, voir 22.10.

Les corps finis ont des propriétés particulières que n'ont pas les corps de caractéristique 0.

L'étude des corps finis commence avec Évariste Galois.

21.1 CORPS FINIS : GÉNÉRALITÉS

Exemples de corps finis

1) $\mathbb{Z}/p\mathbb{Z}$ muni de l'addition et de la multiplication est un corps fini pour tout p premier (voir 12.13).

2) Si K est un corps fini ayant q éléments et si P est un polynôme irréductible de degré n de $K[X]$, l'anneau quotient $K[X]/(P)$ est un corps (voir 20.2), extension de K de degré n qui, en tant que K -espace vectoriel, admet pour base l'ensemble des classes des X^k , $0 \leq k \leq n-1$; il est donc fini et a q^n éléments.

Par exemple, si $K = \mathbb{Z}/2\mathbb{Z}$ et si $P(X) = X^2 + X + 1$, P n'ayant pas de racines dans K est irréductible sur K et $K[X]/(P)$ est un corps à quatre éléments : les classes de 0, 1, X , $X + 1$. Notons-les 0, 1, x , $x + 1$; la multiplication dans K donne :

$$x^2 = -x - 1 = x + 1 ;$$

$$x^3 = x(x + 1) = (x + 1)x = x^2 + x = 1 ;$$

$$x^4 = (x + 1)(x + 1) = x^2 + 1 = x, \text{ etc.}$$

On note souvent $\mathbb{F}_q$ ($\mathbb{F}$ pour *field, corps* en anglais) un corps fini à q éléments, la notation se justifiant par le théorème d'unicité, à isomorphisme près, de tels corps (voir 21.2, théorème 2)

Caractéristique d'un corps fini. On a vu en 19.2 que la caractéristique d'un corps fini K est un nombre premier p et que K contient un sous-corps isomorphe à $\mathbb{F}_p$.

Proposition : Propriétés d'un corps fini. Soit K un corps fini de caractéristique $p > 0$ ayant q éléments.

a) K est une extension de $\mathbb{F}_p$ et un $\mathbb{F}_p$ -espace vectoriel de dimension finie. Si $r \geq 1$ est la dimension de cet espace, on a $q = p^r$.

b) Le groupe additif de K est isomorphe au groupe $(\mathbb{Z}/p\mathbb{Z}, +)^r$.

c) Tout élément x de K^* vérifie

$$x^{q-1} = 1$$

et tout élément x de K vérifie

$$x^q = x.$$

d) L'application $F_p : x \rightarrow x^p$ est un automorphisme de K appelé homomorphisme (on peut dire plus précisément : automorphisme) de Frobenius de K . En particulier, pour tous $x, y \in K$, on a

$$(x + y)^p = x^p + y^p$$

et, plus généralement, pour tout entier $k \geq 1$:

$$(x + y)^{p^k} = x^{p^k} + y^{p^k}$$

e) Le groupe multiplicatif K^* est un groupe cyclique d'ordre $q - 1 = p^r - 1$.

Démonstration. a) Comme K est de caractéristique p , K contient un sous-corps K' isomorphe à $\mathbb{F}_p$. On peut alors munir K d'une structure de $\mathbb{F}_p$ -espace vectoriel ; cet espace est de dimension finie (voir 20.2, proposition). En notant r cette dimension, le cardinal de K est p^r .

b) Cela résulte de la structure d'espace vectoriel précédente.

c) On sait que dans un groupe fini l'ordre d'un élément divise l'ordre du groupe. Comme K^* est d'ordre $q - 1$ on a, pour tout x de K^* , $x^{q-1} = 1$. Pour tout x de K^* on a donc $x^q = x$, propriété encore vraie pour $x = 0$. On retrouve le petit théorème de Fermat comme cas particulier.

d) Les formules ont été montrées en 19.2.

e) Démonstration 1. Suivons la démonstration originale de Gauss, (pages 53-55 des *Recherches arithmétiques*) qui est le premier à dégager cette notion de cyclicité.

Soit a un élément d'ordre maximum s de K^* .

Si $s = q - 1$, a engendre K^* qui est donc cyclique.

Supposons $s < q - 1$. L'ensemble $E = \{a^k, 0 \leq k \leq s - 1\}$ a s éléments distincts, tous racines de $X^s - 1$; comme un polynôme de degré s sur un corps a au plus s racines (voir 13.5), aucun élément b de $K^* \setminus E$ n'est racine de $X^s - 1$; l'ordre t de b ne divise donc pas s et $\text{ppcm}(s, t) > s$.

Écrivons les décompositions de s et t comme produits de facteurs premiers p_i , $1 \leq i \leq r$, élevés à des puissances positives ou nulles : $s = \prod_{1 \leq i \leq r} p_i^{k_i}$ et $t = \prod_{1 \leq i \leq r} p_i^{l_i}$ et supposons que les p_i aient été rangés de telle façon que $k_i < l_i$ pour $1 \leq i \leq j$ et $k_i \geq l_i$ dans le cas contraire (certains k_i et l_i peuvent être nuls). Posons $u = \prod_{1 \leq i \leq j} p_i^{k_i}$, $u' = s/u$, $v = \prod_{1 \leq i \leq j} p_i^{l_i}$, $v' = t/v$. Comme u' et v n'ont pas de facteur premier commun, ils sont premiers entre eux. On voit que a^u est d'ordre u' et que $b^{v'}$ est d'ordre v . L'élément $a^u b^{v'}$ est d'ordre $u'v = \text{ppcm}(s, t) > s$, ce qui donne une contradiction. Donc $s = q - 1$ et a engendre K^* .

f) Démonstration 2. ($D(n)$ désigne l'ensemble des diviseurs de n et φ l'indicateur d'Euler). Les éléments d'ordre d de K^* sont solutions de l'équation $x^d = 1$. Cette équation a au maximum d solutions dans K . S'il existe un élément a d'ordre d , les solutions de $x^d = 1$ sont les d éléments a^k avec $0 \leq k < d$. Parmi ces solutions, les éléments d'ordre d sont les a^k où k est premier avec d ; il y en a $\varphi(d)$. Le nombre $v(d)$ d'éléments d'ordre d est donc 0 ou $\varphi(d)$. Comme tout élément de K^* a un ordre divisant $q - 1$, on a $q - 1 = \sum_{d \in D(q-1)} v(d)$; en comparant avec l'égalité

$$q - 1 = \sum_{d \in D(q-1)} \varphi(d) \quad (\text{voir exercice 21.1 8}), \text{ on voit que } v(d) = \varphi(d) \text{ pour tout divi-}$$

seur d de $q - 1$; en particulier, $v(q - 1) = \varphi(q - 1)$ n'est pas nul, ce qui prouve que K^* contient au moins un élément d'ordre $q - 1$. $\square$

Commentaires sur le d). L'homomorphisme de Frobenius permet de reconnaître une puissance p -ième d'un polynôme P de $\mathbb{F}_p[X]$. Si $P(X) = \sum_{0 \leq k \leq n} a_k X^k$, alors

$P(X)^p = \sum_{0 \leq k \leq n} a_k X^{kp}$, autrement dit $(P(X))^p = P(X^p)$: les seules termes non nuls d'une puissance p -ième sont de degré multiple de p . La réciproque est immédiate.

Commentaires sur le e). Dans le groupe $\left(\frac{\mathbb{Z}}{17\mathbb{Z}}\right)^*$, 2 n'est pas un générateur car $2^4 = -1$ et $2^8 = 1$; on vérifie que 3 est générateur (voir 22.9) ; c'est la clé de la construction à la règle et au compas du polygone régulier à 17 côtés par Gauss (voir l'exercice IX.7 de mon livre de Théorie de Galois).

Il n'existe malheureusement pas d'algorithme pour calculer rapidement un générateur de K^* , même dans le cas où $K = \mathbb{F}_p$. Pour de petites valeurs de p , on peut calculer l'ordre des éléments, en commençant par celui de 2, puis éventuellement celui de 3, etc., en s'arrêtant au premier élément d'ordre $p - 1$. Pour des compléments sur ce sujet, voir 24.4.

21.2 EXISTENCE ET UNICITÉ DES CORPS FINIS

On vient de voir que le cardinal d'un corps fini est toujours une puissance d'un nombre premier. Deux résultats importants sont maintenant à démontrer, les théorèmes 1 et 2 suivants.

Théorème 1. *Pour tout entier $n > 0$ et tout nombre premier p , il existe effectivement un corps ayant $q = p^n$ éléments.*

Démonstration. Montrons que le corps de décomposition L du polynôme $P = X^q - X$ sur $\mathbb{F}_p$ répond à la question (la construction de ce corps a été donnée en 20.2). Pour cela, notons E l'ensemble des racines de P dans L . Comme le polynôme dérivé de P est égal à -1 , les racines de P sont simples, donc $\text{card}(E) = q$. On va montrer que L coïncide avec E . Si a et b sont deux racines de P , on a $(a+b)^q = (a+b)^{p^n} = a^{p^n} + b^{p^n} = a^q + b^q = a + b$, ce qui prouve que $a + b \in E$ et, d'autre part, $(ab)^q = a^q b^q = ab$, ce qui prouve que $ab \in E$; enfin, il est clair que $0, 1 \in E$ et que $a^{-1} \in E$. Par conséquent, E est un corps, donc $E = L$.

Corollaire. *Pour tout entier $n > 0$ et tout nombre premier p , il existe un polynôme irréductible de degré n de $\mathbb{F}_p[X]$.*

Démonstration. Le théorème montre l'existence d'un corps K à p^n éléments, extension de degré n de $\mathbb{F}_p$. Soit x un générateur de K^* et notons $f: \mathbb{F}_p[X] \rightarrow K$ l'homomorphisme défini par $f(X) = x$; f est surjectif ; son noyau est l'idéal des polynômes

de $\mathbb{F}_p[X]$ s'annulant en x ; c'est un idéal principal, engendré par un polynôme $P \in \mathbb{F}_p[X]$. On a donc un isomorphisme $\varphi : \frac{\mathbb{F}_p[X]}{(P)} \rightarrow K$. Par conséquent, P est un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$, d'après 20.2. $\square$

Pour construire un corps de p^n éléments, il suffit donc de savoir trouver un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$.

Théorème 2. *Deux corps finis ayant le même nombre d'éléments sont nécessairement isomorphes.*

Démonstration. Soient K et K' deux corps finis, de caractéristique p et ayant $q = p^n$ éléments. En reprenant la démonstration du corollaire précédent, un générateur x du groupe multiplicatif K^* permet de construire un homomorphisme surjectif $f: \mathbb{F}_p[X] \rightarrow K$ qui se factorise par la projection canonique $\pi : \mathbb{F}_p[X] \rightarrow \mathbb{F}_p[X]/(P)$ en un isomorphisme φ . Comme $x^{q-1} = 1$, x est racine du polynôme $X^{q-1} - 1$ donc P divise $X^{q-1} - 1$; posons $X^{q-1} - 1 = P(X)S(X)$.

$$\begin{array}{ccccc}
 K' & \xleftarrow{f'} & \mathbb{F}_p[X] & \xrightarrow{f} & K \\
 & \swarrow \varphi' & \downarrow \pi & \searrow \varphi & \\
 & & \mathbb{F}_p[X]/(P) & &
 \end{array}$$

Les $q - 1$ éléments non nuls de K' sont racines de $X^{q-1} - 1$. Comme P n'est pas une constante, il existe au moins un élément $y \neq 0$ de K' qui est racine de P ; soit $f' : \mathbb{F}_p[X] \rightarrow K'$ l'homomorphisme défini par $f'(X) = y$; comme $f'(P) = 0$, il existe un homomorphisme de corps $\varphi' : \mathbb{F}_p[X]/(P) \rightarrow K'$ tel que $f' = \varphi' \circ \pi$; φ' est injectif comme homomorphisme de corps donc est un isomorphisme puisque $\mathbb{F}_p[X]/(P)$ et K' ont même nombre d'éléments. Le composé $\varphi' \circ \varphi^{-1}$ est un isomorphisme de K sur K' . $\square$

Le théorème 2 permet d'affirmer que, si on choisit deux polynômes irréductibles P_1 et P_2 de degré n dans $\mathbb{F}_p[X]$, les corps $K_1 = \mathbb{F}_p[X]/(P_1)$ et $K_2 = \mathbb{F}_p[X]/(P_2)$ sont deux corps isomorphes, tous deux à p^n éléments. Voici un exemple de cette situation.

Corps à seize éléments. Posons $K = \mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$.

1) Les polynômes irréductibles de degré 4 de $K[X]$ sont $P_1 = X^4 + X + 1$, $P_2 = X^4 + X^3 + 1$, $P_3 = X^4 + X^3 + X^2 + X + 1$ (voir l'exercice 20.2 1)).

2) Chacun des trois polynômes précédents permet de construire un corps à 16 éléments et le théorème 2 montre que les trois corps obtenus sont isomorphes. Pour construire ces isomorphismes, nous allons d'abord décrire les trois corps $K_i = \frac{K[X]}{(P_i)}$, $i = 1, 2, 3$, en notant x_i la classe de X dans K_i . Les tables des valeurs des puissances x_i^k de x_i , $k = 0, \dots, 5$, exprimées dans la base $(1, x_i, x_i^2, x_i^3)$ sont les suivantes où seules les coordonnées non nulles sont indiquées.

x_1^k dans K_1					x_2^k dans K_2					x_3^k dans K_3				
k	1	x_1	x_1^2	x_1^3	k	1	x_2	x_2^2	x_2^3	k	1	x_3	x_3^2	x_3^3
0	1				0	1				0	1			
1		1			1		1			1		1		
2			1		2			1		2			1	
3				1	3				1	3				1
4	1	1			4	1			1	4	1	1	1	1
5		1	1		5	1	1		1	5	1			

On voit que x_1 et x_2 sont des générateurs de K_1^* et K_2^* respectivement, mais que x_3 est d'ordre 5 (la construction complète des tables n'est pas nécessaire pour cela ; il suffit de s'arrêter aux puissances cinquièmes). Pour construire un isomorphisme de K_2 dans K_1 , on cherche à envoyer x_1 sur un générateur de K_2 vérifiant la même équation. Comme x_2 vérifie $x_2^4 + x_2^3 + 1 = 0$, on voit que $y = x_2^{-1}$ vérifie $y^4 + y + 1 = 0$; on peut donc définir l'isomorphisme $K_1 \rightarrow K_2$ par $x^k \mapsto y^k$. Pour construire un isomorphisme $K_2 \rightarrow K_3$, il faut commencer par trouver un générateur de K_3^* , puisque x_3 ne convient pas ; il y en a $\varphi(15) = 8$; on essaie $1 + x_3$: $(1 + x_3)^3 = 1 + x_3 + x_3^2 + x_3^3 \neq 1$, $(1 + x_3)^5 = 1 + x_3 + x_3^4 + x_3^5 = 1 + x_3^2 + x_3^3 \neq 1$ montrent que $z = 1 + x_3$ est d'ordre 15. On a $z^4 + z^3 + 1 = 1 + x_3^4 + 1 + x_3 + x_3^2 + x_3^3 + 1 = 0$, ce qui permet de définir l'isomorphisme $K_2 \rightarrow K_3$ par $y^k \mapsto z^k$; pour obtenir un isomorphisme $K_1 \rightarrow K_3$, on peut passer par K_2 et on trouve l'isomorphisme défini par $x_1^k \mapsto (1 + x_3)^{-k}$.

21.3 LOI DE RÉCIPROCITÉ QUADRATIQUE

La loi de réciprocité quadratique permet, étant donnés deux nombres premiers impairs p et q , de relier les propriétés : p est un carré modulo q et q est un carré modulo p . Quand un nombre est un carré modulo p , on dit qu'il est résidu quadratique modulo p .

Des cas particuliers de cette loi ont été donnés par Fermat, qui a indiqué pour quelles valeurs de p premier, ± 2 , ± 3 sont des carrés modulo p ; ces affirmations ont été démontrées par Lagrange et Euler. La loi générale fut conjecturée, en 1785, par Adrien-Marie Legendre (1752-1833) qui en donna une démonstration incomplète (il y a plusieurs demi-échecs dans la vie de Legendre...). Le jeune Gauss appelait la loi de réciprocité quadratique *théorème d'or* ; il en a découvert les premières démonstrations en 1796, à 19 ans.

21.3.1 Symbole de Legendre

Pour tout nombre $p > 2$ premier et tout entier $a \in \mathbb{Z}$ non nul modulo p , on définit le symbole de Legendre $\left(\frac{a}{p}\right)$ comme étant égal à 1 si a est un carré modulo p , à -1 dans le cas contraire. Si $a = a' \bmod p$, on a $\left(\frac{a}{p}\right) = \left(\frac{a'}{p}\right)$, ce qui permet de définir le symbole de Legendre pour les éléments de $(\mathbb{Z}/p\mathbb{Z})^*$.

Propriétés du symbole de Legendre. Soit $p > 2$ premier. Les propriétés suivantes sont vraies.

- 1) Pour $a \in \mathbb{Z}$ premier avec p , on a : $\left(\frac{a}{p}\right) = a^{(p-1)/2} \bmod p$, (critère d'Euler).
- 2) Pour $a, b \in \mathbb{Z}$ premiers avec p , on a : $\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$.

Démonstration. 1) Si $a = b^2 \bmod p$, on a $a^{(p-1)/2} = b^{p-1} = 1 \bmod p$.

Supposons maintenant que a ne soit pas un carré modulo p . Soit g un générateur de $(\mathbb{Z}/p\mathbb{Z})^*$. Dans un corps, l'équation $x^2 = 1$ a deux racines : ± 1 ; comme $(g^{(p-1)/2})^2 = g^{p-1} = 1 \bmod p$ et que $g^k \neq 1$ pour $1 \leq k < p-1$, on a $g^{(p-1)/2} = -1 \bmod p$. Il existe un entier $r \leq p-1$ tel que $a = g^r \bmod p$. Comme a n'est pas un carré, r est impair. On a

$$a^{(p-1)/2} = (g^{(p-1)/2})^r = (-1)^r = -1 \bmod p,$$

ce qui démontre le critère d'Euler.

- 2) Cette dernière propriété résulte immédiatement du critère d'Euler. □

21.3.2 Loi de réciprocité quadratique

Énoncé de la loi de réciprocité quadratique.

1) **Loi pour -1 .** Si $p > 2$ est premier, on a $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$.

Par conséquent, -1 est un carré modulo p si $p \equiv 1 \pmod{4}$ et -1 n'est pas un carré modulo p si $p \equiv 3 \pmod{4}$.

2) **Loi pour 2 .** Si $p > 2$ est premier, on a $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$.

Par conséquent, 2 est un carré modulo p si $p \equiv \pm 1 \pmod{8}$ et 2 n'est pas un carré modulo p si $p \equiv \pm 3 \pmod{8}$.

3) **Loi de réciprocité.** Si p, q sont des nombres premiers impairs distincts, on a

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Démonstration. Nous suivons la démonstration donnée par Gauss dans les *Disquisitiones* en 1801.

Loi pour -1 .

Pour $a = -1$, le critère d'Euler montre que $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} \pmod{p}$. Si $p \equiv 1 \pmod{4}$,

on a $\left(\frac{-1}{p}\right) = 1$ et si $p \equiv 3 \pmod{4}$, on a $\left(\frac{-1}{p}\right) = -1$, d'où la loi pour -1 .

Pour montrer les 2) et 3), des préliminaires sont utiles.

Préliminaires.

Soient $p > 2$ un nombre premier et a non nul modulo p .

On pose :

- $r = \frac{p-1}{2}$;
- $I = \{k, k \in \mathbb{N} \text{ et } 1 \leq k \leq r\}$; on a encore $I = \{k, k \in \mathbb{N} \text{ et } 1 \leq k < p/2\}$;
- $J = \{k, k \in \mathbb{Z} \text{ et } -r \leq k \leq r\}$; on a encore $J = \{k, k \in \mathbb{Z} \text{ et } -p/2 < k < p/2\}$;
- $f: \mathbb{Z} \rightarrow J$, fonction telle que $f(n)$ soit l'entier de J égal à n modulo p .

Pour tout $k \in I$, on a $ak \neq 0 \pmod{p}$, donc $|f(a.k)| \in I$, ce qui permet de définir une application $g: I \rightarrow I$ par $g(k) = |f(a.k)|$. Cette application est injective car si $|f(a.j)| = |f(a.k)|$ pour $j, k \in I$, on a $aj = \pm ak \pmod{p}$, d'où $j = \pm k \pmod{p}$, ce qui impose $j = k$. On en déduit que g est surjective.

Au couple (a, p) , on associe le nombre s d'entiers k de I tels que $f(ak) < 0$.

Calculons le produit $P = \prod_{k \in I} f(a.k) \bmod p$ de deux manières différentes. D'une part, l'ensemble des valeurs de $|f(a.k)|$ étant I quand k décrit I d'après la remarque précédente, on a $P = (-1)^s r! \bmod p$. D'autre part, $P = \prod_{k \in I} (ak) \bmod p = a^r r! \bmod p$. L'égalité de ces deux expressions donne $(-1)^s = a^r = \left(\frac{a}{p}\right)$, d'après le critère

d'Euler. La détermination de la parité de s permet donc de savoir si a est ou non un carré modulo p . Elle va remplacer le critère d'Euler dans ce qui suit.

Pour suivre la démarche de Gauss, nous allons prendre deux exemples.

Si $a = 13$ et $p = 67$, on a $r = 33$ et les valeurs de $13k \bmod p$, pour $k \in I = \{1, \dots, 33\}$, sont 13, 26, 39, 52, 65, 11, 24, 37, 50, 63, 9, 22, 35, 48, 61, 7, 20, 33, 46, 59, 5, 18, 31, 44, 57, 3, 16, 29, 42, 55, 1, 14, 27. On trouve $s = 15$; par conséquent, $\left(\frac{13}{67}\right) = (-1)^{15} = -1$, ce qui montre que 13 n'est pas un carré modulo 67.

Si $a = 67$ et $p = 13$, on a $r = 6$ et $a = 2 \bmod 13$; les valeurs de $2k \bmod p$, pour $k \in I = \{1, \dots, 6\}$, sont 2, 4, 6, 8, 10, 12. On trouve $s = 3$; par conséquent, $\left(\frac{67}{13}\right) = \left(\frac{2}{13}\right) = (-1)^3 = -1$, ce qui montre que 67 n'est pas un carré modulo 13.

Loi pour 2.

Le calcul précédent donne un critère pour reconnaître si 2 est ou non un carré modulo p . On peut en effet calculer s suivant les valeurs de p . Si $p = 1 \bmod 4$, $p = 4t + 1$, on a $r = 2t$ et $f(2k) < 0$ pour les k tels que $t < k \leq 2t$, ce qui donne $s = t = (p - 1)/4$. Si $p = 3 \bmod 4$, $p = 4t + 3$, on a $r = 2t + 1$ et $f(2k) < 0$ pour les k tels que $t < k \leq 2t + 1$, ce qui donne $s = t + 1 = (p + 1)/4$. Ces résultats s'expriment en fonction des valeurs de p modulo 8 :

$p \bmod 8$	1	3	5	7
$\left(\frac{2}{p}\right)$	1	-1	-1	1

Par exemple, 2 est un carré modulo 71 et modulo 73, mais ce n'est pas un carré modulo 67. On notera que, quand 2 est un carré, la loi ne permet pas de préciser de quel nombre.

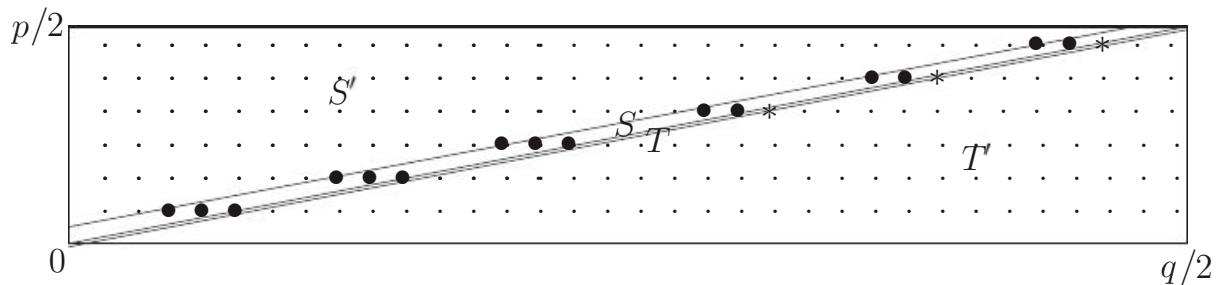
Loi pour des nombres premiers impairs distincts.

L'argument de Gauss est basé sur des dénombrements de points à coordonnées entières dans le rectangle $R = \{(x, y), 0 < x < q/2, 0 < y < p/2\}$. Notons u et v les

entiers s définis comme précédemment pour les couples (p, q) et (q, p) respectivement.

Pour définir u , étudions $f(pk)$ pour $1 \leq k < q/2$; on a $f(pk) < 0$ si on a une inégalité de la forme $qm - q/2 < pk < qm$. Le nombre u est donc le nombre des points (k, m) à coordonnées entières de R tels que $0 < qm - pk < q/2$. Notons S l'ensemble qu'ils forment.

De même, en échangeant les rôles de k et m , on vérifie que v est le nombre des points (k, m) à coordonnées entières de R tels que $0 > qm - pk > -p/2$. Notons T l'ensemble qu'ils forment et notons S' et T' les sous-ensembles de R de points (k, m) à coordonnées entières définis par $qm - pk > q/2$ et $qm - pk < -p/2$. Voici une représentation graphique de ces différents ensembles pour $p = 13$ et $q = 67$; les points de S sont notés avec des $\bullet$, les points de T avec des $\star$. Les bandes dans lesquelles figurent ces points sont étroites.



On vérifie que l'application g définie par $g(x, y) = (q + 1)/2 - x, (p + 1)/2 - y$ échange les points de S' et de T' . Par conséquent, la somme de leurs cardinaux est paire. On en déduit que $u + v$ est de la même parité que le nombre de points de R ,

c'est-à-dire $\frac{p-1}{2} \times \frac{q-1}{2}$. D'où :

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{u+v} = (-1)^{(p-1)(q-1)/4}. \quad \square$$

Il existe aujourd'hui plus d'une centaine de démonstrations différentes du *theorema aureum*. Celle-ci ne vous paraît-elle pas très belle ? La loi de réciprocité quadratique rend extrêmement aisée la réponse à la question : a est-il un carré modulo p ? Par exemple, 13 n'est pas un carré modulo 67 car

$$\left(\frac{13}{67}\right) = \left(\frac{67}{13}\right) (-1)^{(66 \times 12)/4} = \left(\frac{2}{13}\right) = -1 \text{ car } 13 \equiv 5 \pmod{8}.$$

21.4 FACTORISATION DANS $\mathbb{Z}[\iota]$, THÉORÈME DES DEUX CARRÉS

L'anneau $A = \mathbb{Z}[\iota]$ est principal et factoriel, étant euclidien d'après 19.12. Tout élément de cet anneau admet donc une décomposition unique en produit de facteurs

irréductibles. Nous allons maintenant déterminer ces éléments à partir d'une succession de petits résultats.

- 1) Un entier de $\mathbb{Z}$ non premier est réductible dans $\mathbb{Z}$ et, *a fortiori*, dans A .
- 2) Les unités de A sont les éléments de norme 1, c'est-à-dire ± 1 et $\pm i$. L'élément $1 + i$ est irréductible dans A car sa norme est 2 ; ses associés $\pm(1 + i)$ et $\pm(1 - i)$ le sont également. Enfin, $2 = (1 + i)(1 - i)$ n'est pas irréductible dans A et $2 = 1 + 1$ montre que 2 est une somme de deux carrés.
- 3) Si un nombre premier p impair est réductible dans A , c'est que $p = \alpha\beta$ avec $\alpha, \beta \in A$ non inversibles dans A . On a $N(\alpha)N(\beta) = N(p) = p^2$, donc $N(\alpha) = N(\beta) = p$, ce qui prouve que p est la norme d'un élément $a + bi$; on a donc $p = (a + bi)(a - bi) = a^2 + b^2$. Comme les carrés modulo 4 sont 0 et 1 et que $p = 1 \bmod 4$ ou $p = 3 \bmod 4$, on a nécessairement $p = 1 \bmod 4$.
- 4) Montrons maintenant la réciproque du 3) : tout nombre premier p tel que $p = 1 \bmod 4$ est réductible. On sait que -1 est un carré dans $\mathbb{Z}/p\mathbb{Z}$. Par conséquent, $X^2 + 1$ est réductible dans $(\mathbb{Z}/p\mathbb{Z})[X]$, et $\frac{(\mathbb{Z}/p\mathbb{Z})[X]}{(X^2 + 1)}$ n'est pas un corps. Mais ce dernier anneau est isomorphe à $\frac{A}{(p)}$, ce qui prouve que p est réductible dans A .
- 5) Comme le produit de deux sommes de deux carrés est une somme de deux carrés, d'après la formule $(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2$, on voit que les entiers dont la valuation en des nombres égaux à 3 mod 4 est paire peuvent s'écrire comme sommes de deux carrés.

On peut énoncer le théorème suivant.

Théorème des deux carrés. *Soit p un nombre premier.*

- 1) Si $p = 2$, p se décompose dans $\mathbb{Z}[i]$ en produit de deux facteurs irréductibles associés $p = (1 + i)(1 - i) = i(1 - i)^2$.
- 2) Si $p = 3 \bmod 4$, alors p est irréductible dans $\mathbb{Z}[i]$ et n'est pas somme des carrés de deux entiers.
- 3) Si $p = 1 \bmod 4$, alors p est réductible dans $\mathbb{Z}[i]$ et se décompose en produit de deux facteurs irréductibles non associés $p = (a + ib)(a - ib)$. Dans ce cas, p est la somme des carrés de deux entiers.
- 4) Un entier N produit de facteurs premiers tel que les valuations $v_p(N)$ pour p premier, $p = 3 \bmod 4$, soient paires, s'écrit comme somme de deux carrés.

Commentaires. L'exercice 21.7 montre une réciproque du 4) et aborde le problème de la détermination d'entiers a et b tels que $p = a^2 + b^2$ quand $p = 1 \bmod 4$.

Le théorème des deux carrés a été mentionné par Girard en 1625 ; Fermat l'énonce à plusieurs reprises dans ses lettres. Il en donne même une démonstration dans une lettre célèbre d'août 1659 à Pierre de Carcavi (1600-1684) où il fait un bilan de ce qu'il a découvert en arithmétique.

Et pour ce que les méthodes ordinaires, qui sont dans les livres, étaient insuffisantes à démontrer des propositions si difficiles, je trouvai enfin une route tout à fait singulière pour y parvenir.

J'appelai cette manière de démontrer la descente infinie...

Après avoir donné deux exemples du principe de sa méthode, Fermat continue.

Je fus longtemps sans pouvoir appliquer ma méthode aux questions affirmatives, parce que le tour et le biais pour y venir est beaucoup plus malaisé que celui dont je me sers aux négatives. De sorte que, lorsqu'il me fallut démontrer que tout nombre premier, qui surpasse de l'unité un multiple de 4, est composé de deux carrés, je me trouvais en belle peine. Mais enfin, une méditation maintes fois répétée me donna les lumières qui me manquaient... si un nombre premier pris à discrétion, qui surpasse de l'unité un multiple de 4, n'est point composé de deux carrés, il y aura un nombre premier de même nature, moindre que le donné, et ensuite un troisième encore moindre, etc. en descendant à l'infini jusqu'à ce que vous arriviez au nombre 5, qui est le moindre de tous ceux de cette nature, lequel il s'ensuivrait n'être pas composé de deux carrés, ce qu'il est pourtant. D'où on doit inférer, par la déduction à l'impossible, que tous ceux de cette nature sont par conséquent composés de deux carrés.

Bien sûr, l'essentiel pour Fermat est d'indiquer le principe de sa méthode et il laisse au lecteur le soin de trouver comment passer d'un nombre premier à un nombre premier plus petit. Dans cette même lettre, souhaitant laisser avant sa mort (qui survient en 1665) des indications sur ce qu'il a trouvé et démontré, Fermat donne d'autres exemples de ses résultats, comme la résolution du cas $n = 3$ de son grand théorème (voir 19.14).

Voilà sommairement le compte de mes rêveries sur le sujet des nombres. Je ne l'ai écrit que parce que j'appréhende que le loisir de m'étendre et de mettre au long toutes ces démonstrations et ces méthodes me manquera ; en tout cas, cette indication servira aux savans pour trouver d'eux-mêmes ce que je n'étends point... Et peut-être la postérité me saura gré de lui avoir fait connaître que les Anciens n'ont pas tout su...

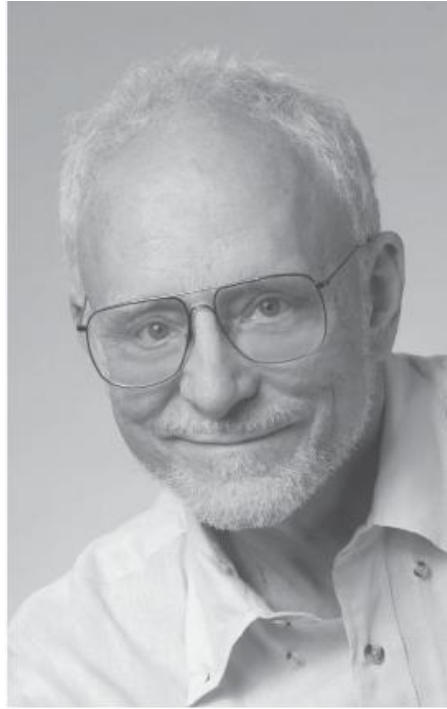
Multi pertransibunt et augebitur scientia (Beaucoup passeront et la science augmentera) concluait Fermat, d'après le livre de Daniel (12.4).

21.5 ALGORITHME DE BERLEKAMP

Nous allons montrer dans cette section comment concevoir un algorithme pour factoriser un polynôme à coefficients dans un corps fini $K = \mathbb{F}_p$, p premier, en produits de polynômes irréductibles de $K[X]$.

L'algorithme a été conçu en 1967 par Elwyn Berlekamp (1940-2019), professeur à l'Université de Berkeley en Californie. Un algorithme probabiliste plus rapide,

l'algorithme de Cantor-Zassenhaus, a été proposé en 1981. Berlekamp a aussi écrit (1994) un livre sur le décompte des points dans les fins de parties de go.



Elwyn Berlekamp

© Photo de Peg Skorpinski.

1) Soient P un polynôme de degré n de $K[X]$, $P = u \prod_{1 \leq i \leq r} P_i^{k_i}$ sa décomposition, définie à des inversibles près, en produit de facteurs irréductibles de $K[X]$. Pour $i = 1, \dots, r$, on pose $Q_i = P_i^{k_i}$: on a $P = u \prod_{1 \leq i \leq r} Q_i$. L'algorithme va déterminer successivement r , les Q_i , puis les P_i . On pose $E = K_{n-1}[X]$, espace vectoriel des polynômes de degré strictement inférieur à n à coefficients dans K ; la restriction à E de la projection canonique $K[X] \rightarrow \frac{K[X]}{(P)}$ est bijective.

2) Comme les Q_i , $i = 1, \dots, r$, sont premiers entre eux deux à deux, le théorème chinois donne un isomorphisme $\varphi : \frac{K[X]}{(P)} \rightarrow \prod_{1 \leq i \leq r} \frac{K[X]}{(Q_i)}$ défini par $\varphi(T \bmod P) = (T \bmod Q_1, \dots, T \bmod Q_r)$.

3) On définit une application $\Phi : E \rightarrow E$ en posant $\Phi(T)$ égal au reste de la division euclidienne de T^p par P ; on a $T^p = (T(X))^p = T(X^p)$ et Φ est linéaire. Le sous-espace $F = \ker(\Phi - \text{id})$ de E est l'espace propre associé à la valeur 1 ; ses éléments sont les polynômes T invariants par Φ , ceux pour lesquels il existe un polynôme Q de $K[X]$ tel que $T^p = QP + T$.

4) Si $T \in F$, P divise $T^p - T = T(T-1) \dots (T-(p-1))$, donc Q_i divise $T(T-1) \dots (T-(p-1))$ pour $i = 1, \dots, r$. Comme les $T-s$ sont premiers entre eux deux à deux, pour $s = 1, \dots, p-1$, il existe un unique $s_i \in K$ tel que Q_i divise $T - s_i$, autrement dit

$T = s_i \bmod Q_i$. La restriction φ' de φ à F est donc à valeurs dans le sous-anneau K^r de $\prod_{1 \leq i \leq r} \frac{K[X]}{(Q_i)}$, ce que décrit le diagramme :

$$\begin{array}{ccc} F = \ker(\Phi - \text{id}) & \xrightarrow{\varphi'} & K^r \\ \text{incl} \downarrow & & \text{incl} \downarrow \\ \frac{K[X]}{(P)} & \xrightarrow{\varphi} & \prod_{1 \leq i \leq r} \frac{K[X]}{(Q_i)} \end{array}$$

5) Réciproquement, si $T \in E$ est tel que $\varphi(T) = (s_1, \dots, s_r) \in K^r$, alors Q_i divise $T - s_i$ pour $i = 1, \dots, r$ et, comme les Q_i sont premiers entre eux deux à deux, leur produit P divise $T(T - 1) \dots (T - (p - 1))$ et $T \in F$. L'application φ' est donc un isomorphisme. On a donc

$$r = \dim(F) = n - \text{rang}(\Phi - \text{id}).$$

6) Pour déterminer le rang de $\Phi - \text{id}$: il suffit d'écrire la matrice M de Φ dans la base $(1, X, \dots, X^{n-1})$ de E et d'appliquer l'algorithme du pivot de Gauss.

7) Si $r = 1$, c'est que le polynôme P est irréductible ou puissance d'un polynôme irréductible dans $K[X]$; on passe à la partie finale de l'algorithme décrite au 11).

8) Si ce n'est pas le cas, on détermine une base $(B_1, \dots, B_r)$ de F . On prend $B_1 = 1$.

9) Comme B_2 n'est pas le polynôme constant, le fait que P divise $B_2(B_2 - 1) \dots (B_2 - (n - 1))$ va nous apporter de nouvelles informations. On calcule $\text{pgcd}(P, B_2 - s)$ pour $s \in K$. On obtiendra ainsi des polynômes Q_i ou des produits de tels polynômes. On pourra éventuellement obtenir une factorisation de P avant d'avoir calculé tous les pgcd si on a obtenu le nombre r de facteurs attendu.

10) Mais il peut se faire qu'un produit de certains des Q_i divise un $B_2 - s$ et on ne pourra, dans ce cas, déterminer complètement les Q_i . Montrons que le calcul des $\text{pgcd}(P, B_i - s)$ pour $s \in K$ et $3 \leq i \leq r$ (au moins de certains d'entre eux) va permettre de les trouver. Sinon, il existerait deux des Q_i , disons Q_1 et Q_2 tels que, pour $i = 1, \dots, r$, il existe des $s_i \in K$ tels que $Q_1 Q_2$ divise $B_i - s_i$. Pour $i = 1, \dots, r$, on aurait $B_i = s_i \bmod Q_1 Q_2$ et les deux premières composantes de $\varphi'(B_i)$ seraient donc toujours égales ; elles le seraient donc pour tout T de F , ce qui est absurde puisque φ' est un isomorphisme.

11) Comment retrouver les P_i connaissant les Q_i ? Si Q_i est une puissance p -ième, il est de la forme $(\sum a_j X^j)^p = \sum a_j X^{pj}$; il faut reconnaître que les coefficients sont des puissances p -ièmes, extraire des racines p -ièmes, éventuellement recommencer, pour obtenir P_i . Si Q_i n'est pas une puissance p -ième, on a $Q_i' \neq 0$ et on peut calculer $\frac{Q_i}{\text{pgcd}(Q_i, Q_i')}$ pour éliminer la multiplicité et obtenir P_i .

Ceci achève la description de l'algorithme. Voici quelques exemples simples, les logiciels de calcul formel ayant évidemment la possibilité de traiter des exemples nécessitant beaucoup de calculs.

Décomposition de $P = X^3 + X^2 + X + 1$ dans $\mathbb{F}_2[X]$. On a $\Phi(1) = 1$, $\Phi(X) = X^2$, $\Phi(X^2) = X^4 = 1 \pmod{P}$, d'où

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ et } M - I = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Comme $M - I$ est de rang 2, on a $r = 1$, ce qui signifie que P est un polynôme irréductible ou une puissance d'un polynôme irréductible. On vérifie d'abord que P n'est pas un carré puisqu'il est de degré impair ; on calcule alors

$$P' = X^2 + 1, \text{pgcd}(P, P') = X^2 + 1 \text{ et } \frac{P}{X^2 + 1} = X + 1$$

pour trouver, si on ne l'avait pas vu depuis le début, $P = (X + 1)^3$.

Décomposition de $P = X^3 - X$ dans $\mathbb{F}_3[X]$. On a $\Phi(1) = 1$, $\Phi(X) = X^3 = X \pmod{P}$, $\Phi(X^2) = X^6 = X^2 \pmod{P}$, d'où $M = I$. Comme $M - I = 0$ est de rang 0, on a $r = 3$, ce qui signifie que P est un produit de trois polynômes du premier degré et que $F = E$. On prend $B_1 = 1$, $B_2 = X$, $B_3 = X^2$. On sait que P divise $B_2(B_2 - 1)(B_2 - 2)$; le calcul de $\text{pgcd}(P, X)$ et $\text{pgcd}(P, X - 1)$ donne les facteurs X et $X - 1$, le troisième s'obtient par quotient sans calculer le dernier pgcd.

Exercices

21.1 Indicateur d'Euler

Pour tout entier $n \geq 1$, on note $\varphi(n)$ le nombre d'entiers naturels k tels que $1 \leq k \leq n$ et k premier avec n ; c'est l'ordre du groupe $U(n)$ des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$.

- 1) Montrer que $\varphi(p) = p - 1$ pour tout nombre premier p .
- 2) Montrer que $\varphi(p^r) = p^r - p^{r-1}$ pour tout nombre premier p et tout entier $r \geq 1$.
- 3) Montrer que, si a et b sont des entiers premiers entre eux, on a :

$$\varphi(ab) = \varphi(a) \varphi(b).$$

- 4) Montrer que $\varphi(n)$ est pair pour $n > 2$.
- 5) Montrer que n divise $\varphi(a^n - 1)$ pour $a > 1$. Faire une vérification directe pour $a = 12$ et $n = 7$.

6) Calcul de $\varphi(n)$: montrer que, si un entier n admet la décomposition en facteurs premiers :

$$n = \prod_{1 \leq i \leq r} p_i^{k_i}, \text{ on a : } \varphi(n) = n(1 - 1/p_1) \dots (1 - 1/p_r).$$

7) Montrer que $\varphi(n)$ est égal au nombre de générateurs de $(\mathbb{Z}/n\mathbb{Z}, +)$.

8) Montrer que, pour tout entier a premier avec n , on a : $a^{\varphi(n)} = 1 \pmod n$ (ce résultat, publié en 1760 par Euler, généralise le petit théorème de Fermat).

9) Soit $n \in \mathbb{N}^*$. On suppose que $n = ds$ avec $1 \leq d, s \leq n$. Montrer qu'il existe $\varphi(d)$ éléments d'ordre d dans $\mathbb{Z}/n\mathbb{Z}$. En déduire que :

$$n = \sum_{d|n} \varphi(d).$$

21.2 Applications de la loi de réciprocité quadratique

1) Soient K un corps fini d'ordre q , g un générateur de K^* et $a \in K$ non nul. Donner un test pour déterminer si a est un carré ou non dans K (on étudiera d'abord le cas où la caractéristique de K est 2). Si la réponse est positive, montrer comment calculer simplement une racine carrée de a dans le cas où $q = 3 \pmod 4$.

2) 71 et 72 sont-ils des carrés modulo 239 ? Si la réponse est positive, préciser de quel nombre.

3) Pour quelles valeurs de p premier, avec $p > 3$, l'entier -3 est-il un carré ?

4) Pour quelles valeurs de p premier, avec $p > 3$, l'entier 3 est-il un carré ?

5) a) Soient p un nombre premier tel que $p \equiv 1 \pmod 3$ et $a \in U(p)$. Montrer que a est un cube dans $U(p)$ si et seulement si $a^{(p-1)/3} = 1 \pmod p$.

b) Soit p un nombre premier tel que $p \equiv 2 \pmod 3$. Montrer que tout élément de $\mathbb{Z}/p\mathbb{Z}$ est un cube.

6) Décrire l'ensemble des nombres premiers p tels que $x^2 - 16x + 24 = 0$ ait une solution dans $\mathbb{F}_p$.

21.3 Polynômes irréductibles de $\mathbb{Z}[X]$

Parmi les polynômes $X^4 - 2X^2 + 4, X^4 + 1, X^4 + 4X^2 + 4$, lesquels sont :

- réductibles dans $\mathbb{Z}[X]$, dans $\mathbb{Q}[X]$?
- réductibles dans $\mathbb{Z}/p\mathbb{Z}$ pour tout entier premier p ?

Pour la deuxième question, on pourra remarquer que dans un groupe cyclique G , étant donnés deux éléments a et b , l'un au moins des trois éléments a , b ou ab est toujours un carré dans G , puis écrire les polynômes proposés comme différence de deux carrés de trois façons différentes, pour pouvoir appliquer ce résultat avec des valeurs bien choisies de a et b .

21.4 Un théorème de Lucas

Édouard Lucas (1842-1891) est connu pour de jolis résultats en théorie des nombres, comme le test de primalité toujours en usage qui porte son nom et avec lequel il montra que le nombre de Mersenne $2^{127} - 1$ est premier en 1876. Ses récréations mathématiques sont un classique, comme le casse-tête des tours de Hanoï qu'il proposa sous le nom de N. Claus de Siam du collège Li-Sou-Stian (anagrammes de Lucas d'Amiens et du lycée Saint-Louis). Il mourut d'une affection due à un streptocoque après avoir été éraflé par un morceau d'un plat cassé au cours d'une réception. Soit p un nombre premier.

1) Soient a, b des entiers positifs avec $0 \leq b < p$. Montrer, en calculant $(1 + X)^{ap+b} \bmod p$ de deux manières différentes, que, pour des entiers u, v avec $0 \leq v < p$, on a $C_{ap+b}^{up+v} = C_a^u C_b^v \bmod p$ si $0 \leq v \leq b$ et $C_{ap+b}^{up+v} = 0 \bmod p$ si $b < v < p$.

2) En déduire que, si k, n sont deux entiers, $k \leq n$, dont les écritures en base p sont $k = \sum_{0 \leq i \leq r} k_i p^i$ et $n = \sum_{0 \leq i \leq r} n_i p^i$, on a $C_n^k = \prod_{0 \leq i \leq r} C_{n_i}^{k_i} \bmod p$ si $0 \leq k_i \leq n_i < p$ pour tout $i \leq r$ et $C_n^k = 0 \bmod p$ sinon.

21.5 Structure du groupe $U(n)$ des unités de $\mathbb{Z}/n\mathbb{Z}$

On rappelle que, si a et b sont des entiers premiers entre eux, $U(ab)$ est isomorphe à $U(a) \times U(b)$.

1) On suppose que n est un nombre premier. Quelle est la structure de $U(n)$?

2) Déterminer la structure de $U(n)$ pour $n = 2, 4, 6, 8, 9, 15$.

3) On suppose que n est de la forme p^k , où p est un nombre premier impair et $k > 1$ un entier naturel.

a) Soit a un entier dont la classe modulo p engendre $U(p)$. On pose $b = a^{p^{k-1}}$. Montrer que la classe de b modulo p^k est d'ordre $p - 1$ dans $U(n)$.

b) Montrer que, pour tout entier $r \geq 0$, on a :

$$(1 + p)^{p^r} = 1 + p^{r+1} \bmod p^{r+2}.$$

c) En déduire que $1 + p$ est d'ordre p^{k-1} dans $U(n)$, puis que $U(n)$ est cyclique.

d) Déterminer un générateur de $U(25)$, puis un générateur de $U(125)$.

4) On suppose que n est de la forme 2^k .

a) Montrer que, pour tout $r \geq 0$, on a : $5^{2^r} = 1 + 2^{r+2} \bmod 2^{r+3}$.

b) En déduire que la classe de 5 est d'ordre 2^{k-2} dans $U(n)$.

c) Montrer que, pour $k \geq 2$, une puissance de 5 n'est jamais égale à -1 modulo n .

d) En déduire que, pour tout $k \geq 2$, $U(n)$ est isomorphe à $\mathbb{Z}/2^{k-2}\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

- 5) a) Quelle est la structure de $U(200)$?
b) Quelle est la structure de $\text{Aut}(\mathbb{Z}/64\mathbb{Z})$?
c) Pour quelles valeurs de n le groupe $U(n)$ est-il cyclique ?
d) Pour quelles valeurs de n le groupe $U(n)$ possède-t-il exactement un élément d'ordre 2 ?

21.6 Nombres de Fermat

On pose $F_n = 2^{2^n} + 1$ pour $n \geq 0$.

- 1) Calculer F_n pour $n \leq 5$. Vérifier que F_n est premier pour $n = 0, 1, 2$.
2) a) Montrer que $F_{n+k} - 2 = (F_n - 2) \prod_{n \leq i \leq n+k-1} F_i$. Montrer que le dernier chiffre de F_n est un 7 pour $n \geq 2$.
b) En déduire que F_n et F_{n+k} sont premiers entre eux pour tout n et tout $k \neq 0$.
c) Donner une démonstration de l'infinitude de la suite des nombres premiers utilisant le résultat du b).
3) a) Soit p un facteur premier de F_n . Montrer que $p = 1 \pmod{2^{n+1}}$ et que, si $n \geq 2$, $p = 1 \pmod{2^{n+2}}$.
b) Montrer que F_3 et F_4 sont premiers et que F_5 ne l'est pas.
Les nombres F_n ont intrigué Fermat ; dans plusieurs lettres, il affirme être absolument convaincu qu'ils sont tous premiers, mais qu'il ne l'a pas encore montré. Euler s'est aperçu de la fausseté de cette affirmation pour $n = 5$ en 1732. Le calcul sur ordinateur de facteurs premiers des F_n pour $n = 6, \dots$ est un sport pratiqué depuis 40 ans ; des sites Internet en donnent les résultats.
4) On suppose F_n premier avec $n \geq 2$.
a) Montrer que si a est résidu quadratique modulo F_n , alors a n'engendre pas $U(F_n)$.
b) Montrer que si a n'est pas résidu quadratique modulo F_n , alors a engendre $U(F_n)$.
c) Montrer que 3, 5, 7 engendrent $U(F_n)$.

- 5) Montrer que F_n est premier pour $n \geq 1$ si et seulement si

$$3^{(F_n-1)/2} = -1 \pmod{F_n}.$$

Ce critère de primalité des nombres de Fermat a été publié par Théophile Pépin (1826-1904) en 1877. Il permet de montrer que le nombre F_7 (un nombre de 39 chiffres) n'est pas premier.

21.7 Nombres premiers et sommes de deux carrés 1

Soit p un nombre premier tel que $p = 1 \pmod{4}$. On sait que -1 est résidu quadratique modulo p et on suppose connaître un entier a , tel que $a^2 = -1 \pmod{p}$, autrement dit tel que $a^2 + 1 = kp$ (on peut supposer $1 < a < p$, d'où $1 \leq k < p - 1$). Le but

du problème est de montrer comment on peut en déduire des entiers a et b tels que $p = a^2 + b^2$.

1) Montrer qu'en choisissant c au hasard entre 1 et $p-1$, on a une chance sur 2 que $c^{(p-1)/2} = -1 \pmod{p}$. Montrer que, si c'est le cas, on peut en déduire un nombre d , $1 < d < p$, et un entier k , $k < p-1$, tels que $d^2 + 1 = kp$.

2) Soient u, v, k des entiers tels que $u^2 + v^2 = kp$ et $1 \leq k < p-1$. On définit w et x par $w = u \pmod{k}$, $x = v \pmod{k}$ et $-\frac{k}{2} < w, x \leq \frac{k}{2}$. Montrer qu'il existe k_1 , $1 \leq k_1 \leq k/2$ tel que $w^2 + x^2 = k_1 k$. En déduire des entiers y et z tels que $y^2 + z^2 = k_1 p$.

3) Utiliser la question précédente pour écrire 29 et 157 comme sommes de deux carrés sachant que $12^2 + 1 = 5 \times 29$ et $28^2 + 1 = 5 \times 157$.

4) Montrer que si un entier n est somme de deux carrés, les diviseurs p de n tels que $p \equiv 3 \pmod{4}$ sont tels que $v_p(n)$ soit pair.

21.8 Nombres premiers et sommes de deux carrés 2

Voici une démonstration récente (publiée en 1990 par Don Zagier) et très élémentaire, du résultat de la section 21.4 : pour tout nombre premier p tel que $p \equiv 1 \pmod{4}$, il existe un couple d'entiers (x, y) tel que $p = x^2 + y^2$.

1) Soient E un ensemble et $f: E \rightarrow E$ une involution de E (c'est-à-dire une application telle que $f^2 = \text{id}$). On appelle point fixe de f un élément x de E tel que $f(x) = x$.

a) Montrer que, si E est fini et si f n'a qu'un seul point fixe, le cardinal de E est impair.

b) Montrer que, si le cardinal de E est impair, f a au moins un point fixe.

2) On pose $S = \{(x, y, z), x, y, z \in \mathbb{N}^3, x^2 + 4yz = p\}$,

$$A = S \cap \{(x, y, z), x < y - z\},$$

$$B = S \cap \{(x, y, z), y - z < x < 2y\},$$

$$C = S \cap \{(x, y, z), x > 2y\}.$$

Vérifier que S est l'union disjointe de A , B et C .

3) On définit $f: S \rightarrow \mathbb{R}^3$ par

$$f(x, y, z) = (x + 2z, z, y - x - z) \text{ si } x \in A,$$

$$f(x, y, z) = (2y - x, y, x - y + z) \text{ si } x \in B,$$

$$f(x, y, z) = (x - 2y, x - y + z, y) \text{ si } x \in C.$$

Montrer que f définit une involution de S .

4) a) Déterminer l'ensemble des points fixes de cette involution.

b) Conclure en considérant l'application $g: S \rightarrow S$ définie par $g(x, y, z) = (x, z, y)$.

Solutions

21.1 1) C'est clair.

2) Les entiers entre 1 et p^r qui ne sont pas premiers avec p^r sont de la forme pu avec $u = 1, \dots, p^{r-1}$; comme il y en a p^{r-1} , l'ordre de $U(n)$ est $p^r - p^{r-1} = p^{r-1}(p - 1)$.

3) L'isomorphisme d'anneaux du théorème chinois $\mathbb{Z}/ab\mathbb{Z} \simeq \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ se restreint à un isomorphisme $U(ab) \simeq U(a) \times U(b)$; d'où $\varphi(ab) = \varphi(a)\varphi(b)$.

4) Si n contient un facteur p^k avec p premier impair et $k = v_p(n)$, $\varphi(p^k)$ est pair d'après le 2), donc $\varphi(n)$ est pair d'après le 3) ; si $n = 2^k$ avec $k > 1$, le 2) montre encore que $\varphi(n)$ est pair.

5) Le groupe $U(a^n - 1)$ est d'ordre $\varphi(a^n - 1)$; dans ce groupe, l'élément a est d'ordre n . Le théorème de Lagrange (voir 11.4) montre que n divise $\varphi(a^n - 1)$.

Si $a = 12$ et $n = 7$, l'identité $a^7 - 1 = (a - 1)(a^6 + \dots + 1)$ montre que $12 - 1 = 11$ divise $12^7 - 1 = 35\,831\,807$; on a donc $12^7 - 1 = 11 \times 3\,257\,437$. Mais le calcul de $\varphi(12^7 - 1)$ demande une factorisation complète de $35\,831\,807$ qui n'est pas évidente et que donne immédiatement des logiciels comme Maple ou Sage : $35\,831\,807 = 11 \times 659 \times 4943$. On en déduit $\varphi(12^7 - 1) = 10 \times 658 \times 4942$, expression dans laquelle les deux derniers facteurs sont divisibles par 7.

6) Si $n = \prod_{1 \leq i \leq r} p_i^{k_i}$, on a : $\varphi(n) = \prod_{1 \leq i \leq r} \varphi(p_i^{k_i}) = \prod_{1 \leq i \leq r} (p_i^{k_i} - p_i^{k_i-1})$, d'où la formule du texte en mettant n en facteur.

7) Si a est un générateur de $(\mathbb{Z}/n\mathbb{Z}, +)$, il existe $b \in \mathbb{Z}$ tel que $ba = 1 \pmod{n}$, ce qui montre que a est inversible modulo n ; réciproquement, si a est inversible modulo n , il existe b tel que $ba = 1 \pmod{n}$, donc $(kb)a = k \pmod{n}$ pour tout k entier, ce qui montre que a engendre $(\mathbb{Z}/n\mathbb{Z}, +)$.

8) Comme l'ordre de a dans le groupe $U(n)$ divise l'ordre $\varphi(n)$ de $U(n)$ on a la formule du texte.

9) Notons E_d le sous-ensemble des éléments d'ordre d de $(\mathbb{Z}/n\mathbb{Z}, +)$; la multiplication par s est un homomorphisme injectif $(\mathbb{Z}/d\mathbb{Z}, +) \rightarrow (\mathbb{Z}/n\mathbb{Z}, +)$ qui met en bijection les générateurs du premier groupe et les éléments de E_d . Comme $(\mathbb{Z}/n\mathbb{Z}, +)$ est la réunion disjointe des E_d , n est somme des cardinaux des E_d ce qui donne la formule demandée.

Voici une seconde démonstration. On écrit les n fractions $\frac{k}{n}$ pour $k = 1, \dots, n$. On les réduit pour obtenir des numérateurs et dénominateurs premiers entre eux. Pour une telle fraction réduite, le dénominateur est un diviseur d de n et le numérateur un nombre premier avec d et inférieur à d . Chaque diviseur d de n apparaît et les numérateurs des fractions de dénominateur d sont les $\varphi(d)$ entiers inférieurs à d et premiers avec d ; d'où le résultat.

21.2 1) Supposons q pair (q est donc une puissance de 2). Dans ce cas, a est toujours un carré car, comme $a^q = a$, a est le carré de $a^{q/2}$.

Supposons q impair. Soit g un générateur de K^* ; a est un carré dans K si et seulement si a est une puissance paire de g . Si $a = g^{2r}$, on a $a^{(q-1)/2} = g^{(q-1)r} = 1$; si $a = g^{2r+1}$, on a $a^{(q-1)/2} = g^{(q-1)r} g^{(q-1)/2} = -1$. Le test est donc de calculer $a^{(q-1)/2}$ et de le comparer à 1.

Si $q = 4s + 3$ et si on sait que a est un carré dans K , on a $a^{(q-1)/2} = a^{2s+1} = 1$ et $b = a^{s+1}$ a pour carré a car $b^2 = a^{2s+2} = a$.

Si $q \equiv 1 \pmod{4}$, le calcul de racines carrées est plus délicat et se fait avec des méthodes probabilistes.

2) Comme $(70 \times 238)/4$ est impair, la loi de réciprocité quadratique donne $\left(\frac{71}{239}\right) = -\left(\frac{239}{71}\right)$; on en déduit $\left(\frac{71}{239}\right) = -\left(\frac{26}{71}\right) = -\left(\frac{2}{71}\right)\left(\frac{13}{71}\right)$. La loi de réciprocité quadratique donne alors $\left(\frac{2}{71}\right) = 1$ et $\left(\frac{13}{71}\right) = \left(\frac{71}{13}\right)$. Les propriétés du symbole de Legendre montrent que $\left(\frac{71}{13}\right) = \left(\frac{6}{13}\right) = \left(\frac{2}{13}\right)\left(\frac{3}{13}\right)$. On applique une nouvelle fois la loi de réciprocité quadratique $\left(\frac{2}{13}\right) = -1$ et $\left(\frac{3}{13}\right) = \left(\frac{13}{3}\right) = 1$. Finalement, $\left(\frac{71}{239}\right) = 1$; 71 est donc un carré modulo 239. De quel nombre ? Il n'y a pas de belle méthode; on peut vérifier que $73^2 = 71 \pmod{239}$. Le calcul de $\left(\frac{72}{239}\right)$ est plus rapide car $72 = 2 \times 2^2 \times 3^2$, d'où $\left(\frac{72}{239}\right) = \left(\frac{2}{239}\right)$; la loi de réciprocité quadratique indique que 2 est un carré modulo 239; par conséquent, 72 est un carré modulo 239; on peut vérifier que $116^2 = 72 \pmod{239}$.

3) On a $\left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = (-1)^{(p-1)/2} (-1)^{(p-1)/2} \left(\frac{p}{3}\right) = \left(\frac{p}{3}\right)$.

On en déduit que -3 est un carré modulo p si $p \equiv 1 \pmod{3}$ et que -3 n'est pas un carré modulo p si $p \equiv 2 \pmod{3}$.

4) On trouve cette fois-ci $\left(\frac{3}{p}\right) = (-1)^{(p-1)/2} \left(\frac{p}{3}\right)$. En raisonnant modulo 12, on construit le tableau suivant.

$p \pmod{12}$	1	5	7	11
$(-1)^{(p-1)/2}$	1	1	-1	-1
$\left(\frac{p}{3}\right)$	1	-1	1	-1
$\left(\frac{3}{p}\right)$	1	-1	-1	1

5) a) Si $a = b^3$, on a $a^{(p-1)/3} = b^{p-1} = 1 \pmod p$. Pour la réciproque, posons $p - 1 = 3k$. Si g est un générateur de $U(p)$ et si $a = g^r$ vérifie $a^k = 1 \pmod p$, on a $g^{rk} = 1 \pmod p$, d'où $p - 1 = 3k$ divise rk ; par conséquent, 3 divise r ; en posant $r = 3s$, on a $a = (g^s)^3$.

b) Montrons que la fonction $f: \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$ définie par $f(x) = x^3$ est injective. Si $f(x) = f(y) = 0$, on a $x = y = 0$. Si $f(x) = f(y) \neq 0$, on a $x, y \neq 0$ et $(x - y)(x^2 + xy + y^2) = 0$. Posons $t = y/x$; l'équation $t^2 + t + 1 = 0$ s'écrit $(t + 1/2)^2 + 3/4 = 0$; elle a des solutions dans $\mathbb{Z}/p\mathbb{Z}$ si et seulement si -3 est un carré dans $\mathbb{Z}/p\mathbb{Z}$. On a vu au **3)** que ce n'est pas le cas. Par conséquent, $x = y$, donc f est injective ; on en déduit que f est surjective et le résultat.

6) Si $p = 2$, l'équation a 0 pour unique solution. Si $p = 5$, on vérifie directement que 3 est l'unique solution. Sinon, $p = 3$ ou $p > 5$. Dans ce cas, le discriminant doit être un carré modulo p ; comme il vaut $160 = 10 \times 4^2$, tout dépend de la valeur de $\left(\frac{10}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{5}{p}\right)$. On a $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right)$; on doit avoir $\left(\frac{2}{p}\right)$ et $\left(\frac{p}{5}\right)$ simultanément égaux à 1 ou à -1 ; le premier dépend de p modulo 8, le second modulo 5, ce qui conduit à raisonner modulo 40. On trouve $p = \pm 1, \pm 3, \pm 9, \pm 13$ modulo 40.

21.3 Résultat préliminaire

Soit g un générateur de G ; posons $a = g^m$, $b = g^n$; on a $ab = g^{m+n}$. L'un des trois nombres $m, n, m + n$ est nécessairement pair et la puissance de g correspondante est un carré.

1) Étude de $P = X^4 - 2X^2 + 4$

Comme $c(P) = 1$, l'étude de l'irréductibilité dans $\mathbb{Q}$ se ramène à l'étude de l'irréductibilité dans $\mathbb{Z}$. Si P a une racine entière a , a^2 divise 4 ; on vérifie que $\pm 1, \pm 2$ ne conviennent pas. Si P se factorise en produit de deux polynômes de degré 2 à coefficients dans $\mathbb{Z}$, on peut supposer ces deux polynômes unitaires ; l'égalité $X^4 - 2X^2 + 4 = (X^2 + aX + b)(X^2 + cX + d)$ équivaut au système :

$$\begin{cases} a + c = 0 \\ ac + b + d = -2 \\ bc + ad = 0 \\ bd = 4 \end{cases} \quad \text{équivaut à} \quad \begin{cases} c = -a \\ b + d - a^2 = -2 \\ a(b - d) = 0 \\ bd = 4 \end{cases}$$

Partons de la troisième équation. Si $a = 0$, $b + d = -2$ et $bd = 4$ conduisent à une impossibilité. Si $b = d$, $bd = 4$ impose $b = d = 2$ ou $b = d = -2$; ces deux possibilités donnent $a^2 = -2$ ou $a^2 = 6$ dans la seconde équation, ce qui est impossible et on peut conclure.

Pour montrer la réductibilité de P modulo p pour tout nombre premier p , on remarque d'abord que

$$P = (X^2 - 1)^2 + 3 = (X^2 + 2)^2 - 6X^2 = (X^2 - 2)^2 + 2X^2;$$

comme le groupe $(\mathbb{Z}/p\mathbb{Z})^*$ est cyclique, l'un des trois nombres $-2, -3, 6$ est un carré modulo p ; l'une des trois écritures de P se présente donc sous la forme d'une différence de deux carrés et permet de factoriser P .

2) Le traitement du cas $P = X^4 + 1$ est analogue. Quand au dernier cas, la méthode s'applique aussi, mais il vaut tout de même mieux s'apercevoir rapidement qu'il s'agit de $(X^2 + 2)^2$ et que tout est trivial !

21.4 1) D'une part, $(1 + X)^{ap+b} = \sum_{0 \leq l \leq ap+b} C_{ap+b}^l X^l$. D'autre part, $(1 + X)^{ap+b} = ((1 + X)^p)^a (1 + X)^b = (1 + X^p)^a (1 + X)^b \pmod{p}$; en développant le second membre, on voit que le coefficient de X^{up+v} dans $(1 + X^p)^a (1 + X)^b$ est $C_a^u C_b^v$, si $v < b$, et 0 sinon, d'où l'égalité demandée.

2) Il suffit d'appliquer le résultat précédent.

21.5 1) On a vu en 21.1 que $U(n)$ est un groupe cyclique d'ordre $n - 1$: $U(n) \simeq (\mathbb{Z}/(n - 1)\mathbb{Z}, +)$.

2) On a $U(2) = \{1\}$, $U(4) \simeq \mathbb{Z}/2\mathbb{Z}$, $U(6) \simeq U(2) \times U(3) \simeq \mathbb{Z}/2\mathbb{Z}$, $U(8) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ car les quatre éléments de $U(8)$: 1, 3, 5, 7 sont d'ordre 2, $U(9) \simeq \mathbb{Z}/6\mathbb{Z}$ car $U(9)$ a six éléments : 1, 2, 4, 5, 7, 8, est abélien et que le seul groupe abélien à six éléments est $\mathbb{Z}/6\mathbb{Z}$, $U(15) \simeq U(3) \times U(5) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ qui n'est pas cyclique.

3) a) Comme a est un générateur de $U(p)$, $a^r = 1 \pmod{p}$ implique que r est un multiple de $p - 1$. Pour $s = 1, \dots, p - 2$, sp^{k-1} n'est pas un multiple de $p - 1$, donc $b^s \neq 1 \pmod{p^k}$; comme $U(n)$ est d'ordre $p^{k-1}(p - 1)$, $b \pmod{p^k}$ est d'ordre $p - 1$ dans $U(n)$.

b) Raisonner par récurrence.

c) Le **b)** montre que $1 + p$ est d'ordre p^{k-1} dans $U(n)$, le **a)** que b est d'ordre $p - 1$. Les groupes engendrés par $1 + p$ et b ont une intersection réduite à 1 puisque leurs ordres sont premiers entre eux. On en déduit que $U(n)$ est produit des groupes cycliques $Gr(b)$ et $Gr(1 + p)$ (exercice 18.9, 4)). Le théorème chinois donne alors $U(n) \simeq (\mathbb{Z}/(p - 1)\mathbb{Z}) \times (\mathbb{Z}/p^{k-1}\mathbb{Z}) \simeq (\mathbb{Z}/p^{k-1}(p - 1)\mathbb{Z})$; $U(n)$ est cyclique.

d) Un générateur de $U(5)$ est 2; comme $2^5 = 7 \pmod{25}$, le **c)** montre que $7 \times 6 = 17 \pmod{25}$ engendre $U(25)$ (dont l'ordre est 20); comme $2^{25} = 57 \pmod{125}$, le **c)** montre que $57 \times 6 = 92 \pmod{125}$ engendre $U(125)$ (dont l'ordre est 100), ce qu'on peut vérifier facilement.

4) a) Raisonner par récurrence.

b, c) Utiliser le a).

d) D'après ce qui précède, -1 et 5 engendrent des groupes d'ordre 2^{k-2} et 2 dont l'intersection est réduite à 1 et dont le produit des ordres est l'ordre de $U(n)$. L'exercice 18.9, 4), donne l'isomorphisme.

5) a) $U(200) \simeq U(8) \times U(25) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/20\mathbb{Z}$.

b) $\text{Aut}(\mathbb{Z}/64\mathbb{Z}) \simeq U(64) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/16\mathbb{Z}$.

c) Le groupe $U(n)$ est cyclique si $n = p^k$ ou $n = 2p^k$ avec $p > 2$ premier et k entier ou encore si $n = 4$; sinon, en décomposant n en produit de facteurs premiers, on voit que $U(n)$ est un produit de plusieurs groupes cycliques d'ordres pairs, produit qui ne peut être cyclique.

d) Si deux groupes G et H ont des éléments d'ordre 2 , x et y respectivement, le groupe produit $G \times H$ a les éléments d'ordre 2 suivants : (x, e_H) , (e_G, y) , (x, y) . Comme $U(n)$ est produit de groupes cycliques ayant chacun un élément d'ordre 2 , on se convainc que n doit être égal soit à une puissance, soit 2 fois une puissance d'un nombre premier supérieur ou égal à 3 , soit égal à 4 .

21.6 1) $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, $F_4 = 65537$, $F_5 = 4294967297$.

2) a), b), et c) Utiliser $a^2 - b^2 = (a - b)(a + b)$ et une récurrence sur k pour montrer l'égalité. Si p premier, $p > 2$, divise F_n , il divise donc $F_{n+k} - 2$ et ne peut diviser F_{n+k} . Comme la suite des nombres de Fermat est infinie et que chacun d'entre eux a au moins un diviseur premier, la suite des nombres premiers est infinie. Enfin, la formule du a) montre que $F_{2+k} \equiv 2 \pmod{5}$ pour tout $k \geq 0$. Comme F_{2+k} est impair, c'est donc que son dernier chiffre est 7 .

3) a) Comme $2^{2^n} \equiv -1 \pmod{p}$, 2 est d'ordre 2^{n+1} dans $U(p)$; 2^{n+1} divise donc $p - 1$, l'ordre de $U(p)$, ce qui montre que $p \equiv 1 \pmod{2^{n+1}}$. Si $n \geq 2$, on a donc $p \equiv 1 \pmod{8}$ et la loi de réciprocité quadratique montre que 2 est un carré dans $U(p)$. Si $2 \equiv a^2 \pmod{p}$, on a donc $a^{2^{n+1}} \equiv -1 \pmod{p}$ dont on déduit comme précédemment que $p \equiv 1 \pmod{2^{n+2}}$.

b) Si n n'est pas un nombre premier, on sait qu'il admet un diviseur premier p tel que $p \leq \sqrt{n}$. Si F_3 n'est pas premier, il admet donc un diviseur premier inférieur à 16 et tel que $p \equiv 1 \pmod{32}$; c'est impossible, donc $F_3 = 257$ est premier, ce qu'on pouvait facilement vérifier directement. Si F_4 n'est pas premier, il admet un diviseur premier inférieur à 256 et tel que $p \equiv 1 \pmod{64}$; 65 et 129 ne sont pas premiers, 193 est premier mais ne divise pas F_4 ; par conséquent, F_4 est premier. Pour montrer que F_5 n'est pas premier, il suffit de trouver un diviseur premier inférieur à $65\,536$ et tel que $p \equiv 1 \pmod{128}$; 129 n'est pas premier, $257 = F_3$ est premier mais ne divise pas F_5 , 385 et 513 ne sont pas premiers, 641 est premier et divise F_5 car $4294967297 = 641 \times 6700417$.

4) a) Si a non nul est un carré modulo F_n , il existe b tel que $a = b^2 \pmod{F_n}$; on a $a^{(F_n-1)/2} = b^{F_n-1} = 1 \pmod{F_n}$, donc a est d'ordre strictement inférieur à $F_n - 1$, l'ordre de $U(F_n)$.

b) Comme $F_n - 1$ est une puissance de 2 et que $a^{F_n-1} = -1 \pmod{F_n}$, a est d'ordre une puissance de 2. Comme a n'est pas un résidu quadratique modulo $F_n - 1$, $a^{(F_n-1)/2} = -1 \pmod{F_n}$; l'ordre de a ne peut être que $F_n - 1$, donc a engendre $U(F_n)$.

c) Il suffit, d'après le **b)**, de montrer que 3, 5, 7 ne sont pas résidus quadratiques modulo F_n . On vérifie que $F_n = 2 \pmod{3}$ par un calcul simple et on a

$$\left(\frac{3}{F_n}\right) = \left(\frac{F_n}{3}\right) = \left(\frac{2}{3}\right) = -1.$$

On procède de même pour 5; pour 7, on remarque que $F_n = 2^{3k+1} + 1$, donc $F_n = 3 \pmod{7}$ ou $F_n = 5 \pmod{7}$; comme 3 et 5 ne sont pas des carrés modulo 7, on en déduit

$$\left(\frac{7}{F_n}\right) = \left(\frac{F_n}{7}\right) = \left(\frac{3}{7}\right) = -1 \pmod{7} \text{ ou } \left(\frac{7}{F_n}\right) = \left(\frac{5}{7}\right) = -1 \pmod{7}.$$

5) Si F_n est premier, on applique le **4)**. Si $3^{(F_n-1)/2} = -1 \pmod{F_n}$, 3 est dans $U(F_n)$ et son ordre est $F_n - 1$, ce qui montre que tout élément non nul de $\mathbb{Z}/F_n\mathbb{Z}$ est inversible, donc que $\mathbb{Z}/F_n\mathbb{Z}$ est un corps; d'où le résultat. Au lieu de 3, on pourrait choisir d'autres entiers qui sont pas résidus quadratiques modulo F_n .

21.7 1) Les carrés et les non carrés sont aussi nombreux dans $(\mathbb{Z}/p\mathbb{Z})^*$ puisque $x^2 = y^2$ implique $y = \pm x$. Si c n'est pas un carré, on a $c^{(p-1)/2} = -1 \pmod{p}$ et $d = c^{(p-1)/4}$ vérifie $d^2 = -1 \pmod{p}$; on peut choisir d tel que $1 < d < p - 1$; il existe alors k , $k < p$, tel que $d^2 + 1 = kp$.

2) On a $w^2 + x^2 = u^2 + v^2 = 0 \pmod{k}$, donc $w^2 + x^2 = k_1 k$, et le choix de w et x montre que $w^2 + x^2 \leq k^2/2$, donc $1 \leq k_1 \leq k/2$.

On a $(u^2 + v^2)(w^2 + x^2) = k^2 k_1 p$; d'autre part, $(u^2 + v^2)(w^2 + x^2) = (uw + vx)^2 + (ux - vw)^2$ et $uw + vx = 0 \pmod{k}$, $ux - vw = 0 \pmod{k}$. Par conséquent, $y = \frac{uw + vx}{k}$ et $z = \frac{ux - vw}{k}$ sont des entiers qui répondent à la question.

3) Cas de 157. L'égalité $28^2 + 1 = 5 \times 157$ donne $u = 28$, $v = 1$, $k = 5$, d'où $w = -2$, $x = 1$, $k_1 = 1$, $y = -11$, $z = 6$ et $11^2 + 6^2 = 157$.

Dans le cas de 29, on trouve $u = 12$, $v = 1$, $w = 2$, $x = 1$, $y = 5$, $z = 2$ et $5^2 + 2^2 = 29$.

4) Supposons qu'il existe p divisant n tel que $p = 3 \pmod{4}$ et $v_p(n)$ impair. Si $n = x^2 + y^2$, on peut se ramener, en divisant par $\text{pgcd}(x, y)$ au cas où x et y sont premiers entre eux. Alors $x^2 + y^2 = 0 \pmod{p}$ et $x, y \neq 0 \pmod{p}$. En divisant par x^2 , on voit que -1 est un carré modulo p , ce qui est faux.

21.8 1) Les implications de **a)** et **b)** s'obtiennent en considérant la relation d'équivalence $\mathcal{R}$ définie par $x\mathcal{R}y$ si et seulement si $y = f(x)$ (et alors $x = f(y)$) et en remarquant que la classe d'équivalence de $x \in E$ a deux éléments si et seulement si x n'est pas un point fixe de f .

2) Comme p est premier, les points de S ne peuvent vérifier l'une des égalités $x = y - z$, ou $x = 2y$.

3) Remarquer que $f(A) \subset C, f(B) \subset B, f(C) \subset A$ et conduire les calculs, très simples, dans chaque cas.

4) a) Les points fixes de f sont dans B ; on trouve qu'ils sont de la forme (x, x, z) ; comme on a $x^2 + 4xz = p$ et que $p = 4k + 1$ avec $k \geq 1$, le seul point fixe de f est $(1, 1, k)$. Le cardinal de S est impair.

b) Il est clair que g est une involution de S . Comme le cardinal de S est impair, g possède un unique point fixe, de la forme (x, y, y) , donc tel que $p = x^2 + 4y^2$, ce qui permet de conclure.

CODES CORRECTEURS ET CRYPTOGRAPHIE

22

Je propose dans ce chapitre d'aborder deux domaines où les mathématiques sont essentielles, la théorie des codes correcteurs d'erreurs et la cryptographie, des théories dont l'importance ne cesse de croître et dont les enseignements se multiplient¹. La théorie des corps finis présentée dans le chapitre précédent, née avec Galois, y a trouvé (ainsi que la théorie des courbes algébriques sur ces corps), des applications magnifiques. Notre vie de tous les jours dépend maintenant de l'usage de ces théories mathématiques défrichées par la curiosité et le génie de quelques-un(e)s, sans financements, sans idée d'applications à un âge où les ordinateurs n'existaient pas.

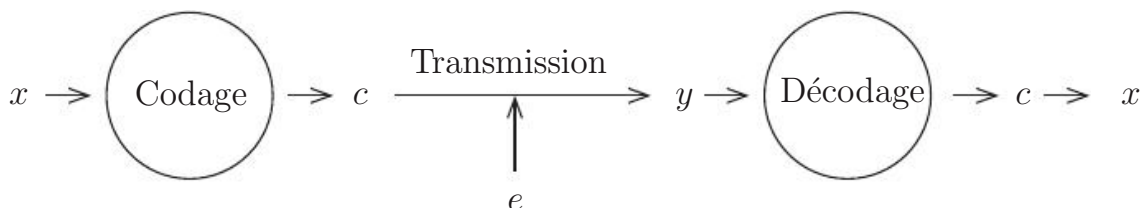
22.1 LES DÉBUTS DES CODES CORRECTEURS

Un peu plus de trente ans après le code de Morse, Émile Baudot (1845-1903) invente au début des années 1870 des codes binaires pour les lettres et les chiffres. L'idée de numériser des sons date de 1937. Puis on cherche à numériser par des suites de 0 et de 1 des données plus complexes : courbes, paroles, photos, séquences de films... En 1943, John Tukey (1915-2000) propose d'abréger l'expression *binary digit* (chiffres binaires : 0 ou 1) en *bit*.

Le problème est celui de la transmission des suites de ces bits : il y a ce qu'on appelle du bruit, de la friture diraient certains, des bits sont modifiés, d'autres perdus. Pour qu'on puisse reconstituer intégralement le message, il faut envoyer des informations supplémentaires en nombre suffisant, comme lorsque vous devez vous répéter plusieurs fois pour vous faire comprendre d'une personne loin de vous alors que la radio braille, qu'un chien aboie et qu'une partie seulement de ce que vous dites est

1. Merci à Delphine Boucher pour ses relectures de ce chapitre.

perçue à chaque fois. Le schéma classique, qui remonte au moins à Shannon, est le suivant².



Le schéma décrit l'envoi d'un message x : on le code en c en ajoutant de l'information ; quand on transmet c , des erreurs e peuvent transformer c en y ; on cherche à décoder y pour retrouver c , dont on déduit x .

Les premiers problèmes des années 1940 ont conduit au développement de la théorie des codes correcteurs d'erreurs (il ne s'agit ni de codes secrets ni de codes pour compresser des données, mais nous parlerons simplement de codes dans la suite) avec :

- Claude Shannon (1916-2001) qui jette les bases de la théorie de l'information dans le fameux article de 1948 : « *A Mathematical Theory of Communication* » où il introduit la notion d'entropie ;
- Marcel Golay (1902-1989), créateur de deux superbes codes parfaits dans un article célèbre, d'une seule page, de 1949 ;
- Richard Hamming (1915-1998) qui publie en avril 1950 l'article fondateur de la théorie : « *Error detecting and error correcting codes* » dans *The Bell Technical Journal* ;
- les codes de Reed-Müller (1954), Reed-Solomon (1960), les codes BCH (1959-1960), les codes de Goppa (1981), etc³.

Les situations se sont diversifiées depuis et on cherche aujourd'hui les moyens de transmettre parfaitement des signaux radio, des conversations téléphoniques, des messages et données par Internet, des émissions de radio ou de télévision, des messages dans les mémoires d'ordinateur...

Pour les enregistrements musicaux, les CD sont mis au point par Philips et Sony autour de 1980 ; c'est un faisceau de lumière qui lit des informations binaires gravées sur la surface du CD ; celles-ci sont traitées avec un code de Reed-Solomon qui peut corriger des paquets (jusqu'à plusieurs milliers) d'erreurs dues à des marques de doigt ou des rayures légères.

Les codes correcteurs sont aussi utiles pour la transmission des données des sondes spatiales envoyées lentement du fond glacé mais pas vraiment silencieux⁴ de

2. Le texte de Michel Coste, Annette Paugam et Ronan Quarez sur les codes correcteurs, publié sur le site très riche <http://agreg-maths.univ-rennes1.fr/> de préparation à l'agrégation de mathématiques de Rennes, m'a été très utile.

3. Irving Reed (1923-2012), David Müller (1924-2008), Gustave Solomon (1930-1996), Valeri Goppa (né en 1939).

4. Les perturbations sont causées par le vent solaire, le passage dans les couches de l'atmosphère, les ondes de toutes sortes saturant l'espace proche...

l'espace infini vers de grandes antennes sur la Terre. En 1971, la mission Mariner 9 vers Mars a utilisé un code de Reed-Müller ; la mission Voyager de 1979 vers Jupiter et Saturne a utilisé un code de Golay pour transmettre des photos avec 4 096 nuances de couleur.

22.2 CODES CORRECTEURS

Le codage

On pose $A = \{0, 1\}$; A est l'alphabet utilisé (la généralisation de ce qui va suivre au cas où les éléments de A sont ceux d'un corps fini plus gros est importante pour les applications actuelles). On suppose que ce qu'on veut transmettre est découpé en suites de 0 et de 1 de longueur fixée k qu'on appelle messages ; les messages sont donc des suites $x = (x_1, \dots, x_k)$ de A^k .

Pour coder les suites de A^k , on se donne un entier $r > 0$ appelé redondance et une injection $\gamma : A^k \rightarrow A^n$, avec $n = k + r$. Pour transmettre un message x , on transmet $c = \gamma(x) = (c_1, \dots, c_n)$. On code donc une suite de k éléments de A par une suite plus longue de n éléments de A . Que γ soit injectif et que r soit strictement positif sont, bien sûr, des nécessités. On pose $\gamma(A^k) = C$. Ces données définissent ce qu'on appelle un code. L'ensemble C est un sous-ensemble de A^n ; ses 2^k éléments sont les mots du code ; ce sont des suites de longueur n et on dit que le code est de longueur n . Le rationnel $\frac{k}{n}$ est appelé taux d'information du code.

Souvent, la suite des k premiers bits de $\gamma(x)$ est la suite des bits de x , autrement dit, $\gamma(x)$ est formé des bits de x suivis de r bits supplémentaires appelés bits de redondance. On parle dans ce cas de codage systématique.

Le décodage

Quand on reçoit un message, la probabilité que beaucoup de bits aient été changés n'est pas nulle et il est donc possible qu'on interprète de façon erronée ce qu'on reçoit. Dans ce qui suit, nous oublierons cet aspect probabiliste et fixerons le nombre d'erreurs maximal pendant la transmission dans chaque cas.

On reçoit donc un message $y = (y_1, \dots, y_n)$. Si $y \in C$, on pose $c = y$ et on retrouve x . Si $y \notin C$, il faut chercher le mot de C qui peut donner y .

Le code

L'objet important de la théorie des codes est le sous-ensemble C de A^n . On doit ajouter une structure sur A^n pour le définir d'une manière agréable ; plusieurs solutions ont été proposées. Il doit avoir de bonnes propriétés comme sous-objet de A^n : ses éléments doivent être suffisamment bien disposés dans A^n afin de permettre la détection et la correction d'erreurs avec un minimum de redondance. La structure

choisie pour A^n doit aussi permettre un décodage sans calculs trop lourds. Enfin, le choix du code dépend aussi de l'usage qu'on va en faire : on ne choisit pas les mêmes codes si les erreurs sont bien réparties ou si elles arrivent par paquets.

22.3 EXEMPLES

Code de répétition

Ce premier exemple est extrêmement simple. Le code consiste à répéter un nombre impair de fois, disons cinq, chaque bit. On convient que le bit correct est celui reçu majoritairement (au moins trois fois) ; si je reçois $y = (1, 0, 1, 1, 0)$, je dois le comprendre comme $c = (1, 1, 1, 1, 1)$ et pose $x = 1$, si je reçois $y = (0, 0, 1, 1, 0)$, je dois le comprendre comme $c = (0, 0, 0, 0, 0)$ et pose $x = 0$. Cette méthode n'est pas satisfaisante car elle demande beaucoup de temps de transmission.

Ajout d'un bit de parité

Cette méthode est intéressante si les erreurs de transmission sont rares, n'apparaissent qu'au plus une fois par mot transmis et que vous avez la possibilité de répéter vos envois. Vous pouvez ajouter à la suite des k bits un bit dit de parité, égal modulo 2 à la somme des k bits précédents de la suite.

Par exemple, vous complétez la suite :

$$x = (1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1)$$

de quinze bits par le bit 1 pour former un mot c du code. Votre correspondant peut recevoir une suite comportant une erreur sur les bits de c , comme :

$$y_1 = (1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1),$$

$$y_2 = (1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0).$$

Dans les deux cas, il constate que le bit de parité ne correspond pas avec les quinze bits précédents et vous demande de renvoyer le message (c'est ce qu'utilise le protocole TCP, Transmission Control Protocol, sur Internet), ce qui, si c'est exceptionnel, ne coûte pas trop cher. Le taux d'information est $k/k+1$.

Code corrigeant une erreur

Dans ce troisième exemple, la valeur de k est le carré d'un entier s : $k = s^2$; on prend $n = (s + 1)^2$, donc $r = 2s + 1$. Les s^2 bits du message x peuvent être considérés comme les éléments d'une matrice carrée $x = (x_{ij})$ d'ordre s . Les $2s + 1$ bits de redondance sont les bits de parité des s lignes, des s colonnes et le bit de parité du tableau entier ; ils forment avec x , en ajoutant une ligne, une colonne et un élément diagonal, une matrice carrée $\gamma(x)$ d'ordre $s + 1$. Si une erreur de transmission se produit pour un élément de $\gamma(x)$, elle est localisée par la ligne et la colonne de l'élément mal transmis. Par conséquent, ce code permet de corriger une erreur de transmission (mais il ne peut en corriger deux, car deux erreurs en (i, j) et (l, m) peuvent être aussi comprises comme des erreurs en (i, m) et (l, j)).

22.4 DISTANCE DE HAMMING

Distance de Hamming et distance minimale d'un code

Soit n un entier. Le poids $\pi(y)$ d'un élément $y = (y_1, \dots, y_n)$ de A^n est le nombre de 1 de la suite. La distance de Hamming sur A^n est la distance d_H définie en posant $d_H(y, y') = \pi(y - y')$ (rappelons que les signes $+$ et $-$ sont interchangeables en caractéristique 2) ; on voit que $d_H(y, y')$ est égal au nombre de différences des suites y et y' de A^n : si $y = (y_1, \dots, y_n)$ et $y' = (y'_1, \dots, y'_n)$, alors

$$d_H(y, y') = \left| \{i, 1 \leq i \leq n \text{ et } y_i \neq y'_i\} \right|.$$

L'inégalité triangulaire est facile à vérifier.

La distance minimale d d'un code C est le minimum des distances de Hamming des éléments distincts de C :

$$d = \inf_{c, c' \in C, c \neq c'} d_H(c, c').$$

Un $[n, k, d]$ code est un code dont la distance minimale est d ; $[n, k, d]$ est le type du code.

Nombre d'erreurs détectées, nombre d'erreurs corrigées

Considérons un tel code C .

Si $d = 1$, il existe deux éléments $c, c' \in C$ tels que $d_H(c, c') = 1$. Quand on envoie le mot c , une erreur dans la transmission peut le transformer en c' , sans qu'on soit capable de s'en apercevoir. Un tel code n'a aucun intérêt.

Si $d = 2$, une erreur de transmission change un mot de C en mot qui n'est pas dans C ; on peut donc détecter une erreur de transmission. Comme $d = 2$, il existe deux éléments $c, c' \in C$ tels que $d_H(c, c') = 2$; on peut former un mot $y \in A^n$ tel que $d_H(y, c) = d_H(y, c') = 1$. Si on reçoit y , on ne peut donc décider si c'est c ou c' qui a été envoyé ; le code C ne permet pas de corriger une erreur de transmission.

Si $d = 3$, on vient de voir qu'on peut détecter une erreur de transmission. Si une erreur de transmission change un mot $c \in C$ en un mot $y \in A^n$, on a $d_H(y, c) = 1$; pour tout mot $c' \neq c$ de C , on a $d_H(c', y) \geq 2$. Comme c est le seul mot de C à distance 1 de y , on peut retrouver c ; le code C permet de corriger une erreur de transmission.

On vérifie facilement que :

- un $[n, k, d]$ code permet de détecter jusqu'à $d - 1$ erreurs de transmission ;
- un $[n, k, d]$ code permet de corriger jusqu'à $\lfloor (d - 1)/2 \rfloor$ (les crochets notent la partie entière) erreurs de transmissions ; en effet, on vérifie que si d est pair, $d = 2e$, alors C permet de corriger jusqu'à $e - 1$ erreurs et que, si d est impair, $d = 2e + 1$, alors C permet de corriger jusqu'à e erreurs.

Codes parfaits

Pour tout $y \in A^n$, la boule $B(y, m)$, m entier, contient, pour tout $i \leq m$, C_n^i éléments à la distance i de y ; son cardinal est donc $1 + C_n^1 + \dots + C_n^m$.

Soit C un $[n, k, d]$ code. Il corrige au plus e erreurs (on vient de voir comment trouver e). Comme les 2^k boules $B(c, e)$ pour $c \in C$, sont disjointes dans A^n , ensemble de cardinal $2^n = 2^{k+r}$, le cardinal de chacune d'entre elles est inférieur ou égal à 2^r , donc :

$$1 + C_n^1 + \dots + C_n^e \leq 2^r.$$

Un $[n, k, d]$ code est dit parfait si A^n est la réunion disjointe des boules de rayon e centrées en les éléments de C , autrement dit si l'inégalité précédente est une égalité.

22.5 CODES LINÉAIRES

C comme sous-espace vectoriel

On munit A^n de sa structure naturelle d'espace vectoriel sur $\mathbb{F}_2$: A^n est isomorphe à $\mathbb{F}_2^n$ et on prend pour C un sous-espace vectoriel de dimension k de A^n .

Le codage $\gamma : A^k \rightarrow A^n$ est une application linéaire injective qu'on peut définir de plusieurs manières. La matrice G de γ par rapport aux bases canoniques de A^k et A^n est appelée matrice génératrice du code. Ses colonnes sont les images dans A^n des vecteurs de la base canonique de A^k et forment une base de C .

Comme $d_H(c, c') = \pi(c - c')$ et que $c - c' \in C$, la distance minimale d d'un code linéaire est égale au poids minimum d'un mot non nul de C .

Notons E le sous-espace vectoriel de A^n des suites commençant par $k - 1$ bits 0. On a $\dim(E) = n - k + 1$, donc $\dim C + \dim(E) > \dim A^n$. L'intersection de C et E n'est pas réduite à 0, par conséquent, il existe dans C des éléments de poids inférieur ou égal à $n + 1 - k$, d'où :

$$d \leq n - k + 1 = r + 1.$$

Cette majoration est naturelle ; elle signifie qu'il faut faire un effort sur la redondance r pour pouvoir corriger plus d'erreurs.

Analyse du message reçu

Le décodage des codes linéaires n'est pas, en général, un problème facile. On utilise les formes linéaires $A^n \rightarrow A$ s'annulant sur C . On sait qu'elles forment un espace vectoriel $C^\perp$ qu'on a appelé, au chapitre 10, dual de l'espace vectoriel C ; $C^\perp$ est de dimension $n - k$. On a vu aussi que les éléments de C sont exactement les éléments de A^n annulés par les formes linéaires de $C^\perp$; par conséquent, C est le noyau d'une application linéaire $\eta : A^n \rightarrow A^{n-k}$ ayant pour composantes les

éléments d'une base de $C^\perp$. La matrice H d'une telle application est appelée matrice de contrôle ; elle est nulle pour tous les vecteurs du code et on a $HG = 0$.

Pour faire parvenir un message x , on envoie $c = \gamma(x)$; si y est le message reçu, l'erreur est la suite $e = y - c$ (on a aussi $e = y + c$ dans A^n) et on a $\eta(e) = \eta(y)$. C'est à partir de $\eta(e)$ qu'il faut retrouver e ; on peut alors trouver $c = y + e$, puis le message initial $x = \gamma^{-1}(c)$. Si $\eta(y) = 0$, on a $y \in C$, on pose $c = y$ et on retrouve x .

Calcul matriciel

Pour utiliser un code linéaire, le calcul matriciel s'impose. Nous adoptons pour cette section la notation suivante : si on a une suite de bits $b = (b_1, \dots, b_m)$, on note avec un prime la matrice colonne associée : $b' = {}^t(b_1 \dots b_m)$ dont les coefficients sont les b_i . Pour envoyer le message $x = (x_1, \dots, x_k)$ de A^k , on transmet le mot du code $c = \gamma(x)$ tel que $c' = Gx'$.

Notons que l'usage est d'écrire, en théorie des codes, les vecteurs en lignes, ce qui conduit à transposer toutes les égalités matricielles qui vont suivre et oblige à une petite gymnastique que je ne vous impose pas.

Code de Hamming (7,4,3)

C'est en 1948, dans son grand article, que Shannon présente le code correcteur de Hamming. On peut s'en servir quand on suppose que la transmission peut créer au plus une erreur. Dans ce code, un message est une suite de 4 bits $x = (x_1, x_2, x_3, x_4) \in A^4$ qu'on code en $c = \gamma(x) = (u, v, x_1, w, x_2, x_3, x_4) \in A^7$, avec :

$$u = x_1 + x_2 + x_4,$$

$$v = x_1 + x_3 + x_4,$$

$$w = x_2 + x_3 + x_4.$$

La matrice de contrôle H est donnée par ces trois égalités ($w + x_2 + x_3 + x_4 = 0$ donne la première ligne, etc.) et la matrice génératrice G a pour colonnes les images par γ des vecteurs de la base canonique, vecteurs formant une base du code C :

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Les seize mots de C s'obtiennent par combinaisons linéaires des vecteurs de cette base. Les voici.

(0,0,0,0,0,0,0) (1,1,1,0,0,0,0) (1,0,0,1,1,0,0) (0,1,0,1,0,1,0)
 (1,1,0,1,0,0,1) (0,1,1,1,1,0,0) (1,0,1,1,0,1,0) (0,0,1,1,0,0,1)
 (1,1,0,0,1,1,0) (0,1,0,0,1,0,1) (1,0,0,0,0,1,1) (0,0,1,0,1,1,0)
 (1,0,1,0,1,0,1) (0,1,1,0,0,1,1) (0,0,0,1,1,1,1) (1,1,1,1,1,1,1)

On peut vérifier que la distance de Hamming de ce code est 3 et que le code est parfait : A^7 a $2^7 = 128$ éléments qui se répartissent dans 16 boules de rayon 1 ayant 8 éléments. Ce code permet donc de corriger une erreur.

Quand on reçoit la suite $y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7)$ de sept bits, on calcule Hy' . Si $Hy' = 0$, le message a été transmis sans erreur et on pose $c = y$. Si $Hy' \neq 0$, c'est qu'une erreur de transmission e a eu lieu et il existe $c \in C$ tel que $d_H(y, c) = 1$ et $c = y + e$. Hamming explique que les colonnes de H sont les écritures en binaire des nombres de 1 à 7. Comme la suite e ne comporte qu'un seul terme non nul, $He' = Hy'$ est la colonne de H correspondant à la position de l'erreur et on trouve e . C'est à partir de cette propriété que le code a été construit.

Par exemple, si on veut envoyer $x = (1, 0, 1, 1)$, on envoie le message codé $Gx' = c'$, avec $c = (0, 1, 1, 0, 0, 1, 1)$. Si on reçoit le message $y = (0, 1, 1, 0, 0, 0, 1)$

avec une erreur $e = (0, 0, 0, 0, 0, 1, 0)$, au rang 6, on a $Hy' = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, la sixième

colonne de H , ce qui indique bien le rang 6. On peut considérer que c'est $x = (1, 0, 1, 1)$ qui a été envoyé.

Code de Hamming avec bit de parité

Hamming propose aussi une amélioration de son code en ajoutant un bit de parité $t = u + v + x_1 + x_2 + w + x_3 + x_4$, ce qui donne

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Le bit de parité est placé au début du message transmis et on peut lui attribuer le rang 0. Le message reçu y est dans C si $Hy' = 0$. S'il y a eu une erreur de transmission e , on la trouve à partir de $Hy' = He'$. Par exemple, si on envoie $(1, 0, 1, 1)$ et que $(1, 0, 1, 1, 0, 0, 1, 1)$ est reçu, Hy' donne la suite $(1, 0, 0, 0)$, ce qui indique que c'est le bit de parité qui est changé. Si y comporte deux erreurs e_1 et e_2 , on a

$Hy' = He'_1 + He'_1$; ce vecteur est la somme de deux colonnes de H ; cette somme n'est, ni nulle, ni égale à une colonne de H , ce qui permet de détecter la présence de deux erreurs ; mais on trouve facilement deux couples de colonnes de H de même somme et on ne peut déterminer sur quel couple portent les erreurs, ce qui montre qu'on ne peut corriger deux erreurs avec ce code.

22.6 CODES CYCLIQUES

Définitions

On suppose que n est un entier premier avec 2, c'est-à-dire impair, dans tout ce qui suit. On définit $\sigma : A^n \rightarrow A^n$ par $\sigma(y_1, \dots, y_n) = (y_2, \dots, y_n, y_1)$; on appelle σ un décalage cyclique ; σ est un automorphisme de l'espace vectoriel A^n et $\sigma^n = \text{id}_{A^n}$.

Un code linéaire de longueur n et de dimension k est cyclique si le sous-espace vectoriel C de A^n est stable par σ , c'est-à-dire si $\sigma(C) = C$.

Le code de Hamming vu ci-dessus peut-être vu comme un code cyclique (voir exercice 22.1.3).

Codes cycliques et polynômes

Mettons encore plus de structure sur A^n . Considérons l'algèbre $\mathbb{F}_2[T]/(T^n - 1)$ quotient de l'anneau de polynômes $\mathbb{F}_2[T]$ par l'idéal engendré par le polynôme $T^n - 1$. Comme on l'a vu au chapitre 20, si on note t la classe de T dans $\mathbb{F}_2[T]/(T^n - 1)$, les éléments du quotient sont des combinaisons linéaires $\sum_{1 \leq i \leq n} y_i t^{n-i}$. Les calculs dans ce quotient se font en utilisant l'égalité $t^n = 1$. Si on se souvient qu'un polynôme n'est pas autre chose que la suite de ses coefficients, on peut identifier les suites de A^n avec des polynômes de $\mathbb{F}_2[T]$ de degré strictement inférieur à n et aussi avec les éléments de l'algèbre quotient $\mathbb{F}_2[T]/(T^n - 1)$, la suite $(y_1, \dots, y_n)$ correspondant au polynôme $\sum_{1 \leq i \leq n} y_i T^{n-i}$ et à $\sum_{1 \leq i \leq n} y_i t^{n-i}$.

Avec cette identification, on vérifie que le décalage σ sur les suites de A^n correspond à la multiplication par t dans $\mathbb{F}_2[T]/(T^n - 1)$ puisque, comme $t^n = 1$, on a

$$t \left(\sum_{1 \leq i \leq n} y_i t^{n-i} \right) = y_1 t^n + \sum_{2 \leq i \leq n} y_i t^{n-i+1} = \sum_{2 \leq i \leq n} y_i t^{n-i+1} + y_1,$$

ce dernier élément correspondant à la suite $(y_2, \dots, y_n, y_1)$.

Codes cycliques et idéaux

Quand on identifie A^n et $\mathbb{F}_2[T]/(T^n - 1)$, on voit qu'un code cyclique $C \subset A^n$ s'identifie à une partie, encore notée C , de $\mathbb{F}_2[T]/(T^n - 1)$ qui est stable par la multiplication par t . Par conséquent, C est stable par la multiplication par toute expression polynomiale en t , c'est-à-dire par tout élément de $\mathbb{F}_2[T]/(T^n - 1)$. On

en déduit que C est un idéal de $\mathbb{F}_2[T]/(T^n - 1)$. On a vu en 19.5 que les idéaux du quotient d'un anneau A par un idéal I correspondent aux idéaux de A contenant I . Par conséquent, C correspond à un idéal de $\mathbb{F}_2[T]$ contenant l'idéal $(T^n - 1)$; mais un tel idéal est défini, comme $\mathbb{F}_2[T]$ est principal, par un diviseur G de $T^n - 1$ dans $\mathbb{F}_2[T]$; comme le coefficient dominant de G est 1, G est unitaire et donc unique. Le polynôme G est appelé polynôme générateur de C .

Réciproquement, étant donné un diviseur G de $T^n - 1$, on définit un code cyclique C en considérant l'idéal engendré par la classe g de G dans le quotient $\mathbb{F}_2[T]/(T^n - 1)$. Le code C consiste en les produits gq où q est la classe d'un polynôme Q tel que $\deg(Q) \leq n - 1 - \deg(G)$. Le code C est donc engendré, comme espace vectoriel, par les éléments $g, tg, \dots, t^{n-1-\deg(G)}g$; si on note k sa dimension, on a $k = n - \deg(G)$.

Codage systématique

Continuons avec les mêmes notations et revenons aux suites de bits pour donner un codage systématique des éléments de A^k . Si $x = (x_1, \dots, x_k)$ est une suite de k bits, on lui associe le polynôme $c_I = x_1 T^{n-1} + \dots + x_k T^{n-k}$ et on effectue la division euclidienne de c_I par G : $c_I = QG + c_R$ où $\deg(c_R) < \deg(G) = n - k$; les lettres I et R indiquent si le polynôme contient l'information où la redondance. La suite $\gamma(x)$ est la suite dont les premiers termes sont ceux de c_I , autrement dits les termes de la suite x , et les derniers termes les coefficients de c_R ; elle correspond à la suite des coefficients de QG ($QG = c_I + c_R$ en base 2) ; on a bien défini un code cyclique.

Correction d'une erreur

Votre correspondant veut vous transmettre un message de k bits. Il calcule les polynômes c_I et c_R et vous transmet la suite $c = \gamma(x)$ définie par ces polynômes. Vous savez qu'il y a au plus une erreur de transmission et vous recevez la suite $y = (y_1, \dots, y_n)$. Vous calculez $\sum_{1 \leq i \leq n} y_i t^{n-i}$. Si vous trouvez 0, c'est que le message codé envoyé est $c = y$; le message original est donné par les k premiers bits de c . Si vous ne trouvez pas 0, c'est qu'une erreur de transmission a eu lieu et il faut retrouver son rang s . Le calcul précédent vous donne ce t^s , mais exprimé dans la base $(1, t, \dots, t^{n-k})$. Il reste à retrouver s .

Exemple

On suppose qu'une erreur de transmission au plus intervient dans une transmission. On prend $n = 7$. La décomposition $T^7 - 1 = (T - 1)(T^3 + T + 1)(T^3 + T^2 + 1)$ conduit à définir le code C de longueur 7 et de dimension 4 associé à $G = T^3 + T + 1$ par la construction précédente. On veut transmettre un message $x = (x_1, x_2, x_3, x_4)$ de A^4 . La division euclidienne de $x_1 T^6 + x_2 T^5 + x_3 T^4 + x_4 T^3$ par G s'écrit $x_1 T^6 + x_2 T^5 + x_3 T^4 + x_4 T^3 = QG + uT^2 + vT + w$. On transmet $\gamma(x) = (x_1, x_2, x_3, x_4, u, v, w)$, un élément de A^7 .

Notons $y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7)$ le message reçu. On calcule $y_1 t^6 + y_2 t^5 + y_3 t^4 + y_4 t^3 + y_5 t^2 + y_6 t + y_7$. Si on trouve 0, c'est que le message envoyé est (y_1, y_2, y_3, y_4) sans correction. Sinon, une erreur unique est intervenue à un certain rang s , $1 \leq s \leq 7$. On a donc $y_1 t^6 + y_2 t^5 + y_3 t^4 + y_4 t^3 + y_5 t^2 + y_6 t + y_7 = t^{7-s}$. On calcule le premier membre dans la base $(t^2, t, 1)$. On dresse la table des puissances de t dans cette base pour trouver s : $t^3 = t + 1$, $t^4 = t^2 + t$, $t^5 = t^2 + t + 1$, $t^6 = t^2 + 1$.

Par exemple, on veut transmettre le message $x = (1, 1, 1, 0)$. On calcule $t^6 + t^5 + t^4$ dans la base $(t^2, t, 1)$; on trouve t^2 , ce qui conduit à envoyer $\gamma(x) = (1, 1, 1, 0, 1, 0, 0)$. Si on reçoit $y = (1, 1, 1, 1, 1, 0, 0)$, on voit que $t^6 + t^5 + t^4 + t^3 + t^2 = t + 1 = t^3$, ce qui conduit à corriger le quatrième terme de y et à retrouver le message envoyé.

Codes cycliques et sous-ensembles de $\mathbb{Z}/n\mathbb{Z}$

Ce qui suit nécessiterait quelques compléments sur les racines de l'unité ; je renvoie à mon livre sur la théorie de Galois pour les lire, en espérant compréhensible sans cela ce qui va suivre. On a vu en 20.2 qu'il existe un corps de décomposition K du polynôme $T^n - 1$ de $\mathbb{F}_2[T]$ et que les racines de $T^n - 1$ sont simples dans K (cette dernière affirmation vient de ce que la dérivée de $T^n - 1$ est nT^{n-1} qui ne s'annule qu'en 0 puisque n est supposé impair et que 0 n'est pas racine de $T^n - 1$). Les racines n -ièmes de l'unité forment un groupe μ_n qui est un sous-groupe de K^* . Puisque K^* est un groupe cyclique (voir 21.1), μ_n est lui-même cyclique ; ses générateurs sont appelés racines primitives n -ièmes de l'unité.

Notons α l'une d'entre elles. Soit G un diviseur de degré m de $T^n - 1$ dans $\mathbb{F}_2[T]$; ses racines sont simples comme celles de $T^n - 1$. Elles forment un ensemble $R = \{\alpha^{i_1}, \dots, \alpha^{i_m}\}$. On peut considérer l'ensemble $\Sigma = \{i_1, \dots, i_m\}$ des exposants de α comme une partie de $\mathbb{Z}/n\mathbb{Z}$. Montrons que, si $G \in \mathbb{F}_2[T]$, R est stable par élévation au carré et Σ par multiplication par 2. Posons $G(T) = z_0 T^m + \dots + z_m$; l'homomorphisme de Frobenius (voir 21.1) donne $(G(T))^2 = z_0^2 T^{2m} + \dots + z_m^2$. Comme les coefficients de G sont dans $\mathbb{F}_2$, ils sont égaux à leur carré, donc $(G(T))^2 = G(T^2)$. Si $\beta \in \Sigma$, on a $G(\beta^2) = (G(\beta))^2 = 0$, ce qui montre la stabilité de R par élévation au carré et celle de Σ par multiplication par 2.

Montrons la réciproque. Supposons que R soit stable par élévation au carré (ou que Σ soit stable par multiplication par 2). L'application $\beta \mapsto \beta^2$ de R dans lui-même étant injective (ne pas oublier qu'on raisonne en caractéristique 2) est une bijection. On vérifie que les polynômes $(G(T))^2$ et $G(T^2)$ ont tous deux les éléments de R comme racines doubles. Ils sont donc égaux, donc leurs coefficients sont égaux et vérifient $z^2 = z$; par conséquent, on a $z = 0$ ou $z = 1$, ce qui montre que $G \in \mathbb{F}_2[T]$.

Récapitulation

On a vu qu'un code cyclique est un idéal de $\mathbb{F}_2[T]$ et que cet idéal correspond à un diviseur de $T^n - 1$ dans $\mathbb{F}_2[T]$. On vient de voir que ce diviseur correspond à un ensemble de puissances de α stable par élévation au carré et enfin que l'ensemble des exposants de ces puissances est un sous-ensemble de $\mathbb{Z}/n\mathbb{Z}$ stable par multiplication par 2.

Finalement, on a une bijection entre les sous-ensembles de $\mathbb{Z}/n\mathbb{Z}$ stables par multiplication par 2 et les codes cycliques de longueur n .

Distance minimale d'un code cyclique

Continuons avec les mêmes notations. Montrons que si Σ contient s éléments consécutifs $a + 1, \dots, a + s$ (dans les exemples, on aura souvent $a = 0$), la distance minimale du code C correspondant, si C est non nul, est supérieure ou égale à $s + 1$.

Notons d la distance minimale du code et supposons $d \leq s$. Il existe un mot c de C de poids $\pi(c) = d$. Le polynôme P de $\mathbb{F}_2[T]$ correspondant à c est de degré strictement inférieur à n et possède un nombre d de coefficients non nuls ; il s'écrit $P = y_1 T^{m_1} + \dots + y_d T^{m_d}$ où $y_1 = \dots = y_d = 1$. Par hypothèse, $P(\alpha^{a+1}) = \dots = P(\alpha^{a+s}) = 0$. Ces équations donnent un système de s relations linéaires :

$$\begin{cases} y_1 \alpha^{(a+1)m_1} + \dots + y_d \alpha^{(a+1)m_d} = 0 \\ y_1 \alpha^{(a+s)m_1} + \dots + y_d \alpha^{(a+s)m_d} = 0 \end{cases}$$

Dans le déterminant des d premières équations de ce système, on peut mettre en facteur les $\alpha^{am_1}, \dots, \alpha^{am_d}$; en posant $\beta_1 = \alpha^{m_1}, \dots, \beta_d = \alpha^{m_d}$, on aboutit à un déterminant de Vandermonde, non nul d'après l'exercice 14.7. Il est donc impossible que tous les $y_i, 1 \leq i \leq d$ soient non nuls ; on en déduit le résultat.

Si Σ est stable par multiplication par 2 et contient $1, \dots, v$ avec $v = 2u - 1$ impair, il contient u donc $v + 1 = 2u$. Par conséquent, Σ contient un ensemble contenant $1, \dots, s$ avec s pair ; la distance minimale du code est donc minorée par le nombre impair $s + 1$.

On peut appliquer ce résultat au cas d'un diviseur irréductible du polynôme cyclotomique Φ_n . Ses racines sont des racines primitives n -ièmes de l'unité, son degré est l'ordre de 2 dans $(\mathbb{Z}/n\mathbb{Z})^*$. Si α est l'une des racines de G , l'ensemble Σ contient α, α^2 , donc le code défini par G permet de corriger une erreur au moins.

22.7 CODES BCH

Présentation

Les codes BCH ont été inventés par Alexis Hocquenghem (1908-1990) en 1959 et, indépendamment, par Raj Chandra Bose (1901-1987), Dwijendra Kumar Ray-

Chaudhuri (né en 1933) en 1960. Des algorithmes de décodage efficaces ont été proposés dans les années suivantes. Le code du Minitel (présenté dans l'exercice 20.2) est un code BCH qui corrige une erreur.

Définition

On se donne un entier n impair de la forme $2^r - 1$. L'ordre de 2 dans $(\mathbb{Z}/n\mathbb{Z})^*$ est r et le corps de décomposition du polynôme $T^n - 1$ de $\mathbb{F}_2[T]$ est $K = \mathbb{F}_{2^r}$ (voir 21.2). On se donne ensuite un entier impair $\delta = 2e + 1$, $3 \leq \delta < n$. On choisit une racine primitive n -ième α dans K , on note Σ le plus petit sous-ensemble de $(\mathbb{Z}/n\mathbb{Z})^*$ contenant $1, 2, \dots, 2e$ et G le polynôme associé à Σ comme ci-dessus.

On définit ainsi un code C qu'on appelle code BCH de longueur n et de distance prescrite (on dit aussi : construite) δ .

La distance minimale de C est supérieure ou égale à sa distance prescrite et le code peut corriger jusqu'à e erreurs et peut-être plus. Un élément de C correspond à un polynôme s'annulant en chacune des puissances α^i , avec $1 \leq i \leq 2e$.

Correction des erreurs 1

La méthode que nous allons décrire est un peu complexe, mais les calculs qu'elle nécessite ne sont pas très gros ; c'est ce qui fait l'intérêt des codes BCH.

On suppose défini un code BCH avec les notations précédentes. On veut envoyer un message x de longueur $k = n - \deg(G)$. On code x en une suite $c = (c_1, \dots, c_n)$ de longueur n correspondant à un polynôme P_0 . On sait que la transmission ne peut introduire plus de e erreurs. Le message reçu $y = (y_1, \dots, y_n)$ correspond à un polynôme P . Les erreurs portent sur e' bits, $0 \leq e' \leq e$ dont les rangs dans l'ordre croissant sont notés $v_1, \dots, v_{e'}$. L'égalité $y = c + e$ donne :

$$P = y_1 T^{n-1} + \dots + y_n = c_1 T^{n-1} + \dots + c_n + T^{n-v_1} + \dots + T^{n-v_{e'}}.$$

On a $P = P_0 + E$ avec le polynôme E correspondant aux erreurs : $E = T^{n-v_1} + \dots + T^{n-v_{e'}}$. Trouver les rangs des erreurs revient à trouver E .

Comme P_0 est un multiple de G , on a $E(\beta) = P(\beta)$ pour toute racine β de G , en particulier $E(\alpha^i) = P(\alpha^i)$ pour $i = 1, \dots, 2e$. Posons $\beta_j = \alpha^{n-v_j}$ pour $j = 1, \dots, e'$. On connaît donc $E(\alpha^i) = \beta_1^i + \dots + \beta_{e'}^i$ pour $i = 1, \dots, 2e$.

Exemple

On prend $n = 15 = 2^4 - 1$. Les éléments non nuls du corps $K = \mathbb{F}_{16}$ forme le groupe des racines quinièmes de l'unité dans ce corps. La décomposition de $T^{15} - 1$ dans $\mathbb{F}_2[T]$ est :

$$(T + 1)(T^2 + T + 1)(T^4 + T^3 + T^2 + T + 1)(T^4 + T^3 + 1)(T^4 + T + 1).$$

On peut obtenir $\mathbb{F}_{16}$ comme quotient de $\mathbb{F}_2[T]$ par l'idéal engendré par l'un des trois polynômes de degré 4 de la factorisation précédente. Si on choisit de faire le quotient par l'idéal engendré par $T^4 + T + 1$, les puissances de la classe t de T s'exprime dans la base $(t^3, t^2, t, 1)$ de la façon suivante.

$$\begin{array}{lll} t & t^6 = t^3 + t^2 & t^{11} = t^3 + t^2 + t \\ t^2 & t^7 = t^3 + t + 1 & t^{12} = t^3 + t^2 + t + 1 \\ t^3 & t^8 = t^2 + 1 & t^{13} = t^3 + t^2 + 1 \\ t^4 = t + 1 & t^9 = t^3 + t & t^{14} = t^3 + 1 \\ t^5 = t^2 + 1 & t^{10} = t^2 + t + 1 & t^{15} = 1 \end{array}$$

On peut poser $\alpha = t$. Pour corriger deux erreurs, on prescrit la distance $\delta = 5$. Parmi les racines de G , on doit avoir t, t^2, t^3, t^4 . On trouve t , donc aussi t^2 et t^4 parmi les racines de $T^4 + T + 1$ dans K . Quant à t^3 , on vérifie que c'est une racine de $T^4 + T^3 + T^2 + T + 1$. On pose donc :

$$G = (T^4 + T^3 + T^2 + T + 1)(T^4 + T + 1) = T^8 + T^7 + T^6 + T^4 + 1.$$

La dimension du code est $k = n - \deg(G) = 7$. Pour coder $x = (x_1, \dots, x_7)$, on fait la division euclidienne du polynôme $c_I = x_1 T^{14} + \dots + x_7 T^8$ par G . Elle s'écrit $c_I = QG + c_R$ avec $c_R = c_8 T^7 + \dots + c_{15}$, donc on envoie la suite $c = (x_1, \dots, x_7, c_8, \dots, c_{15})$.

Par exemple, on veut envoyer $x = (1, 0, 1, 0, 1, 0, 1)$. On a $c_I = T^{14} + T^{12} + T^{10} + T^8$. La division euclidienne donne $Q = T^6 + T^5 + T^4 + T^2 + 1$ et $c_R = T^7 + T^6 + T^5 + T^2 + 1$. On envoie donc la suite $c = (1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1)$.

Supposons qu'on reçoive la suite $y = (1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1)$ avec deux erreurs de transmission aux rangs 3 et 11. Le polynôme P associé à y est $T^{14} + T^{10} + T^8 + T^7 + T^6 + T^5 + T^3 + T^2 + 1$. Le polynôme E associé aux deux termes d'erreurs s'écrit $T^{15-m_1} + T^{15-m_2}$ avec $m_1 < m_2$ qu'on ne connaît pas (on doit trouver $15 - m_1 = 3$ et $15 - m_2 = 11$).

On pose $\beta_1 = t^{15-m_1}, \beta_2 = t^{15-m_2}$. On a, après calculs à l'aide de la table :

$$\beta_1 + \beta_2 = E(t) = P(t) = t^2 + t + 1 ;$$

$$\beta_1^3 + \beta_2^3 = E(t^3) = P(t^3) = t^2 + t.$$

Comme on est en caractéristique 2, les formules $\beta_1^2 + \beta_2^2 = E(t^2) = P(t^2) = (P(t))^2 = t^2 + t$ et $\beta_1^4 + \beta_2^4 = E(t^4) = P(t^4) = (P(t))^4 = t^2 + t + 1$ sont immédiates.

On utilise alors l'identité $a^3 + b^3 = (a + b)(a^2 + ab + b^2) \pmod{2}$; elle donne $t^2 + t = \beta_1^3 + \beta_2^3 = (t^2 + t + 1)(t^2 + t + \beta_1 \beta_2)$, d'où $\beta_1 \beta_2 = 1$.

Connaissant la somme et le produit de β_1 et β_2 , on voit que ce sont les racines du polynôme $S(Z) = Z^2 + (t^2 + t + 1)Z + 1$. La fin manque d'élégance : on trouve les deux racines de S par essais successifs ; on en déduit les rangs des erreurs et on retrouve le message x .

Polynômes symétriques

Terminons cette section sur les codes BCH en donnant une méthode générale pour retrouver les rangs des erreurs. Nous avons besoin de quelques compléments sur les polynômes symétriques étudiés en 20.9.

Nous sommes dans l'anneau $\mathbb{Z}[X_1, \dots, X_n]$ des polynômes à n indéterminées $X_1, \dots, X_n$ et à coefficients dans $\mathbb{Z}$. Pour $k \geq 0$, s_k est le polynôme symétrique élémentaire

$$s_k(X_1, \dots, X_n) = \sum_{H \subset \{1, \dots, n\}, |H|=k} \prod_{i \in H} X_i.$$

Pour $d \geq 0$, p_d est la somme des puissances d -ièmes des indéterminées

$$p_d(X_1, \dots, X_n) = \sum_{1 \leq i \leq n} X_i^d.$$

Nous avons dit que $s_k = 0$ pour $k > n$. Nous avons montré la formule :

$$\prod_{1 \leq i \leq n} (X - X_i) = \sum_{0 \leq k \leq n} (-1)^k s_k X^{n-k}$$

et donné, dans l'exercice 20.12, les formules de Newton qui relient les sommes de puissances p_d aux polynômes s_k :

$$p_d = \sum_{1 \leq k \leq d-1} (-1)^{k-1} s_k p_{d-k} + (-1)^{d-1} d s_d.$$

En caractéristique 2, ces formules donnent :

$$\begin{aligned} \prod_{1 \leq i \leq n} (X + X_i) &= \sum_{0 \leq k \leq n} s_k X^{n-k} \\ p_d &= \sum_{1 \leq k \leq d-1} s_k p_{d-k} + d s_d. \end{aligned}$$

Notons aussi que, pour $d > n$, on a :

$$p_d = \sum_{1 \leq k \leq d-1} s_k p_{d-k}.$$

Pour $l \geq 1$, posons

$$H_l = \begin{pmatrix} p_1 & p_2 & \cdots & p_l \\ p_2 & p_3 & \cdots & p_{l+1} \\ \cdots & \cdots & \cdots & \cdots \\ p_l & p_{l+1} & \cdots & p_{2l-1} \end{pmatrix} \quad V_l = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_l \\ \cdots & \cdots & \cdots & \cdots \\ X_1^{l-1} & X_2^{l-1} & \cdots & X_l^{l-1} \end{pmatrix}$$

et soit D_l la matrice diagonale dont les termes diagonaux sont $X_1, \dots, X_n$.

On vérifie que : $H_l = V_l D_l {}^t V_l$.

Correction des erreurs 2

Revenons à la correction des erreurs.

Posons $S(Z) = (Z - \beta_1) \dots (Z - \beta_{e'})$. En développant ce polynôme, on obtient (puisque l'on est en caractéristique 2) : $S(Z) = Z^{e'} + s_1 Z^{e'-1} + \dots + s_{e'}$, où les s_i sont les valeurs des polynômes symétriques élémentaires en les β_j , $1 \leq j \leq e'$.

On ne connaît pas encore le nombre e' d'erreurs ni leurs rangs, on ne connaît donc pas les $\beta_j = \alpha^{n-v_j}$ ni les s_i . Mais on connaît les $E(\alpha^i)$, donc les sommes $p_i = E(\alpha^i) = \beta_1^i + \dots + \beta_{e'}^i$, pour $i = 1, \dots, 2e'$.

Comme les β_j sont non nuls, $D_{e'}(\beta_1, \dots, \beta_{e'})$ est inversible ; comme les β_j sont distincts, la matrice de Vandermonde $V_{e'}(\beta_1, \dots, \beta_{e'})$ est inversible. Par conséquent, $H_{e'}(\beta_1, \dots, \beta_{e'})$ est inversible.

Considérons maintenant la matrice $H_{e'+1}(\beta_1, \dots, \beta_{e'})$. Sa dernière colonne est combinaison linéaire des e' premières colonnes car les formules de Newton donnent : $p_{e'+m} = s_1 p_{e'+m-1} + s_2 p_{e'+m-2} + \dots + s_{e'} p_m$ pour $m = 1, \dots, e' + 1$. On en déduit que $H_{e'+1}(\beta_1, \dots, \beta_{e'})$ est de rang e' . On vérifie de même que $H_{2e'}(\beta_1, \dots, \beta_{e'})$ est de rang e' .

On vient de trouver un moyen de déterminer e' en calculant le rang d'une matrice qu'on connaît.

L'égalité

$$H_{e'} = (\beta_1, \dots, \beta_{e'}) \begin{pmatrix} s_{e'} \\ \dots \\ s_1 \end{pmatrix} = \begin{pmatrix} p_{e'+1} \\ \dots \\ p_{2e'} \end{pmatrix}$$

permet alors de trouver les s_i ($\beta_1, \dots, \beta_{e'}$), donc le polynôme S .

Après tous ces beaux efforts, la dernière étape du décodage est encore décevante : on essaie les différentes puissances de α pour voir celles qui sont racines de S ... et on a gagné. Voir l'exercice 20.3 pour un exemple.

22.8 HISTOIRE DE LA CRYPTOGRAPHIE

La cryptographie est la science du chiffrement des messages, d'un texte clair en un texte chiffré, et du déchiffrement des messages chiffrés. Pour plus de détails (tout ce que vous avez toujours voulu savoir...) sur l'histoire de la cryptographie, voir Kahn David, *La guerre des codes secrets*, Interéditions, 1980, Simon Singh, *Histoire des codes secrets*, JC Lattès, 1999.

À quoi sert la cryptographie aujourd'hui

L'importance de la cryptographie dans le monde actuel est évidente. Elle est essentielle dans le domaine militaire (liens entre les centres de commandement, les satellites et les troupes en opération, etc.) ainsi que pour les échanges diplomatiques, bancaires et commerciaux. Depuis quelques années, elle sert également pour

chacun d'entre nous, sans qu'on le sache toujours, pour sécuriser les échanges sur la *Toile*, l'usage des cartes bancaires, des téléphones portables, de nos ordinateurs, etc. Elle sert aussi aux trafiquants de drogue, aux mafias, aux terroristes, etc.

Méthodes anciennes

Jusque dans les années 1960, les mathématiques ne sont quasiment pas utilisées en cryptographie.

Voici quelques méthodes anciennes de cryptographie ; jusqu'au XIX^e siècle, la plupart sont en fait très médiocres et ne résistent pas à une analyse un peu serrée.

1) Utiliser une application injective de l'ensemble des lettres et des chiffres dans un ensemble de signes : lettres, chiffres, dessins ; le chiffeur et le déchiffeur doivent disposer tous les deux de l'application, éventuellement, ils doivent se la transmettre, ce que tout ennemi peut exploiter. Le premier à avoir utilisé cette méthode serait Jules César (–100 env. à –44), d'après l'historien romain Suétone (vers 69-126) ; il utilisait un simple décalage de trois lettres, remplaçant le *A* par un *D*, le *B* par un *E*, et ainsi de suite (Juliette appellerait Roméo URPHR avec ce procédé). Un magnifique exemple de déchiffrement d'un texte chiffré avec une bijection de l'ensemble des 26 lettres de l'alphabet sur un ensemble de 26 signes est donné par Edgar Allan Poe (1809-1849) dans sa nouvelle *Le scarabée d'or*, parue les 21 et 28 juin 1843 dans le *Dollar Newspaper* et traduite en français en 1856 par Charles Baudelaire. La méthode est élémentaire : on calcule d'abord les fréquences des différents signes du texte, le plus fréquent, en anglais comme en français, correspond à la lettre *e*, le *th* est alors facilement identifiable et des bricolages donnent ensuite le reste du message. L'absence de la lettre *e*, comme dans *La disparition* de Georges Perec (1936-1982), intriguera sans doute un moment le cryptanalyste, mais ne le troublera pas longtemps.

2) Utiliser des nombres arbitraires pour coder les mots et syllabes. L'idée de la méthode remonte à 1379, quand Gabriel de Lavinde, secrétaire du pape, alors installé en Avignon, codait ainsi certains mots comme pape, roi de France, etc. Cette forme de chiffrage, appelée nomenclature, devait être appelée à de grands développements et utilisée jusqu'à l'époque moderne. Le développement du télégraphe, la pose de câbles transatlantiques conduisent ceux qui veulent échanger des informations confidentielles à utiliser de tels codes.

Le code Baravelli, en 1894, comportait 100 pages consacrées au vocabulaire comportant 100 mots chacune. Il y avait aussi des pages pour les syllabes, les formes grammaticales, les consonnes. Chacun pouvait convenir de sa propre numérotation des pages. On codait en donnant le numéro de page et le numéro dans la page des mots, syllabes, etc. à coder. Fin 1894, le capitaine Alfred Dreyfus (1859-1935) est accusé par l'Etat-major français de trahison, avec l'Allemagne et l'Italie. Dans un télégramme célèbre, l'ambassade italienne à Paris demande immédiatement à son gouvernement de publier un démenti officiel si les services italiens connaissent la fausseté de l'accusation. Cela aurait totalement innocenté Dreyfus.

L'ambassadeur utilise le code Baravelli et les services français, supposant que le nom de Dreyfus apparaissait dans le message, retrouvent la pagination du code adoptée par les Italiens et percent le sens de leur message. Mais l'affaire Dreyfus ne faisait que commencer...

3) Surcoder le texte en appliquant successivement deux procédés différents. La première idée du surcodage remonte à Leon Battista Alberti (1404-1472), un des grands humanistes de son temps, qui disait des mathématiques que *Rien ne l'aidait plus à chasser la tristesse*. Au printemps 1918, les Allemands chiffrèrent leurs messages d'une façon entièrement nouvelle : les 26 lettres de l'alphabet étaient codées par les 25 couples formés avec les lettres ADFVX (deux lettres rares par le même couple), puis, quelques semaines plus tard, ils complexifièrent le système en codant les 26 lettres et les dix chiffres par les 36 couples formés avec les lettres ADFGVX ; les lettres du message chiffré obtenu ainsi étaient alors permutées suivant un autre principe. Le travail acharné du lieutenant Georges Painvin et des services français permit de décrypter, le 3 juin 1918, au plus fort de la seconde bataille de la Marne, un télégramme indiquant le lieu de la prochaine attaque allemande, ce qui détermina la stratégie du Maréchal Foch (1851-1929) les jours suivants. Certains voient dans ce succès des déchiffreurs français l'origine de la défaite allemande et parlent du *télégramme de la victoire*. Le secret militaire fit qu'il fallut attendre 50 ans, en 1968, pour que Painvin pût expliquer au créateur allemand de ces codes qu'il les avait effectivement cassés et comment (voir Sophie de Lastours, *1914-1918 La France gagne la guerre des codes secrets*, Tallandier, 1998).

4) L'idée de changer de code régulièrement est aussi une idée d'Alberti. Il utilisait un disque central tournant dans une couronne fixe, tous deux divisés en secteurs avec des lettres ou des chiffres permettant, en les tournant, de créer des bijections différentes entre les cases du disque et de la couronne. On indiquait à son correspondant quand on changeait le disque de position. Cette idée est reprise, après des Italiens, par Blaise de Vigenère (1523-1596) qui propose d'utiliser les uns après les autres plusieurs décalages de Jules César en utilisant un mot clé. Le risque d'erreur important empêche sans doute ce système de se répandre et la clé doit être longue pour éviter que des messages trop bavards permettent le déchiffrement.

Ces idées sont à l'origine des célèbres machines Enigma équipées de rotors multipliant les combinaisons et utilisées par les Allemands pendant la guerre de 1939-1945. Les messages des différentes Enigma finirent par être déchiffrés par les Anglais, à partir des travaux de mathématiciens polonais dans les années 1930, puis de mathématiciens anglais comme Alan Turing (1912-1954) dans les années 1939-45 ; de multiples publications le racontent. L'importance de la cryptographie dans les guerres est essentielle, mais peu connue car tenue secrète pendant de longues années. En 1940-45, les déchiffrements des messages chiffrés par les Enigma ont considérablement réduit les pertes de navires anglais, sans que les Allemands imaginent un instant que des codes en lesquels ils avaient toute confiance aient été cassés par leurs ennemis. Certains disent que les idées géniales de Turing ont raccourci la Seconde Guerre mondiale de deux années.

DES et AES

Depuis la fin de la seconde guerre mondiale, l'apparition des ordinateurs, de plus en plus miniaturisés, de puissance de calcul de plus en plus considérable, a changé complètement la cryptographie.

Le profane imagine peut-être la cryptographie comme la création de nouveaux systèmes de chiffrage qu'on garde secrets pour échanger des messages avec ses correspondants, les seuls à connaître le système. Mais cela ne peut donner que des systèmes fragiles, qu'il faut changer complètement à la première trahison, au premier succès d'un espion ennemi. Au contraire, les systèmes de chiffrement les plus utilisés aujourd'hui sont connus de tous, étudiés par tous, ce qui assure leur solidité ; le secret est simplement dans les nombres formant les clés que les utilisateurs choisissent.

Un algorithme de chiffrements par blocs de nature combinatoire, le DES (Data Encryption Standard) a servi de standard au gouvernement américain en 1976, aux entreprises privées en 1981 ; il chiffrait alors une grande partie des échanges mondiaux. Son utilisation a été prolongée jusqu'en 2000 (alors qu'on y avait décelé des faiblesses). On ne sait si le NSA (National Security Agency des États-Unis, un organisme américain beaucoup plus important que la CIA ou le FBI) avait ou non les moyens de le casser, la taille choisie pour la clé, 56 bits, l'ayant peut-être été en fonction de ses capacités de déchiffrement.

En 2000, le DES est remplacé par Rijndael, appelé AES (Advanced Encryption Standard) une méthode de chiffrements par blocs de 128, 192 ou 256 bits et des clés de même taille, ce qui donne neuf choix possibles, le choix de plus grandes tailles assurant plus de sécurité, mais moins de rapidité. Rijndael a été conçu par deux chercheurs belges, Joan Daemen (né en 1965) et Vincent Rijmen (né en 1970). Il peut être implémenté sur une grande variété de matériels.

Méthodes nouvelles

Au début des années 1970, des problèmes cruciaux n'étaient pas résolus, comme les suivants.

- La confidentialité des mots-clés, qui devaient être changés fréquemment pour éviter les possibilités de déchiffrement ; on les transportait par des coursiers spécialisés, personnes vraiment sûres et de toute confiance, qui parcouraient le monde avec des valises contenant les précieuses clés pour les livrer en main propre, parfois accompagnés de gardes du corps (de gendarmes en France). Il fallait changer les clés toutes les semaines, il fallait des clés différentes pour les échanges avec des personnes différentes et, comme le nombre d'utilisateurs augmentait très rapidement, cela devenait ingérable.
- L'authentification de l'auteur d'un message, par exemple pour reconnaître l'identité d'un avion qui s'approche d'une base, l'identité d'un ordinateur avec lequel vous entrez en contact, etc. Est-il ami, est-il ennemi ou neutre ? On ne peut se contenter d'une simple affirmation de sa part ; le risque est qu'il vous trompe et vous bombarde.

- La certitude que le message reçu n'a pas été modifié par un ennemi.
- L'assurance que votre correspondant ne pourra pas revenir sur les termes d'un contrat qu'il a accepté, une tierce personne pouvant vérifier que lui seul a pu envoyer le message que vous présentez comme venant de lui.

En 1975, Whitfield Diffie, né en 1944, a une idée novatrice. Il réalise que toute la cryptographie a été basée jusque-là sur une situation de symétrie entre l'expéditeur et le destinataire : tous deux utilisent la même clé. L'idée de Diffie est de changer cette situation : l'expéditeur d'un message à un destinataire donné, possède une clé de chiffrement de messages pour ce destinataire, appelée clé publique, que tout le monde connaît, mais qui ne permet pas de déchiffrer les messages une fois cryptés et, d'autre part, le destinataire est l'unique possesseur d'une clé de déchiffrement appelée clé privée qu'on ne peut calculer à partir de la seule clé publique. Le problème de la distribution des clés est donc, en quelque sorte, résolu, ainsi que les autres problèmes évoqués ci-dessus, sauf que : ni Diffie, ni ses amis n'ont d'idée pour la réaliser !

Diffie présente ses idées avec Martin Hellman (né en 1945) dans diverses conférences et les publie en novembre 1976 dans un article célèbre : « *New directions in cryptography* ». C'est l'article le plus cité de tous les articles récents de cryptographie, commençant sur un ton prophétique : *We stand today on the brink of a revolution in cryptography. (Nous nous trouvons aujourd'hui à l'aube d'une révolution en cryptographie.)*

Des résultats simples de théorie des nombres allaient bientôt donner des moyens aussi bien pour les échanges de clés que pour les chiffrements à clés publiques.

22.9 LOGARITHME DISCRET

Inventé en 1976 par Diffie et Hellman, il est encore utilisé aujourd'hui. Cependant, des chercheurs français ont montré en 2014 que son utilisation avec des corps de caractéristique petite et la puissance des calculateurs actuels n'était plus sûre.

Le logarithme discret est basé sur l'idée qu'une opération facile à effectuer a un inverse impossible à calculer dans un temps raisonnable avec les méthodes et les moyens de calculs actuels.

On se place dans un corps fini $K = \mathbb{F}_q$ où q est un entier grand, premier ou puissance d'un nombre premier. On sait qu'il existe des générateurs du groupe multiplicatif K^* . On en détermine un noté g . On a donc une bijection $f: \{1, 2, \dots, q-1\} \rightarrow K^*$ définie par $f(x) = g^x$. Ce qui est très difficile est de retrouver x à partir de g^x ; par analogie avec la fonction logarithme, la fonction inverse de f est appelée logarithme discret.

Exemple

Prenons $n = 17$ et $g = 3$. On calcule facilement :

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3^x	3	9	10	13	5	15	11	-1	14	8	7	4	12	2	6	1

On peut inverser cette table :

y	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\log y$	16	14	1	12	5	15	11	10	2	3	7	13	4	9	6	8

Ces deux tables permettent de comprendre ce qu'est le logarithme discret, mais l'important est que le calcul de la seconde table dépasse les possibilités des ordinateurs pour les valeurs de q choisies.

Échange de clés par le logarithme discret

Appelons Juliette et Roméo les deux personnes ou organismes qui désirent échanger secrètement une clé. Elles sont séparées mais conviennent, par un canal non sécurisé, d'un grand nombre premier q et d'un générateur g du groupe multiplicatif K^* de $K = \mathbb{Z}/q\mathbb{Z}$ (nous n'entrerons pas dans les problèmes posés par la détermination de g). Juliette choisit aléatoirement un entier $a < q - 1$ qu'elle seule connaîtra et Roméo choisit aléatoirement un entier $b < q - 1$ que lui seul connaîtra. Ce sont les seules choses qu'ils doivent tenir secrètes. Juliette calcule g^a et Roméo calcule g^b . Ils s'envoient l'un à l'autre le résultat de leur calcul. Ils peuvent alors tous deux calculer g^{ab} (Juliette en calculant $(g^b)^a$, Roméo en calculant $(g^a)^b$) qui leur servira de clé secrète (en fait, une dernière opération est nécessaire pour fabriquer une clé de longueur suffisamment petite à partir de g^{ab}). Le problème de savoir si un ennemi peut récupérer cette clé connaissant q , g , g^a et g^b est toujours activement étudié : en 2014, une équipe française a montré la fragilité du logarithme discret dans le cas de corps de petite caractéristique ; il faut donc maintenant ne plus utiliser ces corps dans la méthode du logarithme discret.

22.10 LA MÉTHODE RSA

L'inconvénient du logarithme discret est qu'il ne permet pas un échange immédiat de clé. Si Roméo ne peut pas répondre instantanément à Juliette, par exemple parce qu'il dort aux antipodes, elle ne peut lui envoyer de message. Il fallait trouver une autre idée.

Ron Rivest (né en 1947), Adi Shamir (né en 1952) et Len Adelman (né en 1945) étaient alors mathématiciens au MIT. Ils s'intéressèrent à la recherche de fonctions

à sens unique pour trouver une réponse au problème de Diffie de chiffrement asymétrique. Après beaucoup d'essais, c'est en avril 1977 que Rivest eut la bonne idée ; ses deux collègues ne découvrirent cette fois-ci aucune faille dans la méthode proposée. Martin Gardner (1914-2010) publia la méthode dans le numéro d'août de *Scientific american*, sous le titre « *A new kind of cipher that would take millions of years to break* » avec un problème qui fut, en fait, résolu 17 ans plus tard (et non des millions d'années plus tard, prévoir dans ce domaine est vraiment difficile), grâce au développement des ordinateurs, d'une part, et des méthodes de factorisation d'entiers en produits de facteurs premiers, d'autre part. Le nom de la méthode est donné par les initiales de ses découvreurs.

On choisit au hasard deux nombres premiers p et q (voir ci-dessous comment les obtenir) suffisamment grands pour que le nombre $n = pq$ ne soit pas factorisable avec les méthodes et les moyens de calcul actuels (la taille de n conseillée en 2019 est 2048 bits).

On choisit ensuite un entier e premier avec $\varphi(n) = (p - 1)(q - 1)$ et on calcule son inverse d modulo $\varphi(n)$; on est le seul à pouvoir faire ce calcul parce qu'on connaît p et q .

Les applications $D, E : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ définies par $E(x) = x^e \bmod n$ et $D(x) = x^d \bmod n$ sont des bijections réciproques car $ed = 1 \bmod \varphi(n)$ d'où $D(E(x)) = x^{ed} = x \bmod n$; pour justifier l'égalité précédente, il suffit de vérifier l'égalité modulo p et modulo q en utilisant le petit théorème de Fermat et d'appliquer le théorème chinois.

Les lettres E, e et D, d ont été choisies pour évoquer les opérations d'*encodage* ou *enchiffrement*, on a le choix entre ces deux mots laids, et de *décodage* ou *déchiffrement*.

Pour transmettre un message M , on le découpe en blocs codés par des entiers $< n$. Pour chiffrer un bloc x , on calcule $y = x^e \bmod n$. Pour le déchiffrer, on calcule $y^d \bmod n$. L'émetteur a besoin de connaître e et n , le récepteur d et n . La connaissance des couples (d, n) ou (e, n) ne livre pas la clé du codage, étant donné le choix de n .

En fait, n et e peuvent être publics, mais la clef de déchiffrement d doit être conservée secrète par son possesseur A . On utilise fréquemment $e = 5$ ou $e = 17$. Ainsi tout le monde peut écrire à A et il peut être commode d'avoir réuni dans une sorte d'annuaire, les couples (n, e) permettant d'écrire à différents destinataires au vu et au su de tout le monde. Pour décoder les messages adressés à A , il faut élever à la puissance d ; seul A a connaissance de ce nombre et personne d'autre ne peut le calculer à partir de e et n .

Les calculs nécessités par l'utilisation du système RSA sont longs : le système RSA est 1000 fois plus lent que le DES. Le système RSA ne peut être utilisé pour transmettre de grandes quantités d'information. Par contre, il est, jusqu'à preuve du contraire, absolument sûr, pourvu qu'on utilise des nombres p et q assez grands. Sinon la factorisation de n permet de calculer d .

Exemple 1

Montrons comment le système RSA fonctionne pour coder des messages écrits avec les 26 lettres de l'alphabet en prenant deux petits nombres premiers :

$$p = 47, q = 59.$$

On calcule

$$n = 47 \times 59 = 2773, \phi(n) = 2668.$$

On choisit, par exemple,

$$e = 17$$

et on calcule

$$d = 157$$

Le nombre 2773 a été choisi parce qu'il est supérieur à 2626 qui code le maximum : ZZ, de deux lettres successives.

Pour coder UN, on prend les rangs de U et de N dans l'alphabet, ce qui donne 2114, puis on calcule $2114^{17} \bmod 2773 = 1644$. Pour décoder ce message, on calcule $1644^{157} \bmod 2773 = 2114$ et on retrouve les lettres U et N par leurs rangs.

Pour coder AS, on calcule $119^{17} \bmod 2773 = 2243$. Le message UN AS, sans blanc, serait donc codé 1644 2243 ; $n = 2668$ et $e = 17$ seront publics, $d = 157$ sera tenu secret.

Exemple 2

Si on veut coder un texte par blocs de 10 lettres, on prendra un produit de deux nombres premiers supérieur à 262626262626262626, comme, par exemple

$$p = 3336670033, q = 9876543211.$$

$$n = 3336670033 \times 9876543211 = 32954765748560082720.$$

On aura :

$$\varphi(n) = 32954765748560082720.$$

On prendra, par exemple :

$$e = 1031$$

et on calculera

$$d = 31963885304131991.$$

Dans l'exemple donné par Gardner, n avait 129 chiffres, p et q en avaient 64 et 65.

Quelques applications de la cryptographie

La France est le pays d'origine des cartes à puce. Au début des années 1980, des organismes comme le CCETT (Centre Commun d'Études de Télédiffusion et de Télécommunication à côté de Rennes) définissent les spécifications de la carte à puce, des banques mènent des expériences et la carte à processeur de Bull est choisie. L'authentification des cartes à puce, basée sur le système RSA, est conçue en 1983 pour une durée de 5 ans à partir de nombres de 321 bits. Dès 1988, Louis Guillou du CCETT souligne la petitesse des nombres, alors que les techniques de factorisation et la puissance des ordinateurs progressent. Faute de tenir compte de ces avertissements, les banques installent un système fragile qui dure jusqu'en 2000, alors que la clé de 320 bits a été cassée en 1998 (les détails de cette histoire ont mis longtemps à être connus).

Dans un autre domaine, avec le lancement du satellite Syracuse IIIA le 13 octobre 2005, les militaires français disposent d'un réseau permettant au commandement d'être en relation permanente avec les unités présentes sur le terrain dans n'importe quelle région du monde ou presque (le Pacifique n'est pas couvert), par exemple en Côte d'Ivoire. Les problèmes de cryptographie pour sécuriser totalement les communications sont étudiés au CELAR (Centre Électronique de l'ARmement à Rennes). La demande de communications militaires est en croissance rapide et tout doit être sécurisé.

Les utilisations de méthodes cryptographiques pour sécuriser les échanges sur le réseau, protéger les ordinateurs, etc. sont extrêmement nombreuses.

22.11 GRANDS NOMBRES PREMIERS

22.11.1 Nombres fortement pseudopremiers (fpp)

La notion de nombre fpp s'appuie sur une propriété bien connue : si K est un corps, le polynôme $X^2 - 1$ a deux racines dans K , les éléments 1 et -1 .

Supposons K fini de cardinal q puissance d'un nombre premier impair. On peut écrire $q - 1 = 2^u v$ avec v impair ; 2^u est la plus grande puissance de 2 divisant $q - 1$. Soit un élément x non nul de K ; on a $x^{q-1} = 1$. Considérons alors la suite d'éléments de K : $(x^{2^i v})_{0 \leq i \leq u} = (x^v, \dots, x^{q-1})$. Comme chaque terme se déduit du précédent en l'élevant au carré, on voit que :

- soit tous les termes de la suite sont égaux à 1 ;
- soit les termes de la suite ne sont égaux à 1 qu'à partir d'un certain rang s et $x^{2^{s-1} v} = -1$.

On comprend alors pourquoi on pose la définition suivante.

Définition

Soit n un entier composé impair. On pose $n - 1 = 2^u v$ avec v impair.

On dira que n est fpp pour un entier a , $1 \leq a \leq n - 1$, s'il vérifie :

- 1) $\text{pgcd}(a, n) = 1$;
- 2) la suite $(a^{2^i v})_{0 \leq i \leq u}$ vérifie l'une des deux propriétés précédentes.

Si un entier n n'est pas fpp pour un entier a , on voit immédiatement qu'il n'est pas premier. Si un entier n est fpp pour plusieurs entiers, on va voir qu'il a de fortes chances d'être premier et donc de ne pas être fpp ! Ce test, appelé test de Rabin-Miller est le résultat de travaux de 1976 de Gary Miller (né en 1946) et de Michael Rabin (né en 1931).

Exemples

$341 = 11 \times 31$ n'est pas fpp pour 2 car $340 = 4 \times 85$, $2^5 = -1 \pmod{11}$, $2^5 = 1 \pmod{31}$, donc $2^5 = (-1, 1)$ dans $\mathbb{Z}/11\mathbb{Z} \times \mathbb{Z}/31\mathbb{Z}$, d'où $2^{85} = (2^5)^{17} = (-1, 1) \neq \pm 1 \pmod{341}$ alors que $2^{170} = 1 \pmod{341}$. Le calcul précédent a utilisé l'isomorphisme du théorème chinois : $\mathbb{Z}/11\mathbb{Z} \times \mathbb{Z}/31\mathbb{Z} \simeq \mathbb{Z}/341\mathbb{Z}$.

Par contre, $2047 = 23 \times 89$ est fpp pour 2 car $2046 = 2 \times 1023$, $1023 = 93 \times 11$, $2^{11} = 2048 = 1 \pmod{2047}$, d'où $2^{1023} = 1 \pmod{2047}$.

Pour un entier composé n , la propriété d'être fpp pour un entier a est beaucoup plus forte que la propriété suivante qui est basée sur le petit théorème de Fermat.

Définition

On dit qu'un entier composé impair n est pseudopremier pour un entier a si $(a, n) = 1$ et $a^{n-1} = 1 \pmod{n}$.

On sait par expérience qu'un entier composé n a une faible probabilité d'être fpp pour un entier inférieur à n et premier avec n . Pour un entier de l'ordre de 2^{256} qui passe le test pour 6 nombres tirés au hasard, elle serait (expérimentalement) de $1/2^{51}$, ce qui est extrêmement faible. On a vérifié que le plus petit entier fpp pour 2 est 2047, le plus petit entier fpp pour 2 et 3 est 1 530 787, le plus petit entier fpp pour 2, 3, 5 et 7 est $N = 3\,215\,031\,751$. On peut donc tester la primalité d'un entier $< N$ d'une manière originale : il suffit juste de vérifier qu'il passe le test de Rabin-Miller pour 2, 3, 5, 7. On est loin de l'algorithme long qui cherche s'il existe des diviseurs premiers de N inférieurs à $\sqrt{N}$.

22.11.2 Construction de grands nombres premiers

La dernière remarque permet de construire rapidement de grands nombres premiers. Pour construire un nombre premier N de 100 chiffres environ en écriture décimale, on choisit au hasard une suite de 300 bits en prenant le premier et le dernier égaux à 1 pour que le nombre soit impair de la taille voulue. La probabilité pour qu'un tel

nombre soit premier est de l'ordre de $1/230$ (d'après le théorème des nombres premiers que nous n'expliquons pas ici). On teste si le nombre est divisible par un petit nombre premier (inférieur à 1000 ou 2000, par exemple). Si c'est le cas, on change N . Après, on teste si N est fpp pour 2, 3, 5, 7, 11. Un nombre de 300 chiffres ayant de très fortes chances d'être premier s'obtenait, il y a quelques années, en moins de cinq minutes et un nombre de 80 chiffres ayant de très fortes chances d'être premier en moins de trois secondes. Avec deux nombres obtenus de cette manière, vous pouvez utiliser le système RSA pour communiquer en étant quasiment sûr que vos nombres sont bien premiers et que personne ne les aura utilisés avant vous.

En 2003, un article de chercheurs indiens (Manindra Agrawal, Neeraj Kayal, Nitin Saxena nés respectivement en 1953, 1979 et 1981) a eu un grand écho en proposant le test AKS qui certifie en temps polynomial si un nombre donné est ou non premier (mais il est trop lent en pratique).

22.12 FACTORISATION

Depuis l'apparition du RSA, les recherches sur la factorisation de grands nombres se sont rapidement développées. Les mathématiques utilisées sont de moins en moins triviales. On n'a jamais prouvé que la factorisation est un problème difficile, mais on a toujours constaté que ce n'était pas facile du tout ; c'est donc un fait d'expérience plutôt qu'une certitude mathématique. Donnons une idée de deux des premières méthodes utilisées dont une que Fermat avait déjà envisagée.

22.12.1 Bases de facteurs

L'exemple le plus ancien d'un tel type d'argument est la factorisation par Fermat de $n = 2\,027\,651\,281$; on a $\sqrt{n} \approx 45029$ et Fermat essaie $t = 45029 + k$. Avec $k = 12$ il obtient : $t^2 = 1020^2 \pmod n$; on a donc $(t - 1020)(t + 1020) = kn$, ce qui lui donne 45041 ± 1020 comme diviseurs de n . *Si l'on alloit par la voie ordinaire, pour trouver la composition d'un tel nombre, au lieu de onze additions, il eût fallu diviser par tous les nombres depuis 7 jusqu'à 44021* écrit Fermat au père Marin Mersenne (1588-1648) dans cette lettre, de 1643 sans doute, dont on ne possède plus qu'une copie.

La méthode des bases de facteurs reprend et améliore l'idée de Fermat. On veut factoriser un entier n composé impair.

L'idée est de trouver deux entiers s et t tels que $t^2 = s^2 \pmod n$. On calculera alors $\text{pgcd}(n, t - s)$ ou $\text{pgcd}(n, t + s)$ pour trouver un facteur de n . Il y a cependant un risque : c'est que $t - s$ ou $t + s$ soient multiples de n . Dans ce cas, on ne se décourage pas et on recommence.

On va chercher s et t un peu au hasard, mais pas tellement : on va partir d'entiers proches de $\sqrt{n}, \sqrt{2n}, \sqrt{3n}$, etc. ; leurs carrés seront petits modulo n et permettront de chercher une factorisation.

Exemple. Factorisation de $1829 = 59 \times 31$

Cet exemple est tiré du livre très précis de Neal Koblitz, *A course in number theory and cryptography*, 1994, 235 p.

Comme $\sqrt{1829} \approx 42,7$, $\sqrt{1829 \times 2} \approx 60,4$, $\sqrt{1829 \times 3} \approx 74,07$,
 $\sqrt{1829 \times 4} \approx 85,4$, on fait les essais suivants de calcul modulo 1829 :

$$42^2 = -65$$

$$43^2 = 20$$

$$61^2 = 63$$

$$74^2 = -11$$

$$85^2 = -91$$

$$86^2 = 80$$

On peut avoir l'impression que rien ne marche et cependant, la solution est proche. Ecrivons les décompositions en facteurs premiers de ces nombres, en indiquant le facteur -1 .

	-1	2	3	5	7	11	13
42^2	1			1			1
43^2		2		1			
61^2			2		1		
74^2	1					1	
85^2	1				1		1
86^2		4		1			

L'inspection de ce tableau montre que $(43 \times 86)^2 = (8 \times 5)^2 \pmod{1829}$. A-t-on gagné ? On a $t + s = 43 \times 86 + 40 = 3738$, $t - s = 3658$. Manque de chance : $t - s = 2 \times 1829$, ce qui n'apporte rien.

Mais on a aussi $(42 \times 43 \times 61 \times 85)^2 = (2 \times 3 \times 5 \times 7 \times 13)^2$, ce qui conduit au calcul de :

$$\text{pgcd}(42 \times 43 \times 61 \times 85 \pm 2 \times 3 \times 5 \times 7 \times 13, 1829) = \text{pgcd}(901 \pm 370, 1829).$$

On trouve : $\text{pgcd}(1271, 1829) = 31$ et $\text{pgcd}(531, 1829) = 59$, c'est gagné !

Mais on aurait poussé jusqu'au calcul de 45^2 , à la façon de Fermat, qu'on aurait gagné du premier coup, car $45^2 = 2025 = 196 = 14^2 \pmod{1829}$, donc $(45 - 14)(45 + 14) = 0 \pmod{1829}$, etc.

Bien sûr, en général, la réussite sera plus longue et, si n est grand, on va peut-être avoir à faire 500000 essais avant d'y parvenir (éventuellement), mais cela en vaut la peine.

22.12.2 La méthode ρ de Pollard

C'est en 1975 que cette méthode est proposée par John M. Pollard ; elle a permis la factorisation du nombre de Fermat $F_8 = 2^{2^8} + 1$ en 1980. Pollard a depuis proposé des méthodes beaucoup plus élaborées.

Soit n un entier.

Par exemple, on part de $x = 1$ et on itère la fonction $f : x \mapsto x^2 + 1 \pmod n$. On pose $x_i = f^i(x)$. La fonction f permet de décrire (en général) l'ensemble $\frac{\mathbb{Z}}{n\mathbb{Z}}$ assez aléatoi-

rement ; on peut choisir d'autres fonctions que $x \mapsto x^2 + 1$ en suivant le même principe. On regarde les différents pgcd : $\text{pgcd}(x_i - x_j, n)$ en espérant trouver avec un peu de chance un facteur de n .

Exemples

Prenons $n = 91$. On obtient $x_0 = 1, x_1 = 2, x_2 = 5, x_3 = 26$ et $\text{pgcd}(x_3 - x_2, n) = \text{pgcd}(21, 91) = 7$; on a gagné ! On trouve $91 = 7 \times 13$.

Prenons $n = 221$. On obtient $x_0 = 1, x_1 = 2, x_2 = 5, x_3 = 26, x_4 = 677 = 14 \pmod{221}$ et $\text{pgcd}(x_4 - x_0, n) = \text{pgcd}(13, 221) = 17$; on a gagné ! On trouve $221 = 13 \times 17$.

Prenons $n = 323$. On obtient $x_0 = 1, x_1 = 2, x_2 = 5, x_3 = 26, x_4 = 677 = 31 \pmod{323}, x_5 = 962 = -7 \pmod{323}$ et $\text{pgcd}(x_5 - x_4, n) = \text{pgcd}(-38, 323) = 19$; on a gagné ! On trouve $323 = 17 \times 19$.

22.13 COURBES ELLIPTIQUES

Une question de mode ?

Après avoir connu une vogue extraordinaire, le RSA semble moins à la mode aujourd'hui : il est lent et les nombres premiers auxquels on doit faire appel sont de plus en plus grands, afin de garantir que les messages actuellement chiffrés avec cette méthode ne soient pas déchiffrables non seulement aujourd'hui, mais aussi dans quelques années, au moins pour certains d'entre eux. La méthode à la mode (elle est recommandée par le NSA) est celle des courbes elliptiques, basée sur des travaux indépendants, en 1984-85, de Neal Koblitz (né en 1948) et Victor Miller (né en 1947). Elle n'est pas très rapide non plus, mais ses clés sont plus courtes (170 bits contre 1024 en 2011) pour le même niveau de sécurité.

Les courbes elliptiques intéressent les mathématiciens depuis Newton dans les années 1670. Elles apparaissent dans différents champs des mathématiques et sont, en particulier, au cœur de la démonstration du théorème de Fermat par Wiles. Elles peuvent aussi servir pour tester si un entier est premier ou pour factoriser un entier.

La littérature sur les courbes elliptiques est particulièrement abondante. Les lecteurs et lectrices de ce livre pourront prendre un premier contact avec elles au travers de cette section que nous leur consacrons. Il vous faudra admettre quelques résultats, dont l'énoncé est très accessible, pour pouvoir, je l'espère, comprendre ce qui suit.

La même idée que pour le logarithme discret

Avant de définir ce qu'est une courbe elliptique, expliquons l'intérêt de ces objets en cryptographie. Les points d'une courbe elliptique E sur un corps fini forment un groupe abélien $\Gamma(E)$; on peut choisir (E) de façon que $\Gamma(E)$ ait des points d'ordre fini très grand dans $\Gamma(E)$. Si Juliette et Roméo conviennent d'utiliser un tel point P , ils peuvent choisir en secret, la première un entier a , le second un entier b . Juliette calcule aP qu'elle envoie à Roméo, Roméo calcule bP qu'il envoie à Juliette ; tous deux peuvent calculer $(ab)P = a(bP) = b(aP)$ qui leur servira de clé secrète (ou plus souvent à fabriquer une clé secrète). L'intérêt de la méthode est que personne, comme pour le logarithme discret, n'est aujourd'hui capable quand les données sont bien choisies, de retrouver $(ab)P$ à partir de P, aP, bP .

Courbe elliptique

On appellera ici courbe elliptique sur un corps K , l'ensemble des points (x, y) de K^2 vérifiant :

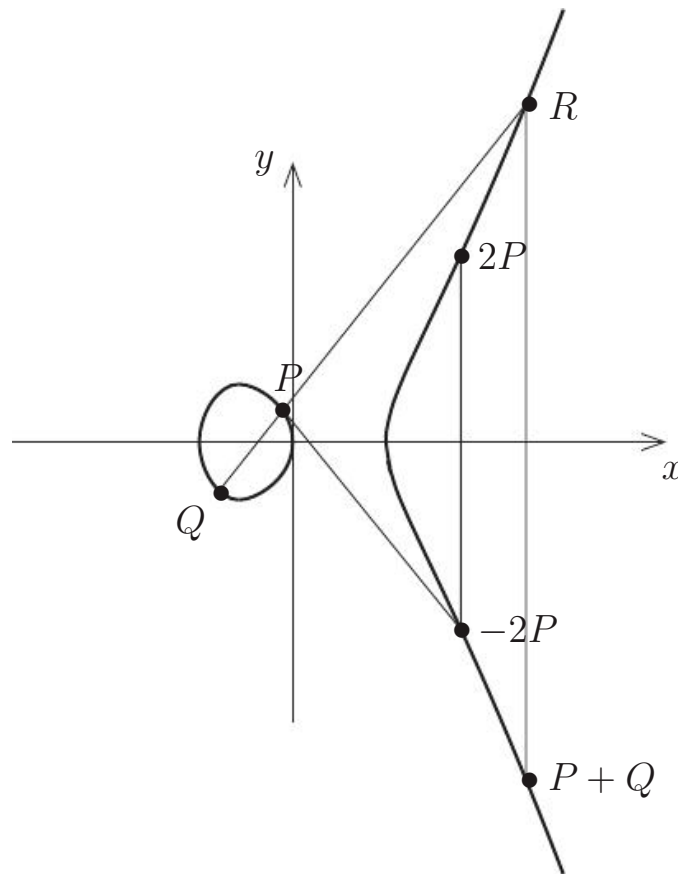
- si $\text{car}(K) \neq 2, 3$: $y^2 = x^3 + ax + b$ avec $x^3 + ax + b = 0$ n'ayant que des racines simples ;
- si $\text{car}(K) = 3$: $y^2 = x^3 + ax^2 + bx + c$ avec $x^3 + ax^2 + bx + c = 0$ n'ayant que des racines simples ;
- si $\text{car}(K) = 2$: $y^2 + cy = x^3 + ax + b$ avec $c \neq 0$ ou $y^2 + xy = x^3 + ax^2 + b$ avec $b \neq 0$.

Dans chaque cas, on ajoute dans la direction de Oy un point O à l'infini. On pourrait donner une définition plus générale englobant tous les cas ; les conditions données assurent que la courbe n'a pas de point singulier (de point où les dérivées partielles par rapport à x et y soient toutes deux nulles).

En cryptographie, les corps ayant un nombre d'éléments égal à une puissance de 2 sont parmi les plus choisis.

Structure de groupe sur une courbe elliptique

Comment définir la structure de groupe sur $\Gamma(E)$? Présentons-le sur l'exemple de la courbe E définie sur $\mathbb{R}$ par l'équation $y^2 = x(x-1)(x+1) = x^3 - x$ dont le graphe est le suivant.



Si P et Q sont deux points distincts de E , la droite (PQ) recoupe E en un troisième point R et la droite (OR) , parallèle à l'axe des y , recoupe la courbe au point qu'on définit comme $P + Q$. Il y a des cas particuliers : si $P = Q$, la tangente en P à E recoupe E en R et on définit $P + Q = 2P$ comme précédemment. La construction justifie que $O + P = P$ pour tout $P \in E$, autrement dit O est élément neutre pour la loi qu'on vient de définir. On vérifie aussi que si $P = (x, y) \in E$, alors $-P = (x, -y)$ (cette formule ne vaut pas pour $\text{car}(K) = 2$). La commutativité de la loi est immédiate ; l'associativité résulte d'une démonstration plus longue. On remarque que, si trois points P, Q, R de E sont alignés, alors $P + Q + R = 0$ et que les points d'ordre 2 sont les points où la tangente à la courbe est verticale.

Supposons $\text{car}(K) \neq 2, 3$. Si $P = (x_1, y_1)$ et $Q = (x_2, y_2)$ avec $x_2 \neq x_1$, l'équation de la droite (PQ) est $y = m(x - x_1) + y_1$ avec $m = \frac{y_2 - y_1}{x_2 - x_1}$. L'abscisse x_3 de l'intersection de (PQ) et E est la troisième racine de $x^3 - (m(x - x_1) + y_1)^2 + ax + b = 0$. La somme des racines est l'opposé du coefficient de x^2 , d'où $x_3 = m^2 - x_1 - x_2$. On en déduit $P + Q = (x_3, y_3)$ avec

$$x_3 = m^2 - x_1 - x_2$$

$$y_3 = -y_1 - m(x_3 - x_1).$$

Si $Q = P$, on a, en dérivant, $2yy' = 3x^2 + a$, d'où $m = \frac{3x_1^2 + a}{2y_1}$. On en déduit que $P + Q = 2P = (x_3, y_3)$ avec

$$\begin{aligned}x_3 &= m^2 - 2x_1 \\ y_3 &= -y_1 - m(x_3 - x_1).\end{aligned}$$

On obtient des formules analogues si $\text{car}(K) = 2, 3$ (voir exercice 22.7).

Courbes elliptiques sur un corps fini

On suppose que $K = \mathbb{F}_q$ est un corps fini de caractéristique p (q est une puissance de p). On va énoncer sans démonstration plusieurs résultats difficiles mais compréhensibles.

Le premier est que la structure du groupe $\Gamma(E)$ est celle d'un produit de deux groupes cycliques.

Pour le second, notons N le nombre d'éléments de E ; un résultat de Helmut Hasse (1898-1979) montre que N est *proche* de q en donnant l'encadrement suivant :

$$q + 1 - 2\sqrt{q} \leq N \leq q + 1 + 2\sqrt{q}.$$

Pour le troisième, on remarque qu'une équation de courbe elliptique E dans le corps K définit une famille de courbes elliptiques E_r sur les extensions $\mathbb{F}_{q^r}$ de K . Notons N_r le nombre de points de E_r dans K . Un résultat de Weil, contenu dans les conjectures qu'il a formulées en 1949, donne l'égalité formelle

$$\sum_{r \geq 1} N_r \frac{X^r}{r} = \ln \frac{1 - \alpha X + qX^2}{(1 - X)(1 - qX)}$$

avec $\alpha = q + 1 - N_1$. Ainsi les N_r sont connus dès qu'on connaît N_1 . On remarque que le $1 - \alpha X + qX^2$ a deux racines complexes conjuguées puisque son discriminant $\alpha^2 - 4q$ est négatif d'après l'inégalité de Hasse.

Un exemple

On part de E d'équation $y^2 + y = x^3$ sur $K = \mathbb{F}_2$. On vérifie que E a trois points : $O, (0, 0), (0, 1)$. On a donc $N_1 = 3$, d'où $\alpha = 0$. Par conséquent :

$$\sum_{r \geq 1} N_r \frac{X^r}{r} = \ln \frac{1 + 2X^2}{(1 - X)(1 - 2X)}.$$

En dérivant, on obtient :

$$\begin{aligned}\sum_{r \geq 1} N_r X^{r-1} &= \frac{4X}{1 + 2X^2} + \frac{1}{1 - X} + \frac{2}{1 - 2X} \\ &= 4X(1 - 2X^2 + 4X^4 + \dots) + 1 + X + X^2 + \dots + 2(1 + 2X + 4X^2 + \dots).\end{aligned}$$

On en déduit que $N_2 = N_3 = N_4 = 9$, $N_5 = 33$, etc. ; en fait, on voit que $N_r = 2^r + 1$ si r est impair et $N_r = 2^r + 1 - 2(-1)^{r/2} 2^{r/2}$ si r est pair.

Exercices

22.1 Codes correcteurs

1) Soient $n = 2m + 1$ un entier impair et C le code consistant à répéter n fois chaque bit. Vérifier que C est un code linéaire parfait ; donner son type, son taux d'information et le nombre d'erreurs qu'il peut corriger.

2) a) Vérifier que, dans le cas d'un codage systématique, la matrice G est de la forme

$$G = \begin{pmatrix} I_k \\ M \end{pmatrix} \text{ où } M \text{ est une matrice à } r = n - k \text{ lignes et } k \text{ colonnes et qu'on peut prendre}$$

comme matrice de contrôle $H = (M \ I_{n-k})$.

b) Montrer que ce résultat s'applique au code de Hamming.

3) Vérifier que le code de Hamming peut être vu comme un code cyclique en effectuant une permutation convenable des vecteurs de base.

4) Que font les codes cycliques définis par $G = 0$, $G = 1$, $G = T$, $G = T + 1$, $G = (T^n - 1)/(T - 1) = T^{n-1} + \dots + 1$?

22.2 Minitel

Le Minitel a connu la célébrité dans les années 1980-2000 alors que le réseau actuel n'existait pas. C'était une magnifique création française, pilotée par le ministère des postes et télécommunications (voir le livre de Valérie Schafer, *Le Minitel*, Nuvis, 2012). Le Minitel proposait l'annuaire téléphonique, la vente par correspondance et la messagerie rose qui connut un immense succès. Les appareils étaient distribués gratuitement. Depuis le milieu des années 1990, la toile a considérablement réduit l'intérêt du Minitel ; deux millions de personnes l'utilisaient encore en 2010 ; il a été fermé en 2012.

Un message Minitel est une suite de 0 et de 1 ; les statistiques montrant que les erreurs de transmission, réception d'un 0 à la place d'un 1 ou vice versa, n'étaient pas très fréquentes, les chercheurs ont mis au point un code correcteur capable de corriger une erreur de transmission sur 120 bits transmis. Ce code est un code BCH utilisant un corps à 128 éléments défini comme le quotient $K = \mathbb{F}_2[T]/(P)$, avec $P = T^7 + T^3 + 1$.

Le code BCH donne des messages de 127 bits. Un message du minitel occupe en fait 17 octets : 15 pour les 120 bits du message, un pour les 7 bits de contrôle et un

bit de parité, un dernier octet de zéros validant le message et détectant d'éventuelles perturbations de la ligne. On suppose qu'une erreur de transmission au maximum intervient.

- a) Vérifier l'irréductibilité de P .
- b) Montrer que la classe t de T engendre K^* .
- c) Expliquer comment transmettre une suite de 120 bits $x = (x_k)_{0 \leq k \leq 119}$ avec le code BCH défini par les données précédentes et comment retrouver le message initial quand on reçoit une suite $y = (y_k)_{0 \leq k \leq 126}$.

22.3 Code BCH

On reprend l'exemple de 22.7 du corps de décomposition de $X^{15} - 1$; on note t la classe de T dans le quotient $K = \frac{\mathbb{F}_2[T]}{(T^4 + T + 1)}$ et on utilisera la table des puissances de t donnée alors.

- 1) Construire un code BCH C avec la distance prescrite 7 tel que $T^4 + T + 1$ divise G .
- 2) Décrire le codage dans C .
- 3) Retrouver le message d'origine en appliquant la méthode de correction d'erreur présentée en 22.7 quand on reçoit les mots :
 $y = (1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1)$;
 $y = (0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1)$.

22.4 Système RSA

Vous pouvez expérimenter le système RSA à l'aide de différents logiciels. Il faut d'abord avoir écrit des procédures de découpage de textes en blocs (de 10 caractères par exemple), de codage et décodage des caractères en chiffres (a codé par 01, b par 02, etc.). Il faut ensuite choisir les entiers p, q, e , calculer d , comme dans l'exemple 2. Vous pouvez recevoir des messages chiffrés de vos amis et amies en leur donnant votre clef publique, etc.

22.5 Nombres pseudo-premiers

- 1) a) Pour quelles valeurs de a entre 1 et 91 le nombre composé $91 = 7 \times 13$ est-il pseudopremier ?
 b) Pour quelles valeurs de a entre 1 et 91 le nombre composé $91 = 7 \times 13$ est-il fpp ?
- 2) Vérifier que 65 est fpp pour 8 et 18 mais non pour leur produit.
- 3) On pose $n = 561 = 3 \times 11 \times 17$.
 a) Calculer $\varphi(n)$.
 b) Pour quelles valeurs de a entre 1 et 561 le nombre n est-il pseudopremier ?

Robert Carmichael (1879-1967) a découvert les propriétés de 561 en 1910, puis d'autres nombres ayant la propriété $a^{n-1} \equiv 1 \pmod n$ pour tout a premier avec n ; ces nombres sont appelés depuis nombres de Carmichael ; on sait depuis 1992 qu'il en existe une infinité.

22.6 Factorisations

- 1) Factorisation de 8989. Utiliser la méthode de Pollard avec $f: x \mapsto x^2 + 1$ et $x_0 = 7$, puis la méthode de Fermat.
- 2) Factorisation de 2701. Utiliser la méthode de Fermat, puis la méthode de Pollard avec $f: x \mapsto x^3 + x + 1$ et $x_0 = 1$ (il faut aller jusqu'à x_6). Vérifier que 2701 est pseudopremier pour 2 et 3.
- 3) Factorisation de $n = 2^{35} - 1$.
 - a) Trouver deux facteurs premiers de n à l'aide d'identités algébriques.
 - b) Montrer que, si $p > 2$ est un nombre premier et si d est l'ordre de a dans $U(p)$, alors $p \equiv 1 \pmod d$; si, de plus, d est impair, alors $p \equiv 1 \pmod{2d}$.
 - c) Terminer la factorisation de n .

22.7 Points des courbes elliptiques

- 1) En 1621, Bachet indique dans son langage que si on connaît une solution rationnelle (x_1, y_1) de l'équation $y^2 = x^3 + c$, avec $x_1 \neq 0$, une autre solution rationnelle de cette équation est $\left(\frac{x_1^4 - 8cx_1}{4y_1^2}, \frac{x_1^6 + 20cx_1^3 - 8c^2}{8y_1^3} \right)$. Relier la formule de Bachet aux courbes elliptiques. Comment peut-on en général obtenir d'autres points rationnels de la courbe ?
- 2) Obtenir des formules donnant les coordonnées x_3, y_3 de $P + Q$ si $\text{car}(K) = 2, 3$.
- 3) Dans la formule de Weil, on note u, v les nombres complexes conjugués tels que $1 - \alpha X + qX^2 = (1 - uX)(1 - vX)$. Vérifier que $N_r = q^r + 1 - u^r - v^r$.

22.8 $y^2 = x^3 - 3x$ modulo 11

Pour tout entier $r \geq 1$, on note E_r la courbe elliptique de $\mathbb{F}_{11^r}$ d'équation $y^2 = x^3 - 3x$; on note N_r le nombre de points de E_r et on note Γ_r le groupe des points de E_r . On définit le corps à $121 = 11^2$ éléments comme le quotient $K = \frac{\mathbb{F}_{11}[X]}{(X^2 - 2)}$.

- 1) Déterminer N_1 ; en déduire N_2, N_3 , puis N_r pour tout entier $r \geq 1$.
- 2) Quels sont les points d'ordre 2 de Γ_1 ? En déduire la structure de Γ_1 .
- 3) a) Montrer qu'un point $P(u, v)$ est d'ordre 3 si et seulement si $u^4 - 6u^2 - 3 = 0$.
 - b) Quels sont les points d'ordre 3 de Γ_1 ?
 - c) Quels sont les points d'ordre 3 de Γ_2 dans K^2 ?

- 4) Montrer qu'un point $P(u, v)$ est d'ordre 4 si et seulement si : $(u^2 + 3)(u^2 - u - 3) \equiv 0 \pmod{4}$.
- b) Quels sont les points d'ordre 4 de Γ_2 ?
- c) Quelle est la structure de Γ_2 ?
- 5) Quelle est la structure de Γ_3 ?

Solutions

22.1 1) Le code C est défini par $\gamma : \mathbb{F}_2 \rightarrow \mathbb{F}_2^n$ avec $\gamma(0) = (0, \dots, 0)$ et $\gamma(1) = (1, \dots, 1)$; C est donc un code linéaire de type $[n, 1, n]$ et de taux d'information $1/n$. La boule de centre $\gamma(0)$ et de rayon m contient les suites de n bits comportant au plus m bits égaux à 1. Son complémentaire est la boule de centre $\gamma(1)$ et de rayon m , ce qui montre que C est parfait et peut corriger jusqu'à m erreurs.

2) a) Dans le calcul des coefficients du produit HG , seuls deux termes interviennent dans chaque cas.

b) Pour faire apparaître les formes précédentes, on peut permuter les lignes de G et les colonnes de H avec la même permutation : les prendre dans l'ordre 3 5 6 7 1 2 4, puis permuter les lignes de H dans l'ordre 3 2 1.

3) Construisons un tableau avec les seize vecteurs du code de Hamming dans la colonne de gauche et leurs transformés par la permutation (1 2)(3 4)(6 7) des vecteurs de base de A^7 .

(0,0,0,0,0,0,0)	(0,0,0,0,0,0,0)
(1,1,1,0,0,0,0)	(1,1,0,1,0,0,0)
(1,0,0,1,1,0,0)	(0,1,1,0,1,0,0)
(0,1,0,1,0,1,0)	(1,0,1,0,0,0,1)
(1,1,0,1,0,0,1)	(1,1,1,0,0,1,0)
(0,1,1,1,1,0,0)	(1,0,1,1,1,0,0)
(1,0,1,1,0,1,0)	(0,1,1,1,0,0,1)
(0,0,1,1,0,0,1)	(0,0,1,1,0,1,0)
(1,1,0,0,1,1,0)	(1,1,0,0,1,0,1)
(0,1,0,0,1,0,1)	(1,0,0,0,1,1,0)
(1,0,0,0,0,1,1)	(0,1,0,0,0,1,1)
(0,0,1,0,1,1,0)	(0,0,0,1,1,0,1)
(1,0,1,0,1,0,1)	(0,1,0,1,1,1,0)
(0,1,1,0,0,1,1)	(1,0,0,1,0,1,1)
(0,0,0,1,1,1,1)	(0,0,1,0,1,1,1)
(1,1,1,1,1,1,1)	(1,1,1,1,1,1,1)

Sous cette nouvelle forme, on vérifie que le code est cyclique.

4) Si $G = 0$, $C = \{0\}$; si $G = 1$, $C = A^n$; si $G = T$, on répète le dernier terme ; si $G = T + 1$, on ajoute un bit de parité ; si $G = (T^n - 1)/(T - 1)$, on trouve le code de la répétition n fois.

22.2 a) La vérification de l'irréductibilité du polynôme $P = T^7 + T^3 + 1$ se fait en plusieurs temps. Si ce polynôme de degré 7 n'était pas irréductible, il aurait un facteur irréductible de degré 1, 2 ou 3. On vérifie successivement que $P(0) = P(1) \neq 0$, ce qui signifie que P n'a pas de facteur du premier degré, que P n'est pas divisible par le seul polynôme irréductible de degré 2 : $T^2 + T + 1$ et enfin que P n'est pas divisible par les deux polynômes irréductibles de degré 3 : $T^3 + T + 1$ et $T^3 + T^2 + 1$.

b) Comme l'ordre du groupe K^* est 127 et que 127 est un nombre premier, tout élément de K^* autre que 1 est un générateur du groupe. En particulier, la classe t de T engendre K^* .

c) Le 7-uplet $B = (t^6, t^5, t^4, t^3, t^2, t, 1)$ est une base de l'espace vectoriel K sur $\mathbb{F}_2$.

Quand on veut transmettre la suite de bits $x = (x_k)_{0 \leq k \leq 119}$, on envoie la suite $\gamma(x) = (c_k)_{0 \leq k \leq 126}$ où $c_k = x_k$ pour $0 \leq k \leq 119$ et dont les derniers termes sont donnés (un dispositif technique astucieux rend ce calcul aisé) par l'expression de $t^7 \sum_{0 \leq k \leq 119} x_k t^k$ dans la base B :

$$t^7 \sum_{0 \leq k \leq 119} x_k t^k = \sum_{0 \leq k \leq 6} c_{120+k} t^k$$

Comme on suppose qu'une erreur au maximum peut avoir lieu dans la transmission de la suite x , un seul des termes, au maximum, de la suite $y = (y_k)_{0 \leq k \leq 126}$ du message reçu diffère des termes de la suite x . On calcule la somme $t^7 \sum_{0 \leq k \leq 119} y_k t^k + \sum_{0 \leq k \leq 6} y_{120+k} t^k$ dans la base B . Si la somme est nulle, on admet qu'aucune erreur de transmission n'a eu lieu et on pose $x = (y_k)_{0 \leq k \leq 119}$. Si une erreur de transmission a eu lieu, la somme est t^s exprimée dans la base B et une table permet de retrouver l'exposant s . Si on a $s \leq 6$, c'est que l'erreur a eu lieu sur un des bits de contrôle ; les bits du message initial ont été transmis sans erreur et on pose $x = (y_k)_{0 \leq k \leq 119}$. Si on a $6 < s \leq 126$, on pose $x = (y_k)_{0 \leq k \leq 119} + e$ où e est une suite ne comportant qu'un seul 1, au rang $s - 7$. On notera que, si on a $s \geq 6$, au moins deux des bits de contrôle ont changé.

22.3 On aura constamment besoin de la table des puissances de t dans la base $(t^3, t^2, t, 1)$ pour effectuer les nombreux calculs de cet exercice.

1) G a pour facteurs : $T^4 + T + 1$ dont les racines sont $t, t^2, t^4, t^8, T^4 + T^3 + T^2 + T + 1$ qui a t^3 (vérifier que $t^{12} + t^9 + t^6 + t^3 + 1 = 0$) et t^6 comme racines et $T^2 + T + 1$ qui a t^5 pour racine (vérifier que $t^{10} + t^5 + 1 = 0$). Comme G a $t, t^2, t^3, t^4, t^5, t^6$ parmi ses racines, la distance minimale du code qu'il définit est supérieure ou égale à 7 et C corrige jusqu'à 3 erreurs.

2) La dimension du code est $k = 15 - \deg(G) = 5$. On code $(x_1, x_2, x_3, x_4, x_5)$ en posant $c_T = x_1 T^{14} + x_2 T^{13} + x_3 T^{12} + x_4 T^{11} + x_5 T^{10}$ et en effectuant la division

euclidienne de $c_I = QG + c_R$ de c_I par G ; les bits suivants du codage de x sont les coefficients de c_R .

3) On doit suivre pas à pas la méthode donnée en 22.7. On note P le polynôme associé à y : $P = T^{14} + T^{13} + T^9 + T^7 + T^5 + T^2 + T + 1$. On calcule $P(t^i)$ pour $i = 1, \dots, 6$: on trouve : $P(t) = t^2$, $P(t^2) = t + 1$, $P(t^3) = t$, $P(t^4) = t^2 + 1$, $P(t^5) = 1$, $P(t^6) = t^2$. Comme la matrice H_3 qu'on forme avec ces valeurs est de rang 3, il y a eu trois erreurs de transmission. On trouve ensuite les trois coefficients du polynôme $S(Z) = Z^3 + t^2Z^2 + (t^3 + 1)Z + t^3 + t^2$ en résolvant le système linéaire indiqué en 22.7. Par essais successifs, on trouve les racines de S : $1, t^9, t^{12}$; les erreurs sont donc aux rangs 14, 6 et 3, ce qui permet de corriger y en $c = (1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0)$. Le message envoyé est $x = (1, 1, 1, 0, 0)$.

4) Reprendre la même démarche qu'au **3)**. Ici H_3 est de rang 1 et l'erreur est au rang 12.

22.5 1) a) Il y a $\varphi(91) = \varphi(7)\varphi(13) = 72$ nombres premiers avec 91 et inférieurs à 91. Si 91 est pseudopremier pour a , on a $a^{90} = 1 \pmod{91}$, soit $a^{6 \times 15} = 1 \pmod{91}$, ce qui équivaut, par l'isomorphisme du théorème chinois, à $a^{6 \times 15} = 1 \pmod{7}$ et $a^{6 \times 15} = 1 \pmod{13}$.

La première égalité est toujours vraie, car $a^{6 \times 15} = (a^6)^{15}$ et le petit théorème de Fermat montre que $a^6 = 1 \pmod{7}$. La seconde donne $a^{6 \times 15} = a^{7 \times 12 + 6} = a^6 = 1 \pmod{13}$; comme $6 = (13 - 1)/2$, la question revient donc à déterminer les a résidus quadratiques modulo 13 : ce sont 1, 3, 4, 9, 10, 12. On obtient donc 36 couples correspondants aux a cherchés qu'on peut facilement énumérer : 1, 3, 4, 9, 10, 12, 16, 17, 22, 23, 25, 27, 29, 30, 36, 38, 40, 43 et leurs opposés.

b) 91 est fpp pour a si et seulement si $a^{45} = \pm 1 \pmod{91}$, soit, d'après l'isomorphisme du théorème chinois, $a^3 = 1 \pmod{7}$ et $a^3 = 1 \pmod{13}$ ou encore $a^3 = -1 \pmod{7}$ et $a^3 = -1 \pmod{13}$. On trouve 3 solutions pour chacune de ces équations, ce qui donne 18 valeurs pour lesquelles 91 est fpp.

2) On a $8 \times 18 = 14 \pmod{65}$, $8^2 = 18^2 = -1 \pmod{65}$, donc $14^2 = 1 \pmod{65}$; puisque $65 - 1 = 2^6$, il faudrait que $14^1 = \pm 1 \pmod{65}$, ce qui est faux.

3) a) $\varphi(n) = 2 \times 10 \times 16 = 320$.

b) On a $n - 1 = 560 = 16 \times 5 \times 7$. Soit a un des 320 nombres premiers avec n . Pour calculer $a^{560} \pmod{561}$, on utilise l'isomorphisme du théorème chinois : $\mathbb{Z}/561\mathbb{Z} = \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/11\mathbb{Z} \times \mathbb{Z}/17\mathbb{Z}$. L'élévation à la puissance 560 donne 1 sur les composantes car 560 est multiple de 2, de 10 et de 16. Par conséquent, 561 est pseudopremier pour les 320 valeurs de a premier avec n .

22.6 1) Avec la méthode de Pollard, on trouve $\text{pgcd}(x_4 - x_1, 8989) = 89$ et $8989 = 89 \times 101$. Avec la méthode de Fermat, on commence par des nombres proches de $\sqrt{8989}$ comme 94 et 95. Comme $95^2 = 36 = 6^2 \pmod{8989}$, on obtient immédiatement la factorisation : $8989 = 89 \times 101$.

2) Pour la méthode de Fermat, on essaie des nombres proches de $\sqrt{2701}$: $52^2 = 3 \bmod 2701$, $53^2 = 2^2 \times 3^3 \bmod 2701$. On est conduit à $(52 \times 53)^2 = 18^2 \bmod 2701$, d'où $2774 \times 2738 = 0 \bmod 2701$; $\text{pgcd}(2701, 2774) = 73$, $\text{pgcd}(2701, 2738) = 37$ et $2701 = 37 \times 73$.

Pour la méthode de Pollard, on trouve la suite $x_1 = 3$, $x_2 = 31$, $x_3 = 112$, $x_4 = 521$, $x_5 = 2325$, $x_6 = 630$, et un calcul aboutit, celui de $\text{pgcd}(2701, 630 - 112) = 37$, d'où la factorisation de 2701.

Pour vérifier que $2701 = 37 \times 73$ est pseudopremier pour 2, on vérifie que $2700 = 4 \times 25 \times 27 = 36 \times 75 = 9 \times 300$, puis que $2^{36} = 1 \bmod 37$, $2^9 = 512 = 1 \bmod 73$, donc, à l'aide du théorème chinois, $2^{36} = 1 \bmod 2701$, donc $2^{2700} = 1 \bmod 2701$. Pour vérifier que 2701 est pseudopremier pour 3, on vérifie que $3^{36} = 1 \bmod 37$, $3^6 = 729 = -1 \bmod 73$, etc.

3) a) Comme $35 = 5 \times 7$ et comme $X^k - 1$ divise $X^{kn} - 1$, les nombres premiers entre eux $2^5 - 1 = 31$ et $2^7 - 1 = 127$ divisent n . On trouve $n = 31 \times 127 \times 8727391$.

b) Comme $a^{p-1} = 1 \bmod p$, si a est d'ordre d dans $U(P)$, alors d divise $p - 1$ et, si d est impair, $2d$ divise $p - 1$, ce qui donne les égalités demandées.

c) Soit p un nombre premier divisant $2^{35} - 1$. L'ordre de 2 dans $U(p)$ est un diviseur de 35.

Si c'est 5, on voit que p divise $2^5 - 1 = 31$, donc $p = 31$; si c'est 7, on voit que p divise $2^7 - 1 = 127$, donc $p = 127$. On vérifie que ces deux nombres ne divisent pas 8 727 391.

Si c'est 35, on tire une autre conclusion : le 2) montre que $p = 1 \bmod 70$. On trouve que $8727391 = 71 \times 122921$. Comme $\sqrt{122921} < 351$, les seuls facteurs premiers possibles de 122 921 sont 71, 141 qui n'est pas premier, 211, 281 (on a réussi à repousser la procédure *bête* de façon à ne plus avoir que très peu d'essais à faire). Comme aucun ne convient, 122 921 est premier et on obtient la factorisation :

$$2^{35} - 1 = 31 \times 127 \times 71 \times 122921.$$

22.7 1) Soit $P = (x_1, y_1)$ la solution connue. C'est un point de la courbe E d'équation $y^2 = x^3 + c$. Un autre point de cette courbe est $2P$. On peut vérifier que les coordonnées indiquées sont celles de $2P$ et qu'elles sont rationnelles. En itérant la construction, on obtiendrait d'autres points rationnels de la courbe, distincts des précédents en général.

2) a) En caractéristique 3, si $x_1 \neq x_2$, on a $m = \frac{y_2 - y_1}{x_2 - x_1}$; on obtient :

$$x_3 = m^2 - a - x_1 - x_2 ;$$

$$y_3 = -m(x_3 - x_1) - y_1.$$

b) Dans le même cas, si $P = Q$, on a $m = \frac{ax_1 - b}{y_1}$ et :

$$x_3 = m^2 - a + x_1;$$

$$y_3 = -m(x_3 - x_1) - y_1.$$

c) En caractéristique 2, dans le cas de l'équation $y^2 + cy = x^3 + ax + b$, on a

$$m = \frac{y_2 + y_1}{x_2 + x_1}; \text{ on obtient si } x_1 \neq x_2:$$

$$x_3 = m^2 + x_1 + x_2;$$

$$y_3 = m(x_3 + x_1) + y_1 + c.$$

Pour trouver y_3 , il faut remarquer que les deux solutions u, v d'une équation de la forme $y^2 + cy = A$ vérifient $u + v = c$.

d) Dans le même cas, si $P = Q$, on a $m = \frac{x_1^2 + a}{c}$ et :

$$x_3 = m^2;$$

$$y_3 = m(x_3 + x_1) + y_1 + c.$$

e) En caractéristique 2, dans le cas de l'équation $y^2 + xy = x^3 + ax^2 + b$, on a

$$m = \frac{y_2 + y_1}{x_2 + x_1}; \text{ on obtient si } x_1 \neq x_2:$$

$$x_3 = m^2 + m + a + x_1 + x_2;$$

$$y_3 = m(x_3 + x_1) + y_1.$$

Pour trouver y_3 , il faut remarquer que les deux solutions u, v d'une équation de la forme $y^2 + xy = A$ vérifient $u + v = x$.

f) Dans le même cas, si $P = Q$, on a $m = x_1 + \frac{y_1}{x_1}$ et :

$$x_3 = x_1^2 + \frac{b}{x_1^2};$$

$$y_3 = x_1^2 + (m + 1)x_3.$$

3) En remplaçant le numérateur par $(1 - uX)(1 - vX)$, en dérivant et en développant en série comme dans l'exemple du cours, l'identification des termes de même degré des deux membres donne immédiatement le résultat.

22.8 1) On dresse un tableau donnant pour tout $x \in \mathbb{F}_{11}$ les valeurs possibles de y .

x	0	1	2	3	4	5	6	7	8	9	10
x^3	0	1	8	5	9	4	7	2	6	3	10
$x^3 - 3x$	0	9	2	7	8	0	0	3	4	9	2
y	0	± 3				0	0	± 5	± 2	± 3	

Le tableau montre que $N_1 = 12$. Dans la formule de Weil, on a $\alpha = 0$, donc :

$$\begin{aligned} \sum_{r \geq 1} N_r X^{r-1} &= \frac{22X}{1+11X^2} + \frac{1}{1-X} + \frac{11}{1-11X} \\ &= 22X(1 - 11X^2 + \dots) + (1 + X + X^2 + \dots) + 11(1 + 11X + 11^2X^2 + \dots) \end{aligned}$$

On en déduit $N_2 = 144$, $N_3 = 1332$. Si $r = 2s + 1$, $N_r = 1 + 11^r$; si $r = 2s$, $N_r = 2 \times (-1)^{s-1} 11^s + 1 + 11^r$.

2) Les points d'ordre 2 de Γ_1 sont donnés par le tableau : (0,0), (5,0), (6,0). Le groupe Γ_1 est abélien d'ordre 12 ; les structures possibles sont $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ et $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$. Seule la première a trois éléments d'ordre 2 ; c'est donc la structure de Γ_1 .

3) a) Que $P(u, v)$ soit d'ordre 3 équivaut à $2P = -P$ avec $P \neq O$. La formule donnant les coordonnées de $2P$ conduit alors à la condition $u^4 - 6u^2 - 3 = 0$.

b) Le **2)** montre que Γ_1 a deux éléments d'ordre 3. La condition précédente s'écrit $(u-2)(u+2)(u^2-2) = 0$; $u^2 = 2$ et $u = 2$ ne donnent aucun point ; $u = -2 = 9$ donne les points $(9, \pm 3)$.

c) Dans K , $u = 2$ donne les points $(2, \pm \sqrt{2})$; $u = \sqrt{2}$ impose $v^2 = -\sqrt{2}$, d'où $v = \pm(2 + 3\sqrt{2})$; de même, $u = -\sqrt{2}$ donne les points $(-\sqrt{2}, \pm(2 - 3\sqrt{2}))$. On trouve huit points d'ordre 3 dans Γ_2 .

4) Si $P(u, v)$ est d'ordre 4, on a $2P = -2P$ et $2P \neq O$, ce qui impose que l'ordonnée de $2P$ soit nulle, soit $v + \frac{3u^2-3}{2v} \left[\left(\frac{3u^2-3}{2v} \right)^2 - 3u \right] = 0$, ce qui se ramène, après calculs, à la condition donnée.

b) La résolution de trois équations du second degré dans K conduit aux 12 points $(2\sqrt{2}, \pm(2 - 3\sqrt{2}))$, $(-2\sqrt{2}, \pm(2 + 3\sqrt{2}))$, $(6 + 6\sqrt{2}, \pm(5 + 5\sqrt{2}))$, $(6 - 6\sqrt{2}, \pm(5 - 5\sqrt{2}))$, $(-6 + 6\sqrt{2}, \pm(4 - 2\sqrt{2}))$, $(-6 - 6\sqrt{2}, \pm(4 + 2\sqrt{2}))$.

c) Comme Γ_2 a 144 éléments, 12 points d'ordre 4 et 8 points d'ordre 3, il ne peut avoir comme structure que celle de $(\mathbb{Z}/4\mathbb{Z})^2 \times (\mathbb{Z}/3\mathbb{Z})^2$, soit celle de $(\mathbb{Z}/12\mathbb{Z})^2$.

6) L'ordre de Γ_3 est $1332 = 4 \times 9 \times 37$. On voit, en résolvant les équations précédentes dans une extension de degré 3 de $\mathbb{F}_{11}$, que ce groupe a trois points d'ordre 2 et deux d'ordre 3. Sa structure est donc celle de $(\mathbb{Z}/2\mathbb{Z})^2 \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/37\mathbb{Z}$; on peut la voir comme produit de deux groupes abéliens, par exemple $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/666\mathbb{Z}$.

FORMES BILINÉAIRES SYMÉTRIQUES ET QUADRATIQUES

L'étude élémentaire des produits scalaires, des espaces euclidiens, des transformations orthogonales a été commencée au chapitre 16. Elle va être poursuivie dans ce chapitre et élargie à des espaces munis de formes bilinéaires symétriques qui ne sont plus nécessairement des produits scalaires. Sauf mention contraire, on suppose que la caractéristique des corps est différente de 2 dans ce chapitre.

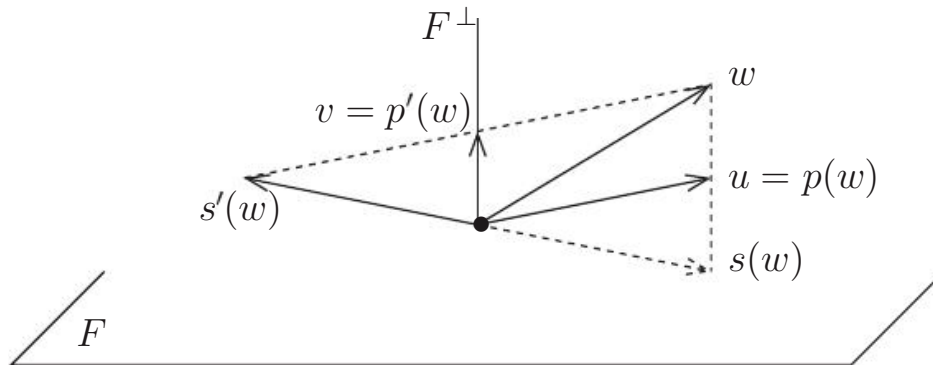
23.1 COMPLÉMENTS SUR LE GROUPE ORTHOGONAL D'UN ESPACE EUCLIDIEN

Soit E un $\mathbb{R}$ -espace vectoriel euclidien. Rappelons (voir 16.3) que cela signifie que E est de dimension finie et qu'un produit scalaire a été défini sur E . L'espace $\mathbb{R}^n$ muni du produit scalaire canonique est un modèle (voir proposition 5 de 16.9) des espaces euclidiens de dimension n ; l'étude générale des espaces euclidiens se ramène donc à l'étude de ces espaces.

Les relations $E = F \oplus F^\perp$, $(F^\perp)^\perp = F$ pour F sous-espace de E ont été démontrées en 16.7, les projections et les symétries orthogonales ont été définies en 16.8, 16.9. Nous avons alors limité l'étude des groupes $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$ aux cas $n = 2$ et $n = 3$ (voir 16.10, 16.11). Nous allons dans cette section faire une étude plus complète.

Symétries orthogonales, retournements. Soit F un sous-espace de E , $F^\perp$ son orthogonal. Comme $E = F \oplus F^\perp$, un vecteur w de E se décompose en une somme $w = u + v$ avec $u \in F$ et $v \in F^\perp$ uniques.

Les projections orthogonales $p : E \rightarrow F$, $p' : E \rightarrow F^\perp$ sont définies par $p(w) = u$, $p'(w) = v$; la symétrie orthogonale $s : E \rightarrow E$ par rapport à F est définie par $s(w) = u - v$, la symétrie orthogonale $s' : E \rightarrow E$ par rapport à $F^\perp$ est définie par $s'(w) = v - u = -s(w)$. Ces applications sont bien sûr linéaires. On peut illustrer la situation par le schéma suivant.



En posant $n = \dim(E)$, $\dim(F) = k$, on peut trouver une base de E dans laquelle la matrice de s est diagonale avec k fois 1 et $n - k$ fois -1 sur la diagonale (il suffit de prendre la réunion d'une base de F et d'une base de $F^\perp$) ; le déterminant de cette matrice est $(-1)^{n-k}$.

Symétries orthogonales par rapport à un hyperplan ou réflexions. Si la dimension de $F^\perp$ est 1, F est un hyperplan de E , s est une symétrie par rapport à cet hyperplan ; on dit souvent que s est une réflexion. On peut trouver une base de E dans laquelle la matrice de s est diagonale avec $n - 1$ fois 1 et une fois -1 sur la diagonale, par conséquent, $s \notin \text{SO}(E)$.

Retournements. Si la dimension de $F^\perp$ est 2, s est appelé un retournement par rapport au sous-espace F . En posant $n = \dim(E)$, F est de dimension $n - 2$ et on peut trouver une base de E dans laquelle la matrice de s est diagonale avec $n - 2$ fois 1 et deux fois -1 sur la diagonale ; le déterminant de cette matrice est 1 ; un retournement est donc dans $\text{SO}(E)$.

Si $n = 2$, il n'existe qu'un retournement : $-\text{id}$; si $n = 3$, un retournement est un demi-tour, autrement dit une rotation d'angle π par rapport à un axe de $\mathbb{R}^3$.

Proposition 1 : centre de $\text{O}(E)$ et de $\text{SO}(E)$. Soit E un espace euclidien ; posons $n = \dim(E)$.

Si $n = 2$, le centre de $\text{O}(E)$ est $\{\text{id}, -\text{id}\}$.

Si $n = 2$, le groupe $\text{SO}(E)$ est commutatif.

Si $n \geq 3$, le centre de $\text{O}(E)$ est $\{\text{id}, -\text{id}\}$.

Si $n \geq 3$ et n pair, le centre de $\text{SO}(E)$ est $\{\text{id}, -\text{id}\}$.

Si $n \geq 3$ et n impair, le centre de $\text{SO}(E)$ est $\{\text{id}\}$.

Démonstration. Dans le cas $n = 2$, les résultats de 16.10 pour l'espace euclidien $\mathbb{R}^2$ se généralisent aux espaces euclidiens de dimension 2, puisqu'ils sont isométriques à $\mathbb{R}^2$ (proposition 5 de 16.9).

Supposons maintenant $n \geq 3$. Un élément f du centre de $O(E)$ ou du centre de $SO(E)$ doit commuter avec tous les retournements.

Soient u un vecteur non nul de E , $v, v' \in E$ tels que u, v, v' soient linéairement indépendants, $F = (\text{Vect}(u, v))^\perp$, $F' = (\text{Vect}(u, v'))^\perp$, s et s' les retournements par rapport aux sous-espaces F et F' respectivement. On doit avoir $f \circ s = s \circ f$ donc $-f(u) = f(s(u)) = s(f(u))$, ce qui montre que $f(u) \in \text{Vect}(u, v)$; de même, on doit avoir $f(u) \in \text{Vect}(u, v')$. On en déduit que $f(u)$ est colinéaire à u et qu'il existe un scalaire λ_u dépendant *a priori* de u , tel que $f(u) = \lambda_u u$. Si u et u' ne sont pas colinéaires, l'égalité $f(u + u') = f(u) + f(u')$ implique $\lambda_{u+u'}(u + u') = \lambda_u u + \lambda_{u'} u'$ donc $\lambda_{u+u'} = \lambda_u = \lambda_{u'}$; par conséquent, f est une homothétie. Dans $O(E)$, les seules homothéties sont $\pm \text{id}$. Dans $SO(E)$, il faut distinguer deux cas; si n est pair, $\pm \text{id} \in SO(E)$ et si n est impair, id seule appartient à $SO(E)$. $\square$

Complexification de $\mathbb{R}^n$. Pour réduire les matrices des endomorphismes orthogonaux de l'espace euclidien $\mathbb{R}^n$ à une forme *presque* diagonale (on va préciser ce que cela veut dire dans l'énoncé de la proposition ci-après), il est intéressant d'approfondir une construction de $\mathbb{C}$ -espace vectoriel associé à un $\mathbb{R}$ -espace vectoriel.

Notons $\text{can}_{\mathbb{R}} = (e_1, \dots, e_n)$ la base canonique du $\mathbb{R}$ -espace vectoriel $\mathbb{R}^n$, $\text{can}_{\mathbb{C}} = (\varepsilon_1, \dots, \varepsilon_n)$ la base canonique du $\mathbb{C}$ -espace vectoriel $\mathbb{C}^n$, $\varphi : \mathbb{R}^n \rightarrow \mathbb{C}^n$ le plongement défini par $\varphi(x_1, \dots, x_n) = (x_1, \dots, x_n)$.

On identifiera dans ce qui suit les éléments de $\mathbb{R}^n$ et leur image par φ dans $\mathbb{C}^n$, comme on identifie $\mathbb{R}$ à une partie de $\mathbb{C}$.

En remarquant que le $\mathbb{C}$ -espace vectoriel $\mathbb{C}$ peut être vu comme un $\mathbb{R}$ -espace vectoriel de dimension 2 de base $(1, \iota) = (e_1, \iota e_1)$, on voit qu'on peut considérer le $\mathbb{C}$ -espace vectoriel $\mathbb{C}^n$ comme un $\mathbb{R}$ -espace vectoriel de dimension $2n$ ayant pour base $(e_1, \iota e_1, \dots, e_n, \iota e_n)$; $\mathbb{R}^n$ s'identifie au $\mathbb{R}$ -sous-espace vectoriel engendré par $(e_1, \dots, e_n)$.

Un vecteur $w = (z_1, \dots, z_n)$ de $\mathbb{C}^n$ avec, pour $1 \leq k \leq n$, $z_k = x_k + \iota y_k$, x_k, y_k réels, s'écrit $w = u + \iota v$ avec : $u = \sum_{1 \leq k \leq n} x_k e_k$, $v = \sum_{1 \leq k \leq n} y_k e_k$ et on a $u, v \in \mathbb{R}^n$.

On définit l'application de conjugaison $c : \mathbb{C}^n \rightarrow \mathbb{C}^n$ en posant $c(w) = (\overline{z_1}, \dots, \overline{z_n})$; l'application c est $\mathbb{R}$ -linéaire, $c(\alpha(u + \iota v)) = \overline{\alpha}(u - \iota v)$ pour $\overline{\alpha} \in \mathbb{C}$ et $c^2 = \text{id}_{\mathbb{C}^n}$.

Si f est un endomorphisme de $\mathbb{R}^n$ de matrice A par rapport à la base canonique, on définit un endomorphisme de $\mathbb{C}^n$ tel que : $M(g, \text{can}_{\mathbb{C}}, \text{can}_{\mathbb{C}}) = A$. On a $g \circ c = c \circ g$. Les polynômes caractéristique de f et g sont tous deux égaux au polynôme caractéristique de A lequel est à coefficients réels. Si λ est une valeur propre non réelle de A , alors $\overline{\lambda}$ est aussi une valeur propre de A de même multiplicité; si $u + \iota v$ est un vecteur propre de g associé à λ , alors $u - \iota v$ est un vecteur propre de g associé à $\overline{\lambda}$ puisque :

$$g(u - \iota v) = g(c(u + \iota v)) = c(g(u + \iota v)) = c(\lambda(u + \iota v)) = \overline{\lambda}(u - \iota v).$$

L'espace vectoriel $\mathbb{C}^n$ est appelé complexifié de $\mathbb{R}^n$. La construction précédente s'étend à tout espace vectoriel réel E en choisissant une base de E . Cette construction du complexifié dépend donc du choix d'une base de E et cela a longtemps troublé les mathématiciens, comme le raconte Marcel Berger, dans le chapitre 7 de son merveilleux livre de géométrie. Il est possible d'éviter ce recours à une base, par exemple en construisant ce qu'on appelle le produit tensoriel $\mathbb{C} \otimes_{\mathbb{R}} E$.

Proposition 2 : Réduction des endomorphismes orthogonaux.

1) Soit $f \in O(\mathbb{R}^n)$. Il existe une base orthonormée de $\mathbb{R}^n$ dans laquelle f a pour matrice une matrice diagonale par blocs de la forme :

$$D = \begin{pmatrix} I_p & & & & \\ & -I_q & & & \\ & & R_1 & & \\ & & & \ddots & \\ & & & & R_r \end{pmatrix}$$

où I_p , I_q sont les matrices unités d'ordre p et q et les R_k $1 \leq k \leq r$ sont des matrices carrées d'ordre 2 de la forme

$$R_k = \begin{pmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{pmatrix}$$

avec $\theta_k \neq 0 \pmod{\pi}$; autrement dit, les R_k sont les matrices de rotation de $SO(\mathbb{R}^2)$ différentes de $\pm I_2$.

2) On a $f \in SO(\mathbb{R}^n)$ si et seulement si n est pair.

3) Soit A une matrice orthogonale d'ordre n . Il existe une matrice orthogonale P telle que ${}^t P A P$ soit une matrice ayant la forme de la matrice D du 1).

Démonstration. 1) Raisonnons par récurrence sur n . La proposition a été montrée pour $n = 2$ et $n = 3$ en 16.11 et 16.12, mais le cas $n = 2$ nous suffit.

Supposons $n \geq 3$ et cherchons un sous-espace F de $\mathbb{R}^n$ distinct de $\{0\}$ et de $\mathbb{R}^n$, stable par f . Une fois cet espace trouvé, comme $F^\perp$ est stable par f et distinct de $\{0\}$ et de $\mathbb{R}^n$, on pourra appliquer l'hypothèse de récurrence aux restrictions de f à F et $F^\perp$ et conclure.

Le polynôme caractéristique $\chi(f)$ de f est à coefficients réels.

Si f admet une valeur propre λ réelle (ce qui est le cas si n est impair), on sait que $\lambda = \pm 1$ (voir 16.9). On prend pour F l'espace propre associé à λ et on peut appliquer l'hypothèse de récurrence.

Si f n'admet pas de valeur propre réelle, on prolonge f en un endomorphisme $\mathbb{C}$ -linéaire g de $\mathbb{C}^n$; l'endomorphisme g admet deux valeurs propres non réelles conjuguées λ et $\bar{\lambda}$. Si $w = u + \iota v$, avec $u, v \in \mathbb{R}^n$ et v non nul, est un vecteur propre associé à λ , $c(w) = u - \iota v$ est un vecteur propre associé à $\bar{\lambda}$. L'intersection du $\mathbb{C}$ -sous-espace vectoriel G engendré par w et $c(w)$ avec $\mathbb{R}^n$ est un $\mathbb{R}$ -espace F de dimension 2 sur $\mathbb{R}$, engendré par $u = \frac{w + c(w)}{2}$ et $v = \frac{w - c(w)}{2\iota}$. On vérifie facilement que F est invariant par f et on peut appliquer l'hypothèse de récurrence.

2) Cela résulte de $\det(D) = (-1)^q$.

3) C'est une traduction du 1) en terme de matrices. □

Topologie sur les groupes $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$. L'espace $M_n(\mathbb{R})$ des matrices carrées d'ordre n est isomorphe à l'espace $\mathbb{R}^{n^2}$ puisque se donner une matrice carrée d'ordre n équivaut à se donner ses n^2 coefficients. Dans cette section, nous munirons cet espace de la distance définie par la norme euclidienne.

On considère le groupe $O(\mathbb{R}^n)$ comme le groupe des matrices orthogonales d'ordre n et le groupe $SO(\mathbb{R}^n)$ comme le sous-groupe du précédent formé des matrices de déterminant 1. Ainsi définis, ces groupes sont des sous-ensembles de $M_n(\mathbb{R})$ qu'on peut munir de la topologie induite.

Proposition 3 : Résultats de topologie sur les groupes $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$. Avec les définitions précédentes, les groupes $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$ ont les propriétés suivantes.

1) Les groupes $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$ sont des compacts de $\mathbb{R}^{n^2}$.

2) Le groupe $SO(\mathbb{R}^n)$ est connexe par arcs.

3) Le groupe $O(\mathbb{R}^n)$ a deux composantes connexes par arcs.

Démonstration. 1) Considérons l'application $M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$ définie par $X \mapsto {}^tXX$. Comme application de $\mathbb{R}^{n^2}$ dans lui-même, cette application a n^2 composantes qui sont toutes des applications polynomiales donc continues. Le groupe $O(\mathbb{R}^n)$ est l'image réciproque par cette application de l'ensemble fermé à un élément défini par la matrice unité I ; c'est donc un fermé de $M_n(\mathbb{R})$. Si $A = (a_{ij})$ est une matrice de $O(\mathbb{R}^n)$, on a $\sum_{1 \leq i \leq n} a_{ij}^2 = 1$ pour $j = 1, \dots, n$, donc $\sum_{1 \leq i, j \leq n} a_{ij}^2 = n$, ce qui montre que $O(\mathbb{R}^n)$ est un sous-ensemble borné de $M_n(\mathbb{R})$. On en déduit que $O(\mathbb{R}^n)$, étant un sous-ensemble fermé borné de $M_n(\mathbb{R})$, est un compact de $M_n(\mathbb{R})$.

L'application $\det : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ est une application continue puisque c'est une application polynomiale (voir la formule de Cramer donnée en 14.1). L'image réciproque du fermé $\{1\}$ est un fermé dont l'intersection avec $O(\mathbb{R}^n)$ est $SO(\mathbb{R}^n)$. On en déduit que $SO(\mathbb{R}^n)$ est compact.

2) Soit A une matrice de $\text{SO}(\mathbb{R}^n)$. Il existe une matrice orthogonale P telle que $P^{-1}AP$ soit une matrice D de la forme donnée dans la proposition précédente, avec q pair ; on pose $q = 2q'$. On peut construire une application continue, appelée chemin, $\gamma : [0, 1] \rightarrow \text{SO}(\mathbb{R}^n)$ d'origine I et d'extrémité D en posant

$$\gamma(t) = D(t) = \begin{pmatrix} I_p & & & & \\ & r_1(t) & & & \\ & & \ddots & & \\ & & & r_{q'}(t) & \\ & & & & R_1(t) & & \\ & & & & & \ddots & \\ & & & & & & R_r(t) \end{pmatrix}$$

où $r_j(t) = \begin{pmatrix} \cos(t\pi) & -\sin(t\pi) \\ \sin(t\pi) & \cos(t\pi) \end{pmatrix}$ pour $j = 1, \dots, q'$

et où $R_k(t) = \begin{pmatrix} \cos(t\theta_k) & -\sin(t\theta_k) \\ \sin(t\theta_k) & \cos(t\theta_k) \end{pmatrix}$ pour $k = 1, \dots, r$.

Pour tout $t \in [0, 1]$, la matrice $PD(t)P^{-1}$ est dans $\text{SO}(\mathbb{R}^n)$. L'application $F : \text{SO}(\mathbb{R}) \rightarrow \text{SO}(\mathbb{R})$ définie par $F(X) = PXP^{-1}$ est une application continue. Le composé $F \circ \gamma$ est donc un chemin de $\text{SO}(\mathbb{R}^n)$ d'origine I et d'extrémité A :

$$\begin{array}{ccc} [0, 1] & \xrightarrow{\gamma} & \text{SO}(\mathbb{R}^n) \\ & \searrow F \circ \gamma & \downarrow F \\ & & \text{SO}(\mathbb{R}^n) \end{array}$$

Par conséquent, il existe des chemins de $\text{SO}(\mathbb{R}^n)$ d'origine I et d'extrémité n'importe quelle matrice de $\text{SO}(\mathbb{R}^n)$. On en déduit la connexité par arcs de $\text{SO}(\mathbb{R}^n)$: si A et A' sont deux matrices de $\text{SO}(\mathbb{R}^n)$, on construit un chemin de A à A' en composant un chemin de A vers I (inverse d'un chemin de I vers A) avec un chemin de I vers A' .

3) La multiplication par une matrice de déterminant -1 (comme, par exemple, la matrice diagonale avec une fois -1 et $n - 1$ fois 1 sur la diagonale) définit un homéomorphisme de $\text{O}(\mathbb{R}^n)$ envoyant les matrices de déterminant $+1$ sur les matrices de

déterminant -1 . Par conséquent, l'ensemble des matrices de déterminant -1 est connexe par arcs.

Enfin, on ne peut relier une matrice de déterminant 1 et une matrice de déterminant -1 de $O(\mathbb{R}^n)$ par un chemin γ de $O(\mathbb{R}^n)$ car, par le théorème des valeurs intermédiaires, la fonction composée $\det \circ \gamma$, qui est continue comme composée de fonctions continues, prendrait la valeur 0 dans $O(\mathbb{R}^n)$, ce qui est impossible. Le groupe $O(\mathbb{R}^n)$ a donc deux composantes connexes par arcs : les matrices de déterminant 1 et celles de déterminant -1 . $\square$

Proposition 4 : générateurs de $O(\mathbb{R}^n)$ et $SO(\mathbb{R}^n)$

- 1) Toute isométrie f de $O(\mathbb{R}^n)$ est un produit de réflexions. Le nombre minimum de réflexions nécessaires est $n - \dim(\ker(f - \text{id}))$, c'est-à-dire $n - p$ avec les notations de la proposition 2.
- 2) Si $n \geq 3$, toute isométrie de $SO(\mathbb{R}^n)$ est un produit de retournements.

Démonstration. Soient f une isométrie de $\mathbb{R}^n$ et $(e_i)_{1 \leq i \leq n}$ une base orthonormée de $\mathbb{R}^n$ par rapport à laquelle la matrice D de f a la forme donnée dans la proposition 2.

La matrice D est le produit de q matrices diagonales M_j , $1 \leq j \leq q$, avec $n - 1$ fois 1 et une fois -1 sur la diagonale pour le coefficient de la ligne $p + j$ et de matrices N_k , $1 \leq k \leq r$, presque diagonales avec $n - 2$ fois 1 et une matrice R_k sur la diagonale placés comme dans D .

- 1) Les matrices M_j sont des matrices de réflexions. Les matrices N_k se décomposent en produit de deux matrices de réflexions d'après le 2) de la proposition 2 de 16.10.

Le nombre de ces réflexions est $q + 2r = n - p$ avec les notations de la proposition 2. On ne peut faire mieux, puisque le produit de réflexions par rapport à k hyperplans fixe leur intersection, laquelle est de dimension supérieure ou égale à $n - k$: si $n - k = p$, on a bien $k = n - p$.

- 2) On suppose maintenant que f est positive, ce qui impose que q soit pair. Le produit de deux matrices M_j est la matrice d'un retournement. Le produit des matrices M_j définit donc un produit de retournements.

On vient de voir que les matrices N_k se décomposent en produit de deux matrices de réflexions : $N_k = S_k S'_k$. Comme $n \geq 3$, il existe un vecteur e_i invariant par ces réflexions. Notons M_i la matrice de la réflexion définie par e_i . On a $N_k = (S_k M_i)(M_i S'_k)$, ce qui montre que N_k est produit de deux matrices de retournements.

Le produit des différents retournements ainsi définis est égal à f . $\square$

23.2 FORMES BILINÉAIRES ET BILINÉAIRES SYMÉTRIQUES

Nous avons déjà rencontré la notion de forme bilinéaire symétrique en 16.2, dans le contexte des produits scalaires, sans la développer. Redonnons les définitions de 16.2.

Définition. Soient K un corps, E un K -espace vectoriel. Une forme bilinéaire sur E est une fonction $\varphi : E \times E \rightarrow K$ vérifiant la propriété dite de bilinéarité :

► pour tous u, u', v, v' de E et tout λ de $\mathbb{R}$:

$$\varphi(u + u', v) = \varphi(u, v) + \varphi(u', v)$$

$$\varphi(u, v + v') = \varphi(u, v) + \varphi(u, v')$$

$$\varphi(\lambda u, v) = \lambda \varphi(u, v) = \varphi(u, \lambda v)$$

La forme bilinéaire φ est dite symétrique si elle vérifie la propriété de symétrie :

► pour tous u, v de E :

$$\varphi(u, v) = \varphi(v, u).$$

Exemples.

- 1) Tous les exemples de produits scalaires du chapitre 16 sont des formes bilinéaires symétriques.
- 2) Sur K^n , on peut définir les formes bilinéaires symétriques

$$(u, v) \mapsto \sum_{1 \leq i \leq r} a_i x_i y_i$$

avec $r \leq n$, $u = (x_1, \dots, x_n)$, $v = (y_1, \dots, y_n)$ et les a_i dans K .

D'autres exemples seront donnés dans ce qui suit.

Propriétés d'une forme bilinéaire symétrique. Soient K un corps, E un K -espace vectoriel, φ une forme bilinéaire symétrique sur E . Les propriétés de φ impliquent les formules :

$$\varphi(u + v, u + v) = \varphi(u, u) + 2\varphi(u, v) + \varphi(v, v)$$

$$\varphi(u - v, u - v) = \varphi(u, u) - 2\varphi(u, v) + \varphi(v, v)$$

$$\varphi(u, v) = \frac{1}{2} (\varphi(u + v, u + v) - \varphi(u, u) - \varphi(v, v))$$

$$\varphi(u, v) = \frac{1}{4} (\varphi(u + v, u + v) - \varphi(u - v, u - v)).$$

Si on demande dans ce chapitre que les corps soient de caractéristique différente de 2, c'est pour pouvoir faire ces divisions par 2.

Expression d'une forme bilinéaire dans une base. Soient K un corps, E un K -espace vectoriel de dimension finie n , $B = (e_1, \dots, e_n)$ une base de E , $\varphi : E \times E \rightarrow K$ une forme bilinéaire sur E . Si $u = x_1 e_1 + \dots + x_n e_n$ et $v = y_1 e_1 + \dots + y_n e_n$ sont deux vecteurs de E , la bilinéarité de φ donne :

$$\varphi(u, v) = \sum_{1 \leq i, j \leq n} x_i y_j \varphi(e_i, e_j).$$

La forme φ est donc entièrement déterminée par les n^2 valeurs $\varphi(e_i, e_j)$, $1 \leq i, j \leq n$. Si φ est symétrique, elle est déterminée par les $n(n+1)/2$ valeurs $\varphi(e_i, e_j)$, $1 \leq i \leq j \leq n$, puisque $\varphi(e_i, e_j) = \varphi(e_j, e_i)$ pour $i > j$ et on a :

$$\varphi(u, v) = \sum_{1 \leq i \leq n} x_i y_i \varphi(e_i, e_i) + 2 \sum_{1 \leq i < j \leq n} x_i y_j \varphi(e_i, e_j).$$

Matrice d'une forme bilinéaire dans une base. Avec les mêmes notations, la donnée de φ définit la matrice carrée d'ordre n : $A = (a_{ij}) = (\varphi(e_i, e_j))$ dite matrice associée à la forme bilinéaire φ dans la base B . Si u et v sont deux vecteurs de E et si U et V sont les matrices colonnes des coordonnées de u et de v dans la base B , on vérifie que

$$\varphi(u, v) = {}^t U A V,$$

en identifiant $\varphi(u, v)$ à la matrice ${}^t U A V$ qui n'a qu'une seule ligne et une seule colonne.

Si la forme bilinéaire est symétrique, la matrice A est symétrique et on a

$$\varphi(u, v) = {}^t U A V = {}^t V A U = \varphi(v, u).$$

Inversement, la base B étant fixée, toute matrice A carrée d'ordre n définit une forme bilinéaire φ sur E dont la matrice dans la base B est A : il suffit de poser $\varphi(u, v) = {}^t U A V$; la forme bilinéaire est symétrique si et seulement si A est symétrique.

Changement de base. La valeur de $\varphi(u, v)$ ne change pas quand on change de base pour exprimer u et v .

Avec les mêmes notations que précédemment, soient B' une autre base de E , $P = M(\text{id}, B', B)$ la matrice de passage. Si on note U' et V' les matrices colonnes des coordonnées de u et v dans la base B' , on a $U = P U'$, $V = P V'$ d'où $\varphi(u, v) = {}^t U A V = {}^t (P U') A P V = {}^t U' ({}^t P A P) V'$. Ainsi, la matrice associée à φ dans la base B' est :

$${}^t P A P.$$

On remarque la différence avec la formule $P^{-1} A P$ de 8.9 où il s'agissait de changement de base pour la matrice d'une application linéaire.

Rang d'une forme bilinéaire symétrique. Les matrices A et ${}^t P A P$ associées à φ dans les bases B et B' ont même rang puisque P est inversible. Par conséquent, le rang de la matrice associée à une forme bilinéaire symétrique ne change pas quand on change de base. On définit le rang d'une forme bilinéaire symétrique comme le rang de la matrice associée à φ dans une base quelconque de E .

On peut également remarquer que le déterminant de la matrice associée à φ change d'un carré de K suivant la base, puisque $\det({}^t P A P) = (\det(P))^2 \det(A)$.

Exemple (plan hyperbolique). Plaçons-nous dans $\mathbb{R}^2$ muni de la base canonique $B = (e_1, e_2)$. La forme bilinéaire symétrique $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $\varphi(e_1, e_1) = \varphi(e_2, e_2) = 0$, $\varphi(e_1, e_2) = \varphi(e_2, e_1) = 1$, a pour matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans B .

Si $u = (x_1, x_2)$, on a $\varphi(u, u) = 2x_1 x_2$ et si $v = (y_1, y_2)$, on a $\varphi(u, v) = x_1 y_2 + x_2 y_1$.

Considérons la base $B' = (e'_1, e'_2)$, définie par

$$\begin{cases} e'_1 &= e_1 \\ e'_2 &= e_2/2 \end{cases}$$

La formule précédente donne la matrice A' de φ dans la base B' :

$$A' = {}^t P A P = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix},$$

Si on note avec des $'$ les coordonnées de u et v dans la base B' , il est plus facile ici de substituer $x_1 = x'_1$, $x_2 = x'_2 / 2$, etc. dans les formules donnant φ pour obtenir direc-

tement $\varphi(u, u) = x'_1 x'_2$, $\varphi(u, v) = \frac{1}{2}(x'_1 y'_2 + x'_2 y'_1)$.

L'expression de φ dans la base $B'' = (e''_1, e''_2)$ définie par

$$\begin{cases} e''_1 &= e'_1 + e'_2 \\ e''_2 &= e'_1 - e'_2 \end{cases}$$

est donnée par la matrice :

$$A'' = {}^t P A P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

soit $\varphi(u, u) = x''_1{}^2 - x''_2{}^2$, $\varphi(u, v) = x''_1 y''_1 - x''_2 y''_2$, en notant avec des $''$ les coordonnées de u, v dans la base B'' . On peut simplement obtenir ces expressions en substituant $x'_1 = x''_1 + x''_2$, $x'_2 = x''_1 - x''_2$, etc. dans les formules précédentes.

Cet exemple est à la base de la notion de plan hyperbolique étudiée en 23.8.

Formes équivalentes. Soient K un corps, E un K -espace vectoriel, φ_1 et φ_2 deux formes bilinéaires symétriques sur E . Les formes φ_1 et φ_2 sont dites équivalentes s'il existe un isomorphisme $f: E \rightarrow E$ tel que, pour tous $u, v \in E$:

$$\varphi_2(u, v) = \varphi_1(f(u), f(v)).$$

Cette relation est bien une relation d'équivalence.

On a $\text{im}(\varphi_1) = \text{im}(\varphi_2)$; l'ensemble des valeurs de φ_1 est égal à l'ensemble des valeurs de φ_2 .

Soient B une base de E et les matrices A_1 de φ_1 , A_2 de φ_2 , P de f par rapport à la base B . Pour tous u, v de E , de matrices colonnes U, V par rapport à la base B , on a $\varphi_2(u, v) = {}^t U A_2 V$ et $\varphi_1(f(u), f(v)) = {}^t (PU) A_1 (PV)$; on a donc $A_2 = {}^t P A_1 P$. Notons $B' = f(B)$ la base image de B par f ; la matrice de passage $M(\text{id}, B', B)$ est égale à P . La situation est décrite par le diagramme commutatif suivant où sont indiquées les applications, les bases et les matrices.

$$\begin{array}{ccc} (E, B) & \xrightarrow{(f, I)} & (E, B') \\ & \searrow (f, P) & \downarrow (\text{id}, P) \\ & & (E, B) \end{array}$$

La matrice de φ_1 dans la base B' est égale à la matrice de φ_2 dans la base B , comme le résume le tableau suivant.

	u	v	$f(u)$	$f(v)$	φ_1	φ_2
matrice dans B	U	V	PU	PV	A_1	$A_2 = {}^t P A_1 P$
matrice dans B'			U	V	${}^t P A_1 P$	

Réciproquement, si deux formes bilinéaires symétriques φ_2 et φ_1 sur E ont même matrice A par rapport à des bases $B = (e_1, \dots, e_n)$ et $B' = (e'_1, \dots, e'_n)$ respectivement, elles sont équivalentes car, si on note f l'endomorphisme de E défini par $f(e_i) = e'_i$ pour $1 \leq i \leq n$, les matrices colonnes U et V associées à u et v dans la base B sont les matrices colonnes de $f(u)$ et $f(v)$ dans la base B' et on a $\varphi_2(u, v) = {}^t U A V = \varphi_1(f(u), f(v))$.

La détermination des classes de formes équivalentes est un problème important de la théorie des formes bilinéaires symétriques.

Forme bilinéaire symétrique et dualité. Soient K un corps, E un K -espace vectoriel de dimension finie n muni d'une base $B = (e_1, \dots, e_n)$, φ une forme bilinéaire symétrique sur E de matrice $A = (a_{ij})$ par rapport à B .

On définit une application linéaire $\psi : E \rightarrow E^*$ en associant à tout $u \in E$, la forme linéaire $\psi(u) : v \mapsto \varphi(u, v)$ (les vérifications de linéarité sont faciles).

Si on munit E^* de la base duale B^* de B (voir 10.3), on a :

$$M(\psi, B, B^*) = A.$$

En effet, d'une part, pour $i, j = 1, \dots, n$, on a $\psi(e_j)(e_i) = \varphi(e_j, e_i) = a_{ij}$; d'autre part, si on écrit $\psi(e_j) = \sum_{1 \leq k \leq n} s_{kj} e_k^*$, on a

$$a_{ij} = \psi(e_j)(e_i) = \sum_{1 \leq k \leq n} s_{kj} e_k^*(e_i) = s_{ij}.$$

La correspondance entre φ et ψ est donc bijective.

Le rang de la forme φ est égal au rang de l'application linéaire ψ .

La controverse entre Jordan et Kronecker. Au lendemain de la guerre de 1870, une vive controverse scientifique oppose Camille Jordan (voir 3.1) et Leopold Kronecker (1823-1891) ; l'un est à Paris, l'autre à Berlin. Ils n'envisagent pas de la même façon le traitement des polynômes bilinéaires (la notion de forme bilinéaire n'est pas encore dégagée et ils travaillent sur les expressions de ces formes) pour en vérifier l'équivalence au sens ci-dessus. Pour Kronecker, les résultats de Jordan ne font que reprendre sous une autre forme et en moins bien ceux que Karl Weierstrass (1815-1897) et lui ont obtenu à Berlin. En réponse, Jordan explique que ses résultats ont une généralité qui fait défaut à ceux de ses confrères allemands ; mais Kronecker n'a pas la même idée de la notion de généralité. La querelle dure pendant les années 1873-1874, chacun publiant sa réponse à l'attaque de l'adversaire. À court terme, les idées de Kronecker semblent s'imposer (recherches de quantités dont l'invariance garantit l'équivalence des formes) ; dans les années 1930, le développement des formalismes de l'algèbre linéaires montre la profondeur des idées de Jordan.

23.3 FORMES QUADRATIQUES

Forme bilinéaire symétrique et forme quadratique associée. Soient K un corps (toujours avec $\text{car}(K) \neq 2$) et E un K -espace vectoriel. À une forme bilinéaire symétrique $\varphi : E \times E \rightarrow K$, on associe l'application $q : E \rightarrow K$ appelée forme quadratique sur E , définie par $q(u) = \varphi(u, u)$.

Une forme quadratique se définit donc comme associée à une forme bilinéaire symétrique. Nous verrons plus loin comment définir une forme quadratique par un polynôme homogène de degré 2.

On peut retrouver φ à partir de q avec les formules données en 23.2 :

$$\begin{aligned} \varphi(u, v) &= \frac{1}{2}(q(u+v) - q(u) - q(v)); \\ \varphi(u, v) &= \frac{1}{4}(q(u+v) - q(u-v)). \end{aligned}$$

Il est donc équivalent de se donner une forme quadratique ou une forme bilinéaire symétrique sur un espace.

La forme bilinéaire symétrique associée à une forme quadratique par l'une des formules précédentes est appelée forme polaire associée à la forme quadratique.

Certains auteurs notent avec la même lettre une forme bilinéaire symétrique et la forme quadratique associée, écrivant, par exemple, $q(u)$ et $q(u, v)$, le nombre d'arguments permettant seul de décider ce qui est exactement désigné par q .

Expression d'une forme quadratique dans une base. Soient K un corps, E un K -espace vectoriel de dimension n , $B = (e_1, \dots, e_n)$ une base de E , $\varphi : E \times E \rightarrow K$ une forme bilinéaire symétrique de forme quadratique associée q . Si, pour $u = \sum_{1 \leq i \leq n} x_i e_i, v = \sum_{1 \leq j \leq n} y_j e_j$, on a $\varphi(u, v) = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j$, alors :

$$q(u) = \varphi(u, u) = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j = \sum_{1 \leq i \leq n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j.$$

Par exemple, la forme bilinéaire de l'exemple de 23.2, définie sur $\mathbb{R}^2$ muni de la base canonique $B = (e_1, e_2)$ par $\varphi(u, v) = x_1 y_2 + x_2 y_1$ donne la forme quadratique définie par $q(u) = 2x_1 x_2$.

Soit $A = (a_{ij})$ la matrice associée à φ dans la base B ; pour tout vecteur u de E de matrice colonne U dans la base B , on a :

$$q(u) = {}^t U A U.$$

La matrice A est appelée matrice associée à la forme quadratique q dans la base B .

Formes quadratiques et polynômes homogènes de degré 2. Les formules précédentes montrent que q est une fonction polynomiale homogène de degré 2 en les coordonnées par rapport à la base B et on a la propriété d'homogénéité : pour tout $\alpha \in K$ et tout $u \in E$:

$$q(\alpha u) = \alpha^2 q(u) ;$$

en particulier, $q(-u) = q(u)$.

Réciproquement, si E un espace vectoriel de dimension n sur un corps K muni d'une base $B = (e_1, \dots, e_n)$ et si

$$q : u = (x_1, \dots, x_n) \mapsto \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$$

est une fonction polynomiale homogène de degré 2 en les coordonnées par rapport à la base B , q est l'expression dans B de la forme quadratique associée à la forme bilinéaire symétrique définie par $\varphi(e_i, e_j) = \varphi(e_j, e_i) = \frac{a_{ij}}{2}$, pour $1 \leq i < j \leq n$, et, pour $1 \leq i \leq n$, $\varphi(e_i, e_i) = a_{ii}$.

On peut donc définir une forme quadratique indépendamment d'une forme bilinéaire symétrique et appeler forme quadratique une fonction polynomiale

homogène de degré 2 en n variables ; c'est ainsi qu'on a d'abord défini les formes quadratiques.

Écriture matricielle, changement de base. La formule de changement de base est la même que pour les formes bilinéaires symétriques ; si B' est une seconde base de E et si $P = M(\text{id}, B', B)$, on a :

$$q(u) = {}^t U A U = {}^t (P U') A P U = {}^t U' ({}^t P A P) U',$$

et la matrice associée à q dans la base B' est ${}^t P A P$.

En reprenant le même exemple qu'en 23.2, la forme quadratique définie sur $\mathbb{R}^2$ par $2x_1 x_2$ dans la base B , a pour expressions $x'_1 x'_2$ dans la base B' et $x''_1{}^2 - x''_2{}^2$ dans la base B'' .

Rang d'une forme quadratique. Le rang d'une forme quadratique est le rang de la forme polaire associée, c'est-à-dire le rang de n'importe quelle matrice associée à q .

En particulier, si $q(x_1, \dots, x_n) = \sum_{1 \leq i \leq n} a_i x_i^2$, la matrice de q est diagonale et le rang de q est le nombre de coefficients a_i non nuls.

Espace quadratique. Soit K un corps. Un couple (E, q) formé d'un K -espace vectoriel et d'une forme quadratique sur E est appelé K -espace quadratique et nous oublierons en général le K . On utilise le même mot pour désigner un couple (E, φ) formé d'un K -espace vectoriel et d'une forme bilinéaire symétrique sur E , ce qui ne présente aucun inconvénient puisqu'il est équivalent de se donner une forme quadratique ou une forme bilinéaire symétrique sur E (on a supposé la caractéristique de K différente de 2).

Si F est un sous-espace de E , la restriction $q|_F$ de la forme quadratique q à F (ou de la forme φ à $F \times F$) définit un espace quadratique $(F, q|_F)$.

23.4 MÉTHODE DE GAUSS POUR LA DÉCOMPOSITION EN CARRÉS

Soit q une forme quadratique sur un corps K . Nous allons présenter une méthode, due à Gauss, permettant de trouver une base dans laquelle la matrice de q soit diagonale, donc dans laquelle l'expression de q soit de la forme $\sum_{1 \leq i \leq n} a_i x_i^2$.

Proposition. Soit K un corps (on suppose toujours $\text{car}(K) \neq 2$). Pour tout entier $n \geq 0$ et toute forme quadratique q sur K^n : $q(x_1, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$, il existe des éléments $\alpha_1, \dots, \alpha_n$ de K et des formes linéaires linéairement indépendantes sur K : $l_1, \dots, l_n$ telles que :

$$q(x_1, \dots, x_n) = \sum_{1 \leq k \leq n} \alpha_k (l_k(x_1, \dots, x_n))^2.$$

Les vecteurs de la base $B' = (e'_1, \dots, e'_n)$, définie par le changement de variables $x'_k = l_k(x_1, \dots, x_n)$ pour $1 \leq k \leq n$, sont orthogonaux (voir 23.7) entre eux deux à deux pour q et, si $u = \sum_{1 \leq i \leq n} x'_i e'_i$, alors $q(u) = q(x'_1, \dots, x'_n) = \sum_{1 \leq k \leq n} \alpha_k x'^2_k$.

Démonstration. On raisonne par récurrence. Pour $n = 0$ ou $n = 1$, le résultat est vrai. Soit $n \geq 2$ un entier et soit q une forme quadratique sur K^n définie par un polynôme homogène $q(x_1, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} \alpha_{ij} x_i x_j$. Plusieurs cas peuvent se présenter.

- 1) Il existe i tel que $a_{ii} \neq 0$, disons $a_{11} \neq 0$. On considère les termes de $q(x_1, \dots, x_n)$ où apparaît x_1 : $a_{11}x_1^2 + \sum_{2 \leq i \leq n} \alpha_{1i} x_1 x_i$. C'est le début du développement du carré : $a_{11}(x_1 + \sum_{2 \leq i \leq n} \frac{\alpha_{1i}}{2a_{11}} x_i)^2 = a_{11}(l_1(x_1, \dots, x_n))^2$ où $l_1(x_1, \dots, x_n)$ est une

forme linéaire dont le terme en x_1 est non nul (ça nous rappelle la mise sous forme canonique d'un trinôme du second degré au lycée). En regroupant les termes en $x_2, \dots, x_n$, on obtient :

$$q(x_1, \dots, x_n) = a_{11}(l_1(x_1, \dots, x_n))^2 + q'(x_2, \dots, x_n)$$

et q' est une forme quadratique sur K^{n-1} à laquelle on peut appliquer l'hypothèse de récurrence : $q'(x_2, \dots, x_n) = \sum_{2 \leq k \leq n} \alpha_k (l_k(x_2, \dots, x_n))^2$; les formes linéaires l_k , $1 \leq k \leq n$ sont définissables sur K^n et linéairement indépendantes sur K .

- 2) Il n'existe pas d'entier i tel que $a_{ii} \neq 0$, mais q n'est pas nulle ; l'un des coefficients a_{ij} , $i \neq j$, n'est pas nul, par exemple, $a_{12} \neq 0$. On considère les termes de $q(x_1, \dots, x_n)$ où apparaissent x_1 ou x_2 : $a_{12}x_1x_2 + \sum_{3 \leq i \leq n} a_{1i}x_1x_i + \sum_{3 \leq i \leq n} a_{2i}x_2x_i$. Cette somme est de la forme $a_{12}x_1x_2 + Bx_1 + Cx_2$ (avec B et C formes linéaires), début du développement de $a_{12} \left(x_1 + \frac{C}{a_{12}} \right) \left(x_2 + \frac{B}{a_{12}} \right)$. En regroupant les termes en $x_3, \dots, x_n$, on obtient :

$$q(x_1, \dots, x_n) = a_{12} \left(x_1 + \frac{C}{a_{12}} \right) \left(x_2 + \frac{B}{a_{12}} \right) + q'(x_3, \dots, x_n)$$

où q' est une forme quadratique sur K^{n-2} à laquelle on peut appliquer l'hypothèse de récurrence : $q'(x_3, \dots, x_n) = \sum_{3 \leq k \leq n} \alpha_k (l_k(x_3, \dots, x_n))^2$. On transforme le premier produit à l'aide de la formule $xy = \frac{1}{4}[(x+y)^2 - (x-y)^2]$, ce qui conduit à :

$$\begin{aligned} q(x_1, \dots, x_n) &= \frac{a_{12}}{4} (x_1 + x_2 + l'_1(x_3, \dots, x_n))^2 - \frac{a_{12}}{4} (x_1 - x_2 + l'_2(x_3, \dots, x_n))^2 \\ &\quad + \sum_{3 \leq k \leq n} \alpha_k (l_k(x_3, \dots, x_n))^2. \end{aligned}$$

En posant :

$$l_1(x_1, \dots, x_n) = x_1 + x_2 + l'_1(x_3, \dots, x_n)$$

$$l_2(x_1, \dots, x_n) = x_1 - x_2 + l'_2(x_3, \dots, x_n),$$

on vérifie que les formes linéaires l_k , $1 \leq k \leq n$ sont définissables sur K^n et linéairement indépendantes sur K .

3) Dans le cas où $q = 0$, on peut poser $l_k(x_1, \dots, x_n) = x_k$ pour $1 \leq k \leq n$.

4) **Changement de base.** Comme les formes linéaires l_k , $1 \leq k \leq n$ sont indépendantes, elles définissent un changement de base. Les nouvelles coordonnées sont définies en fonction des anciennes par $x'_k = l_k(x_1, \dots, x_n)$ pour $1 \leq k \leq n$. On peut écrire ces équations $X' = QX$ où Q est la matrice de passage de $\text{id} : (K^n, B) \rightarrow (K^n, B')$, soit $Q = M(\text{id}, B, B')$. Si on note D la matrice dont la diagonale a pour coefficients $\alpha_1, \dots, \alpha_n$, on a ${}^t Q D Q = A$. Les vecteurs de B' sont les colonnes de la matrice $P = Q^{-1} = M(\text{id}, B', B)$ et on a ${}^t P A P = D$.

5) La matrice de q dans la base B' étant diagonale, les vecteurs de B' sont q -orthogonaux deux à deux et l'expression de q dans la base B' est $q(x'_1, \dots, x'_n) = \sum_{1 \leq k \leq n} \alpha_k x'^2_k$. $\square$

Exemple. Considérons la forme quadratique

$$q(x, y, z, t) = x^2 + 9y^2 + 4z^2 + 6xy + 4xz + 16yz + 4yt + 8zt$$

définie sur $\mathbb{R}^4$ dont l'écriture matricielle dans la base canonique est ${}^t U A U$ où U est la matrice colonne définie par $u = (x, y, z, t)$ et où :

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 \\ 3 & 9 & 8 & 2 \\ 2 & 8 & 4 & 4 \\ 0 & 2 & 4 & 0 \end{pmatrix}$$

On applique le 1) en choisissant la variable x , ce qui conduit à considérer tous les termes où apparaît x : $x^2 + 6xy + 4xz$ et à introduire la forme linéaire donnée par $x + 3y + 2z$. On obtient $q(x, y, z, t) = (x + 3y + 2z)^2 + 4yz + 4yt + 8zt$. On applique alors le 2) à $4yz + 4yt + 8zt$ en choisissant le terme en yz : $4yz + 4yt + 8zt = 4(y + 2t)(z + t) - 8t^2$. Finalement,

$$q(x, y, z, t) = (x + 3y + 2z)^2 + (y + z + 3t)^2 - (y - z + t)^2 - 8t^2.$$

Les quatre formes linéaires sont

$$l_1 : (x, y, z, t) \mapsto x + 3y + 2z, \quad l_2 : (x, y, z, t) \mapsto y + z + 3t,$$

$$l_3 : (x, y, z, t) \mapsto y - z + t, \quad l_4 : (x, y, z, t) \mapsto t.$$

Elles sont linéairement indépendantes sur K . Le changement de variable est défini par $U' = QU$ avec

$$Q = \begin{pmatrix} 1 & 3 & 2 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

et ${}^tQAQ = D$ avec

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -8 \end{pmatrix}$$

On calcule alors (le mieux est sans doute ici de résoudre un système linéaire)

$$P = Q^{-1} = -\frac{1}{2} \begin{pmatrix} -2 & 5 & 1 & -16 \\ 0 & -1 & -1 & 4 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

et on obtient ${}^tPAP = D$.

Décomposition en carrés par diagonalisation. Plaçons-nous maintenant dans le cas $K = \mathbb{R}$ et supposons $\mathbb{R}^n$ muni du produit scalaire défini par la base canonique B de $\mathbb{R}^n$. Soit q une forme quadratique sur $\mathbb{R}^n$, de matrice A par rapport à B . On a vu en 16.12.2 qu'il existe une base orthonormale B' telle que, si $P = M(\text{id}, B', B)$, la matrice $D = {}^tPAP = P^{-1}AP$ est diagonale, ce qui permet d'écrire q dans la base B' sous forme de combinaison linéaire de carrés comme la méthode de Gauss.

Le désavantage de cette méthode par rapport à la méthode de Gauss est qu'elle demande des calculs beaucoup plus compliqués. Elle apporte néanmoins des informations précieuses car la matrice de passage P est cette fois orthogonale (donc ${}^tPP = I$) et le calcul du produit scalaire est le même dans B et dans B' puisque si u, v sont deux vecteurs de matrices colonnes U, V par rapport à B , leurs matrices par rapport à B' sont PU et PV et on a $\langle u, v \rangle = {}^tUV = {}^tU{}^tP PV = {}^t(PU)PV$. Les calculs de longueurs ou d'angles sont donc les mêmes dans les deux bases.

Il s'agit là d'un cas particulier de diagonalisation simultanée (voir 23.7).

23.5 DÉCOMPOSITION D'UNE FORME QUADRATIQUE SUR $\mathbb{C}$ OU $\mathbb{R}$

Soient n un entier et q une forme quadratique sur $\mathbb{C}^n$ ou $\mathbb{R}^n$. On vient de voir que q peut s'écrire dans une base $B = (e_1, \dots, e_n)$ convenable sous la forme

$q(x_1, \dots, x_n) = \sum_{1 \leq k \leq n} a_k x_k^2$. Si le rang de q est r , on peut supposer $a_1, \dots, a_r$ non nuls et on a :

$$q(x_1, \dots, x_n) = \sum_{1 \leq k \leq r} a_k x_k^2.$$

Cas d'une forme quadratique sur $\mathbb{C}$. Dans le cas où le corps de base est $\mathbb{C}$, comme tout nombre complexe est un carré, il existe pour $k = 1, \dots, r$, b_k tel que $b_k^2 = a_k$. Dans la base définie par $e'_k = e_k/b_k$ pour $k = 1, \dots, r$ et $e'_k = e_k$ sinon, la forme quadratique q a pour expression :

$$q(x'_1, \dots, x'_n) = \sum_{1 \leq k \leq r} x_k'^2$$

et sa matrice est

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

avec le cas particulier où $r = 0$ dans lequel I_r n'apparaît pas.

Cas d'une forme quadratique sur $\mathbb{R}$. Dans le cas où le corps de base est $\mathbb{R}$, les coefficients a_k pour $k = 1, \dots, r$, peuvent être positifs ou négatifs. On peut supposer que $a_k > 0$ pour $k \leq s$ et $a_k < 0$ si $s < k \leq r = s + t$ (on a éventuellement $s = 0$ ou $t = 0$). En choisissant b_k tel que $b_k^2 = a_k$ si $a_k > 0$ et $b_k^2 = -a_k$ si $a_k < 0$ et en définissant la base B' par $e'_k = e_k/b_k$ pour $k = 1, \dots, r$ et $e'_k = e_k$ sinon, la forme quadratique q a pour expression :

$$q(x'_1, \dots, x'_n) = \sum_{1 \leq k \leq s} x_k'^2 - \sum_{s < k \leq s+t} x_k'^2$$

et sa matrice est

$$\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

avec les cas particuliers où $s = 0$ ou $t = 0$ dans lesquels I_s ou I_t n'apparaissent pas.

Signature d'une forme quadratique sur $\mathbb{R}$. On vient de voir qu'une forme quadratique sur $\mathbb{R}^n$ peut s'écrire sous la forme $q(x_1, \dots, x_n) = \sum_{1 \leq k \leq s} x_k^2 - \sum_{s+1 \leq k \leq s+t} x_k^2$. Le rang de la forme q est $s + t$. On va montrer que les entiers s et t donnant le nombre de carrés avec un signe $+$ et le nombre de carrés avec un signe $-$ dans l'écriture de q ne dépendent pas de la base dans laquelle on a réussi à écrire q sous la forme précédente. Le couple (s, t) s'appelle signature de la forme q .

Supposons que q s'écrive par rapport à une base $B' = (e'_1, \dots, e'_n)$ sous la forme $q(x'_1, \dots, x'_n) = \sum_{1 \leq k \leq s'} x_k'^2 - \sum_{s'+1 \leq k \leq s'+t'} x_k'^2$. On a $s + t = s' + t'$, le rang de q .

Posons $F = \text{Vect}(e_1, \dots, e_s)$ et $G = \text{Vect}(e'_{s'+1}, \dots, e'_n)$. Comme q est strictement positive sur $F - \{0\}$ et négative sur G , on a $F \cap G = \{0\}$; d'où $s \leq s'$; par symétrie, on a aussi $s' \leq s$; d'où $s = s'$ et $t = t'$.

Classes d'équivalence de formes quadratiques sur $\mathbb{C}$. La décomposition de Gauss en sommes de carrés montre que, pour toute forme quadratique de rang r sur $\mathbb{C}^n$, il

existe une base dans laquelle la matrice de q est de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. D'autre

part, deux formes quadratiques équivalentes ont même rang. Donc les classes d'équivalence de formes quadratiques sur $\mathbb{C}^n$ (ou sur un $\mathbb{C}$ -espace vectoriel de dimension n) correspondent aux entiers r tels que $0 \leq r \leq n$. Il existe donc $n + 1$ classes d'équivalence de formes quadratiques sur $\mathbb{C}^n$.

Classes d'équivalence de formes quadratiques sur $\mathbb{R}$. Dans ce cas, la décomposition de Gauss en sommes de carrés montre que, pour toute forme quadratique de signature (s, t) sur $\mathbb{R}^n$, il existe une base dans laquelle la matrice de q est de la forme

$\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$. D'autre part, deux formes quadratiques équivalentes ont même

signature. Donc les classes d'équivalence de formes quadratiques sur $\mathbb{R}^n$ (ou sur un $\mathbb{R}$ -espace vectoriel de dimension n) correspondent aux couples d'entiers (s, t) tels que $0 \leq s, t \leq n$ et $s + t \leq n$; on les dénombre facilement en remarquant que pour

$s = n, \dots, 0$, il existe $1, 2, \dots, n + 1$ valeurs de t possibles. Il existe donc $\frac{(n+1)(n+2)}{2}$ classes d'équivalence de formes quadratiques sur $\mathbb{R}^n$.

23.6 DIAGONALISATION SIMULTANÉE DE DEUX FORMES QUADRATIQUES

Proposition 1. Soient K un corps, E un K -espace vectoriel, B une base de E , φ_1, φ_2 deux formes bilinéaires symétriques sur E , φ_1 étant non dégénérée. On note A_1, A_2 les matrices associées à φ_1, φ_2 dans la base B ; A_1 est inversible. Alors, il existe une base de E orthogonale pour φ_1 et φ_2 simultanément si et seulement si la matrice $A_1^{-1}A_2$ est diagonalisable.

Démonstration. 1) Supposons d'abord $A_1^{-1}A_2$ diagonalisable et notons E_k les espaces propres de $A_1^{-1}A_2$ associés aux différentes valeurs propres. Dans chaque E_k , il existe une base orthogonale pour φ_1 ; notons C la base de E réunion de ces bases. Soient u et v deux vecteurs de C , U, V les matrices colonnes les représentant dans B .

Si u et v sont associés à la même valeur propre λ , u et v sont orthogonaux pour φ_1 par construction et pour φ_2 car

$$\varphi_2(u, v) = {}^t U A_2 V = {}^t U A_1 A_1^{-1} A_2 V = {}^t U A_1 (\lambda V) = \lambda \varphi_1(u, v) = 0.$$

Si u et v appartiennent à des espaces propres distincts, associés aux valeurs propres λ et μ respectivement, on a :

$$\varphi_2(u, v) = {}^t U A_2 V = {}^t U A_1 A_1^{-1} A_2 V = {}^t U A_1 \mu V = \mu \varphi_1(u, v).$$

De même, $\varphi_2(v, u) = \lambda \varphi_1(v, u)$. On en déduit que $\lambda \varphi_1(u, v) = \mu \varphi_1(u, v)$, puis que $\varphi_1(u, v) = \varphi_2(u, v) = 0$.

La base C est donc orthogonale à la fois pour φ_1 et pour φ_2 .

2) Supposons qu'il existe une base C de E orthogonale pour φ_1 et φ_2 simultanément. Soient $u \in C$, U la matrice colonne représentant u dans la base B , $W = A_1^{-1} A_2 U$ et w le vecteur représenté par W dans B . Pour tout v de C distinct de u , représenté par la matrice colonne V dans B , on a $\varphi_1(w, v) = {}^t W A_1 V = {}^t U A_2 A_1^{-1} A_1 V = \varphi_2(u, v) = 0$. On en déduit que w est orthogonal pour φ_1 à tous les vecteurs autres que u de la base C ; il est donc colinéaire à u ; on en déduit que $A_1^{-1} A_2$ est diagonalisable. $\square$

Pour la recherche des valeurs propres et vecteurs propres de $A_1^{-1} A_2$, on peut remarquer que $A_1^{-1} A_2 U = \lambda U$ équivaut à $(A_2 - \lambda A_1)U = 0$; on peut déterminer λ et U à partir de cette dernière équation, ce qui évite le calcul de l'inverse de A_1 .

Proposition 2. Soient E un espace euclidien, φ une forme bilinéaire symétrique sur E . Alors, il existe une base orthogonale de E qui est également orthogonale pour φ .

Démonstration. Soient B une base orthogonale de E et A la matrice de φ dans la base B . Comme A est une matrice symétrique à coefficients réels, le théorème de 16.2.2 montre que A est diagonalisable et on peut appliquer la proposition 1 (la matrice du produit scalaire est l'identité dans ce cas). $\square$

La diagonalisation simultanée de deux formes quadratiques apparaît donc comme une généralisation de la diagonalisation des matrices symétriques réelles. Elle justifie l'existence de directions conjuguées d'une conique dans le plan qui soient orthogonales pour le produit scalaire ordinaire. Elle justifie aussi, en mécanique, l'existence d'axes orthogonaux (appelés axes principaux) des ellipsoïdes d'inertie des solides et en statistique l'existence d'axes principaux en analyse en composantes principales.

23.7 ORTHOGONALITÉ

L'orthogonalité a déjà été définie entre vecteurs et formes linéaires (voir 10.4) et dans les espaces munis d'un produit scalaire (voir 16.3) ; nous la généralisons au cas des formes bilinéaires symétriques et des formes quadratiques.

Vecteurs orthogonaux. Soient (E, φ) un espace quadratique et q la forme quadratique associée à φ ; on notera simplement E si cela ne crée pas d'ambiguïté. On dit que deux vecteurs u et v de E sont orthogonaux (ou φ -orthogonaux ou q -orthogonaux) par rapport à φ (ou à q) si $\varphi(u, v) = 0$.

En dimension finie, en choisissant une base B de E et en notant A la matrice associée à φ dans B , U, V les matrices colonnes associées à u et v , la condition d'orthogonalité s'écrit encore ${}^tUAV = 0$ (ou ${}^tVAU = 0$).

On dit que deux parties de E sont orthogonales par rapport à φ si tout vecteur de l'une est orthogonal à tout vecteur de l'autre.

Si F est un sous-espace de E , l'orthogonal de F est l'ensemble des vecteurs de E orthogonaux à F ; c'est un sous-espace vectoriel de E qu'on note $F^\perp$.

Ces définitions copient celles données pour les produits scalaires. Mais des phénomènes nouveaux vont apparaître.

Vecteurs isotropes. Avec les mêmes notations, un vecteur u non nul de E est dit isotrope si $q(u) = 0$, autrement dit s'il est orthogonal à lui-même. Une forme sans vecteur isotrope a été appelée *définie* en 16.2 ; on dit aussi que la forme est anisotrope.

Forme non dégénérée, noyau. Une forme φ est dite non dégénérée si $E^\perp = \{0\}$; elle est dégénérée dans le cas contraire, c'est-à-dire quand il existe des vecteurs orthogonaux à tous les vecteurs de E ; ces vecteurs sont bien sûrs isotropes. Quand la forme est dégénérée, on dit parfois que l'espace est isotrope.

On appelle $E^\perp$ le noyau (ce nom sera justifié à la fin de cette section) de φ .

Exemple. Pour la forme quadratique $q : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $q((x_1, x_2)) = x_1^2 - x_2^2$, de forme polaire définie par $\varphi((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2$, le vecteur $(1, 1)$ est isotrope ; cependant, la forme φ est non dégénérée, car si $u = (x_1, x_2)$ est dans le noyau de φ , on a $\varphi(u, e_1) = \varphi(u, e_2) = 0$, ce qui impose $x_1 = x_2 = 0$.

Orthogonalité et dualité. L'étude de l'orthogonalité définie par les formes bilinéaires et quadratiques va pouvoir, grâce à la correspondance entre formes bilinéaires et applications dans le dual, utiliser les résultats d'orthogonalité pour les vecteurs et formes linéaires vus au chapitre 10.

On a vu en 23.2 qu'on pouvait associer à une forme bilinéaire symétrique φ une application linéaire $\psi : E \rightarrow E^*$, définie par $\psi(u)(v) = \varphi(u, v)$. Dire que u et v sont orthogonaux pour φ , c'est dire que $v \in \ker(\psi(u))$, ou encore, avec la définition de 10.4, que v et $\psi(u)$ sont orthogonaux.

Si B est une base de E , de base duale B^* , et si A la matrice de φ dans la base B , alors $A = M(\psi, B, B^*)$.

L'orthogonal d'un sous-espace F de E est l'ensemble des $v \in E$ tels que, pour tout $u \in F$, $\psi(u)(v) = 0$, autrement dit l'orthogonal de F au sens de l'orthogonalité défini par φ dans E est l'orthogonal de $\psi(F)$ au sens de 10.4, ce qu'on peut écrire $F^\perp = (\psi(F))^\perp$; ces deux sous-espaces sont inclus dans E , mais la notation mélange deux relations d'orthogonalités, la première étant celle définie par φ dans E , la seconde étant celle définie par la dualité entre vecteurs de E et formes linéaires de E^* ; notons que certains auteurs choisissent deux symboles différents et notent $F^\perp$ et F° les sous-espaces orthogonaux à F suivant qu'il s'agit de l'une ou l'autre des relations d'orthogonalité. L'utilisation du même signe n'est pas gênante, pourvu qu'on sache de quoi on parle, mais nous essaierons d'éviter ces ambiguïtés.

L'orthogonal $F^\perp$ de F dans E est égal à $\ker(\psi)$, ce qui justifie l'appellation de noyau pour cet espace. On en déduit la proposition suivante.

Proposition. *Une forme bilinéaire symétrique φ est non dégénérée si et seulement si l'application linéaire $\psi : E \rightarrow E^*$ est injective. En dimension finie, cette propriété équivaut à dire que ψ est bijective ou que la matrice de φ est inversible.* $\square$

23.8 ESPACES QUADRATIQUES RÉGULIERS

Espace quadratique régulier. On appelle espace quadratique régulier un espace quadratique (E, q) tel que q soit une forme quadratique non dégénérée ou, ce qui revient au même, tel que la forme bilinéaire symétrique φ associée à q soit non dégénérée.

Orthogonalité dans les espaces quadratiques réguliers. Soient (E, φ) un espace quadratique régulier de dimension finie n . On vient de voir que l'application linéaire $\psi : E \rightarrow E^*$ définie par φ est un isomorphisme. Si F est un sous-espace de E , les résultats de 10.4 donnent $\dim(F) + \dim(\psi(F)^\perp) = n$ et $(F^\perp)^\perp = F$ pour l'orthogonalité entre vecteurs et formes linéaires, ce qui donne pour l'orthogonalité définie par φ :

$$\begin{aligned}\dim(F) + \dim(F^\perp) &= n ; \\ (F^\perp)^\perp &= F.\end{aligned}$$

Une conséquence immédiate est la suivante : si F et G sont deux sous-espaces de E , $F^\perp = G^\perp$ implique, en prenant les orthogonaux, $F = G$.

Sous-espace régulier d'un espace quadratique. Un sous-espace F d'un espace quadratique (E, q) est dit régulier si $(F, q|_F)$ est un espace quadratique régulier, autrement dit si la restriction de q à F est non dégénérée.

Par exemple, une droite engendrée par un vecteur non isotrope est un sous-espace régulier.

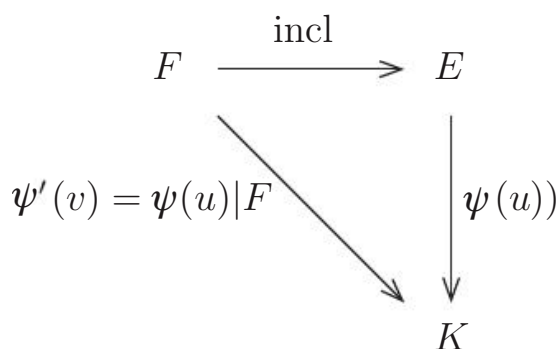
Nous allons montrer, dans cette section et la suivante, qu'un certain nombre de notions et de résultats sur les espaces euclidiens peuvent s'étendre, au moins partiellement, aux espaces quadratiques réguliers et aux sous-espaces réguliers.

Dans le cas d'un espace euclidien E , on a vu (voir 16.7) que tout sous-espace F permettait de définir une décomposition de E comme somme directe de F et de son orthogonal. Ce résultat ne peut subsister en général dans un espace quadratique puisque si F contient un vecteur isotrope, l'intersection de F avec son orthogonal n'est pas réduite à $\{0\}$. Cependant, on a le résultat suivant.

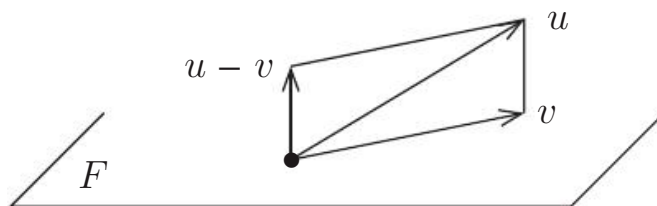
Proposition 1 : propriétés des sous-espaces réguliers. Soient (E, q) un espace quadratique régulier de dimension finie n et F un sous-espace de E .

- 1) F est régulier si et seulement si $E = F \oplus F^\perp$.
- 2) F est régulier si et seulement si $F^\perp$ est régulier.

Démonstration. 1) Si F est régulier, on a $F \cap F^\perp = \{0\}$. Soit $u \in E$; d'une part, en notant toujours $\psi : E \rightarrow E^*$ l'isomorphisme défini par $\varphi, \psi(u)$, forme linéaire sur E , définit par restriction une forme linéaire sur F ; d'autre part, ψ induit par restriction une application $\psi' : F \rightarrow F^*$ qui est un isomorphisme puisque F est régulier. Il existe donc $v \in F$ tel que $\psi'(v) = \psi(u)|_F$. La situation est la suivante.



D'où $\psi(u - v)|_F = 0$ et $u - v \in F^\perp$.



On en déduit $u = v + (u - v) \in F + F^\perp$ et $E = F \oplus F^\perp$.

Réciproquement, si $E = F \oplus F^\perp$, on a $F \cap F^\perp = \{0\}$, donc F est régulier.

2) L'égalité $F = (F^\perp)^\perp$ montre que $F \cap F^\perp = (F^\perp)^\perp \cap F^\perp$. Le 1) donne alors l'équivalence de l'énoncé. $\square$

La proposition montre que l'hyperplan orthogonal à un vecteur non isotrope est un sous-espace régulier.

Somme orthogonale. Soient (E_1, q_1) et (E_2, q_2) deux espaces quadratiques. On définit sur l'espace $E_1 \oplus E_2$ la forme quadratique q par $q(u_1, u_2) = q_1(u_1) + q_2(u_2)$. On appelle (E, q) la somme orthogonale des deux espaces et on écrit $E = E_1 \perp E_2$.

Si (E_1, q_1) et (E_2, q_2) sont des espaces quadratiques réguliers, on vérifie que (E, q) est un espace quadratique régulier et que $E_1^\perp = E_2, E_2^\perp = E_1$.

Cette définition et cette propriété se généralisent à une famille finie d'espaces quadratiques.

Si (E, q) est un espace quadratique et si F est un supplémentaire de $E^\perp$, alors $E = E^\perp \perp F$, car l'orthogonalité des deux sous-espaces est claire. De plus, le sous-espace F est régulier, car il ne peut contenir de vecteur non nul orthogonal à F puisque ce vecteur serait élément du noyau $E^\perp$ de q .

Proposition 2 : existence de bases orthogonales. Soit (E, q) un espace quadratique de dimension finie n . Il existe une base orthogonale pour q .

Démonstration. 1) La proposition est vraie si $q = 0$ puisqu'alors une base quelconque de E est orthogonale.

Supposons q non nulle.

2) Raisonnons par récurrence sur n . La proposition est vraie pour $n = 0$ ou $n = 1$. Soit $n > 1$ et supposons que le résultat soit vrai pour tout espace de dimension strictement inférieure à n .

3) Envisageons d'abord le cas où le noyau $E^\perp$ de q n'est pas réduit à $\{0\}$. Notons F un supplémentaire de $E^\perp$ dans E ; on a $\dim(F) < n$. L'hypothèse de récurrence assure l'existence d'une base orthogonale B de F ; sa réunion avec une base quelconque de $E^\perp$ donne une base orthogonale de E .

4) Étudions maintenant le cas où $E^\perp = \{0\}$; la forme q est non dégénérée et il existe un vecteur non isotrope u . L'espace Ku est régulier et on a $E = Ku \perp (Ku)^\perp$ d'après la proposition 1. L'hypothèse de récurrence assure l'existence d'une base orthogonale B de $(Ku)^\perp$; cette base formée avec u est une base orthogonale de E , ce qui achève le raisonnement par récurrence. $\square$

Plan hyperbolique. Un espace quadratique (E, q) de dimension 2 est un plan hyperbolique s'il existe une base $B = (e_1, e_2)$ de E tel que la matrice de q soit

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

On a alors $q(x_1, x_2) = 2x_1 x_2$. On sait (voir exemple de 23.2) qu'il existe une base $B'' = (e_1'', e_2'')$ de E dans laquelle la forme quadratique q est donnée par

$$q(x_1'', x_2'') = x_1''^2 - x_2''^2 \text{ et a pour matrice associée } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Un plan hyperbolique est un espace quadratique régulier.

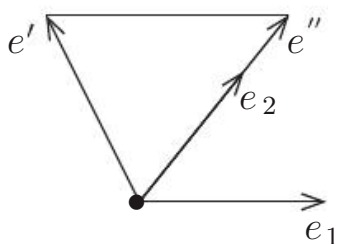
Le nom de plan hyperbolique est le plus fréquemment donné. Marcel Berger utilise dans sa *Géométrie* le nom de plan artinien (en hommage à Emil Artin) qui présente l'avantage de ne pas évoquer la géométrie hyperbolique qui est un tout autre sujet.

Proposition 3 : caractérisation des plans hyperboliques. *Un espace quadratique régulier (E, q) de dimension 2 est un plan hyperbolique si et seulement si E contient un vecteur isotrope.*

Démonstration. Si (E, q) est un plan hyperbolique et si $B = (e_1, e_2)$ est une base de E telle que la matrice de q soit $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, les vecteurs e_1 et e_2 sont isotropes.

Réciproquement, soit (E, q) un espace quadratique régulier de dimension 2 contenant un vecteur isotrope e_1 . Comme q n'est pas dégénérée, il existe un vecteur $e' \in E$ tel que $\varphi(e_1, e') \neq 0$ où φ est la forme polaire associée à q . On peut choisir $\lambda \in K$ tel que $e'' = e' + \lambda e_1$ soit isotrope : on a $q(e' + \lambda e_1) = q(e') + 2\lambda\varphi(e_1, e')$ et il suffit de prendre $\lambda = -\frac{q(e')}{2\varphi(e_1, e')}$. On pose alors $e_2 = \alpha e''$ avec α tel que $\varphi(e_1, e_2) = 1$.

On peut illustrer par le dessin suivant.



La matrice de q par rapport à la base $B = (e_1, e_2)$ est de la forme désirée. $\square$

Proposition 4. *Soit (E, q) un espace quadratique régulier de dimension $n \geq 3$ contenant un vecteur isotrope e . Alors il existe un plan vectoriel H de E tel que la restriction de q à H donne un plan hyperbolique.*

Démonstration. Comme q est non dégénérée, il existe un vecteur e' tel que $\varphi(e, e') \neq 0$. On pose $H = \text{Vect}(e, e')$ et on applique la proposition 3. $\square$

Chapitre 23 • Formes bilinéaires symétriques et quadratiques

Les plans hyperboliques sont des briques de construction d'espaces quadratiques ; l'expliquer nous mènerait aux théorèmes de Ernst Witt (1911-1991), nous ne pouvons aller aussi loin.

Élie et Henri Cartan



Élie Cartan (1869-1951)
© Académie des Sciences.



Henri Cartan (1904-2008)
© Académie des Sciences.

Élie Cartan (1869-1951) était d'une famille très pauvre de forgerons né à Dolomieu, entre Lyon et Chambéry. Il eut la chance d'être remarqué par un jeune inspecteur remarquable (Antonin Dubost, futur président du Sénat) qui lui assura le financement de ses études jusqu'à l'École normale supérieure. Ses travaux sont parmi les plus importants de la première moitié du vingtième siècle. Ils portent sur les équations différentielles, la géométrie des spineurs (fondamentale pour le développement de la mécanique quantique), les espaces quadratiques, etc.

C'est ici l'occasion de dire deux mots de son fils Henri Cartan, né en 1904 et qui est mort centenaire en 2008. Quand Henri Cartan a commencé sa carrière, la guerre venait de tuer bon nombre des jeunes scientifiques français ; passionné de musique, extrêmement chaleureux et enthousiaste, Henri Cartan a animé la vie mathématique pendant de longues années par ses nombreuses recherches, sa participation à la fondation et aux travaux du groupe Bourbaki, par des livres dont l'influence a été profonde, par son enseignement à des générations de normaliens, par la direction pendant 15 ans du fameux séminaire Cartan à Paris, etc.

23.9 QUATERNIONS

La recherche d'Hamilton. Une fois les nombres réels et les nombres complexes découverts, que faire ? Depuis longtemps, Hamilton se posait la question. Il avait

remarqué que les nombres complexes permettaient de multiplier des couples de nombres réels en obtenant une loi de corps, remarque banale aujourd'hui. Il entreprit de trouver une loi de corps sur l'ensemble des triplets de nombres réels, manipulant ce qu'il appella plus tard des *vecteurs*. Mais la tâche se révéla difficile ; les années passaient et Hamilton ne savait toujours pas comment multiplier ses triplets ; il eût une idée, puis une autre, puis encore une autre, aucune ne marchait. Le matin, au petit déjeuner, ses enfants l'interrogeaient : *Alors, papa, est-ce que tu sais aujourd'hui multiplier les triplets ?* Le lundi 16 octobre 1843, au cours d'une promenade avec sa femme, *ce fut comme un éblouissement* : c'est en multipliant des quadruplets qu'il pouvait obtenir une loi de corps, enfin presque.

Plus tard, en 1877, Frobenius montrera qu'on ne pouvait définir de multiplication sur les triplets donnant une structure de corps à $\mathbb{R}^3$.



Sir William Rowan Hamilton (1805-1865)
d'après le portrait de la Royal Irish Academy.

Définition des quaternions. Plusieurs définitions sont possibles. En voici une : un quaternion est une matrice de $M_2(\mathbb{C})$ de la forme

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

où $\alpha, \beta \in \mathbb{C}$; un quaternion est donc défini par deux nombres complexes (ou quatre nombre réels).

Les quaternions forment un $\mathbb{R}$ -espace vectoriel noté $\mathbb{H}$. En suivant les notations traditionnelles d'Hamilton, on définit les quaternions :

$$1 = I \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Les injections de $\mathbb{R}$ et $\mathbb{C}$ dans $\mathbb{H}$ définies par $a \mapsto a.1$, $a + ib \mapsto a.1 + bi$ permettent d'identifier les $\mathbb{R}$ -espaces vectoriels $\mathbb{R}$ et $\mathbb{C}$ à des sous-espaces vectoriels de $\mathbb{H}$.

Chapitre 23 • Formes bilinéaires symétriques et quadratiques

On notera la différence des caractères utilisés : $i \in \mathbb{C}$ et $i \in \mathbb{H}$, $j \in \mathbb{C}$ et $j \in \mathbb{H}$. En posant $\alpha = a + ib$, $\beta = c + id$, on voit que tout quaternion peut s'écrire d'une façon unique sous la forme

$$x = a + bi + cj + dk$$

avec a, b, c, d réels et que $(1, i, j, k)$ forme une base de $\mathbb{H}$ en tant que $\mathbb{R}$ -espace vectoriel ; $\mathbb{H}$ est donc de dimension 4 sur $\mathbb{R}$.

Le corps non commutatif des quaternions. Il est facile de vérifier que $\mathbb{H}$ est une $\mathbb{R}$ -algèbre, sous-algèbre de la $\mathbb{R}$ -algèbre $M_2(\mathbb{C})$ (en particulier, la multiplication est associative dans $\mathbb{H}$) ; en effet, la multiplication de deux quaternions s'écrit :

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ -\bar{\beta}' & \bar{\alpha}' \end{pmatrix} = \begin{pmatrix} \alpha\alpha' - \beta\bar{\beta}' & \alpha\beta' + \beta\bar{\alpha}' \\ -\bar{\alpha}\bar{\beta}' - \bar{\beta}\alpha' & \bar{\alpha}\bar{\alpha}' - \bar{\beta}\beta' \end{pmatrix}$$

où le résultat est le quaternion défini par les deux complexes $\alpha\alpha' - \beta\bar{\beta}'$ et $\alpha\beta' + \beta\bar{\alpha}'$. Cette relation donne la relation de multiplication des quadruplets :

$$\begin{aligned} (a + bi + cj + dk)(a' + b'i + c'j + d'k) = & aa' - bb' - cc' - dd' \\ & + (ab' + a'b + cd' - c'd)i \\ & + (ac' - bd' + a'c + b'd)j \\ & + (ad' + bc' - b'c + a'd)k. \end{aligned}$$

On vérifie facilement que :

$$\begin{aligned} i^2 = j^2 = k^2 = -1, \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{aligned}$$

On a donc perdu la commutativité.

En revanche, tout quaternion non nul est inversible : son inverse s'obtient en calculant l'inverse d'une matrice :

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}^{-1} = \frac{1}{|\alpha|^2 + |\beta|^2} \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix}$$

et c'est un quaternion, défini par $\frac{\bar{\alpha}}{|\alpha|^2 + |\beta|^2}$ et $\frac{-\beta}{|\alpha|^2 + |\beta|^2}$. Le résultat s'écrit encore :

$$(a + bi + cj + dk)^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2} (a - bi - cj - dk).$$

Ces propriétés font de $\mathbb{H}$ un corps non commutatif. Le centre de ce corps, c'est-à-dire l'ensemble des éléments commutant avec tous les éléments du corps, est $\mathbb{R}$.

Conjugaison et norme. On appelle quaternion pur un quaternion dont la partie réelle est nulle ; c'est un quaternion de la forme $bi + cj + dk$. On note $\mathbb{H}_p$ l'ensemble des quaternions purs.

On définit une conjugaison dans $\mathbb{H}$, notée de la même façon que la conjugaison dans $\mathbb{C}$, mais qui doit en être distinguée : si $x = a + bi + cj + dk$, alors $\bar{x} = a - bi - cj - dk$, autrement dit, si $x = a + y$ avec y pur, alors $\bar{x} = a - y$. Il est clair que $\overline{\bar{x}} = x$, que $x = \bar{x}$ si et seulement si x est réel et que $x = -\bar{x}$ si et seulement si x est un quaternion pur.

Revenons à la notation matricielle ; si $x = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$, la matrice $\bar{x}$ est la conjuguée de la transposée de la matrice de x . Comme ${}^t(AB) = {}^tB {}^tA$, on en déduit $\overline{xx'} = \bar{x'}\bar{x}$.

On définit une application $N : \mathbb{H} \rightarrow \mathbb{R}^+$, appelée norme, en posant $N(x) = x\bar{x}$.

La matrice de $N(x)$ est $\begin{pmatrix} \alpha\bar{\alpha} + \beta\bar{\beta} & 0 \\ 0 & \alpha\bar{\alpha} + \beta\bar{\beta} \end{pmatrix}$, ce qui montre que $N(x) = \alpha\bar{\alpha} + \beta\bar{\beta} = a^2 + b^2 + c^2 + d^2$ est un réel positif ou nul et que $N(x) = \det(x)$. On en déduit que $N(x) = N(\bar{x})$ et que $N(xx') = N(x)N(x')$ d'après la formule du déterminant d'un produit (voir 14.7).

On peut encore remarquer que $x^{-1} = \frac{\bar{x}}{N(x)}$, ou encore, si $x = a + y$ avec y quaternion pur, $x^{-1} = (a + y)^{-1} = \frac{a - y}{N(x)}$.

Sommes de quatre carrés. Sous la formule $N(x)N(x') = N(xx')$ se cache une superbe formule découverte par Euler en 1748, bien avant Hamilton, montrant que les sommes de quatre carrés sont stables par produit :

$$\begin{aligned} (a^2 + b^2 + c^2 + d^2)(a'^2 + b'^2 + c'^2 + d'^2) &= (aa' - bb' - cc' - dd')^2 \\ &\quad + (ab' + a'b + cd' - c'd)^2 \\ &\quad + (ac' - bd' + a'd + b'c)^2 \\ &\quad + (ad' + bc' - b'd + a'c)^2 \end{aligned}$$

La formule permet, pour démontrer le théorème :

Tout entier n est somme de quatre carrés d'entiers,

de se ramener au cas où n est premier. Ce dernier résultat, peut-être connu de Diophante, conjecturé par Bachet de Méziriac (1621), dont Fermat affirmait en 1636 avoir une démonstration, qu'Euler ne put démontrer complètement (années 1740-1750), fut obtenu par Lagrange en 1770.

Orthogonalité. On a vu que $\mathbb{H}$ a une structure d'espace vectoriel de dimension 4 sur $\mathbb{R}$. L'application $N : \mathbb{H} \rightarrow \mathbb{R}$ est une forme quadratique dont la forme polaire φ est un produit scalaire défini par $\varphi(x, x') = aa' + bb' + cc' + dd'$ et on a $\varphi(x, x') = \frac{1}{2}(x\bar{x}' + x'\bar{x})$ (pour vérifier cette dernière expression, il suffit de voir qu'elle est réelle car égale à sa conjuguée et de rechercher, dans l'expression des deux produits, les termes donnant des réels, sans calculer les autres).

On vérifie que la base $(1, i, j, k)$ est orthonormée pour la forme quadratique N et que $x \mapsto \bar{x}$ est une symétrie orthogonale par rapport à la droite $\text{Vect}(1)$.

La sphère S_3 de $\mathbb{R}^4$ comme groupe topologique. Les quaternions de norme 1 forment un sous-ensemble U de $\mathbb{H}^*$ en bijection avec la sphère S_3 de $\mathbb{R}^4$: au quaternion $a + bi + cj + dk$ de U on associe le point (a, b, c, d) de S_3 . Si on munit U de la topologie induite par la topologie de $\mathbb{R}^4$, les formules donnant le produit de deux quaternions et l'inverse d'un quaternion non nul montrent que la multiplication et l'inversion sont continues ; U est donc un groupe topologique connexe. Par transport de structure, on peut donc munir la sphère S_3 d'une structure de groupe topologique connexe. On peut montrer que c'est la seule sphère S_n avec S_1 (dont la structure de groupe topologique est donnée par les complexes de module 1), qui possède cette propriété remarquable.

La sphère S_3 a été l'objet d'une conjecture célèbre formulée en 1904 par Henri Poincaré (1854-1912) : *Toute variété compacte simplement connexe sans bord de dimension 3 est homéomorphe à la sphère S_3 .* Cette conjecture s'est révélée extrêmement profonde (*Cette question nous entraînerait trop loin*, écrivait Poincaré) ; elle a stimulé le travail des mathématiciens pendant un siècle. Elle a été résolue en 2002-2003 par Grigori Perelman (né en 1966). Pour des raisons qui lui appartiennent, Perelman a refusé les récompenses : il a refusé la médaille Fields en 2006, il a refusé le prix (1 million de dollars) de la fondation Clay en 2010 pour la résolution de l'un des problèmes dit du millénaire. On ne sait pas non plus sur quoi travaille Perelman aujourd'hui, peut-être sur un autre problème du millénaire, la résolution des équations de Navier-Stokes, qui sait ?

Quaternions purs. Soit $x \in \mathbb{H}$. Montrons deux équivalences :

- x est réel si et seulement si x^2 est réel positif ou nul ;
- x est un quaternion pur si et seulement si x^2 est réel négatif ou nul.

D'abord, si x est réel, x^2 est bien un réel positif ou nul ; si $x = bi + cj + dk$ est un quaternion pur, $x^2 = -x(-x) = -x\bar{x} = -N(x) = -b^2 - c^2 - d^2$ est bien un réel négatif ou nul.

Réciproquement, si $x = a + y$ avec a réel et y quaternion pur, on a $x^2 = a^2 + 2ay + y^2$ où $a^2 + y^2$ est réel ; si x^2 est réel, on a $ay = 0$; si $a = 0$, on a $x^2 \leq 0$ et $x = y$ est un quaternion pur ; si $y = 0$, on a $x^2 \geq 0$ et x est réel.

Le seul élément pur et réel de $\mathbb{H}$ est 0 (il faut être nul pour être pur et réel !).

Action de $\mathbb{H}^*$ sur $\mathbb{R}^3$. On suppose $\mathbb{R}^3$ muni du produit scalaire défini par la base canonique. L'application $\tau : \mathbb{H}_p \rightarrow \mathbb{R}^3$ définie par $\tau(bi + cj + dk) = (b, c, d)$ est un isomorphisme d'espaces vectoriels qui conserve la norme :

$$N(bi + cj + dk) = b^2 + c^2 + d^2 = \|(b, c, d)\|^2.$$

On peut identifier par τ les quaternions purs aux vecteurs de $\mathbb{R}^3$ et on notera y à la fois le quaternion $bi + cj + dk$ et le triplet $\tau(y) = (b, c, d)$ correspondant.

Considérons l'action de $\mathbb{H}^*$ sur $\mathbb{H}$ par conjugaison : pour $x \in \mathbb{H}^*$ et $x' \in \mathbb{H}$, on pose $x.x' = xx'x^{-1}$. Cette action se restreint en une action de $\mathbb{H}^*$ sur $\mathbb{H}_p$, car si $x' = y \in \mathbb{H}_p$, on a $(xyx^{-1})^2 = xy^2x^{-1} = y^2$ puisque y^2 est un réel et que, comme y^2 est négatif, on a $xyx^{-1} \in H_p$.

On définit un homomorphisme de groupes $\rho : \mathbb{H}^* \rightarrow \text{GL}(\mathbb{R}^3)$ en posant, pour tout $x \in \mathbb{H}^*$ et tout $y \in \mathbb{H}_p$, $\rho(x)(y) = xyx^{-1}$; l'application $\rho(x) : \mathbb{H}_p \rightarrow \mathbb{H}_p$ est $\mathbb{R}$ -linéaire et bijective pour tout $x \in \mathbb{H}^*$.

Nous allons montrer que ρ est à valeurs dans $\text{SO}(\mathbb{R}^3)$ (voir 16.11 pour la description des éléments de ce groupe). On va vérifier que $\rho(x)$ est orthogonale et que son déterminant est 1.

La première condition résulte (d'après la proposition 1 de 16.9) de

$$\|\rho(x)(y)\|^2 = N(xyx^{-1}) = N(x)N(y)N(x^{-1}) = N(y) = \|y\|^2.$$

La seconde condition est plus délicate à montrer. On calcule d'abord la matrice $M(x)$ de $\rho(x)$ pour $x = a + bi + cj + dk$; ses colonnes sont données par $\rho(x)i = xix^{-1}$, $\rho(x)j = xjx^{-1}$, $\rho(x)k = xkx^{-1}$. Après calculs, on obtient pour $M(x)$:

$$\frac{1}{a^2 + b^2 + c^2 + d^2} \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(ac + bd) \\ 2(bc + ad) & a^2 - b^2 + c^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(ab + cd) & a^2 - b^2 - c^2 + d^2 \end{pmatrix}$$

On vient de montrer que cette matrice est orthogonale (on aurait pu faire une vérification directe, un peu fastidieuse, mais sans difficulté), donc son déterminant vaut ± 1 . La formule précédente montre que ρ est une application continue (en considérant $\text{GL}(\mathbb{R}^3)$ muni de la topologie induite par celle de $\mathbb{R}^9$), donc l'application $\det \circ \rho : \mathbb{H}^* \rightarrow \{-1, 1\}$ est continue. Comme sa source est connexe, son image est connexe ; comme $\det(\rho(1)) = 1$, cette image est $\{1\}$, ce qui prouve que $\rho(x) \in \text{SO}(\mathbb{R}^3)$.

Le noyau de ρ est l'ensemble des $x \in \mathbb{H}^*$ tels que $\rho(x)$ laisse i, j, k fixes, ce qui donne $x \in \mathbb{R}^*$. Pour déterminer l'image de ρ , on remarque que, si $y = bi + cj + dk \in \mathbb{H}_p$, alors $\rho(y)(y) = yyy^{-1} = y$; $\rho(y)$ est donc une rotation d'axe

$\text{Vect}(b, c, d)$. Comme y^2 est réel, $\rho(y)^2 = \text{id}$, donc $\rho(y)$, qui n'est pas l'identité, est un demi-tour d'axe $\text{Vect}(b, c, d)$. Ainsi, tout demi-tour est dans l'image de ρ . Comme les demi-tours engendrent le groupe $\text{SO}(\mathbb{R}^3)$ des rotations de $\mathbb{R}^3$ (voir 23.1), on voit que ρ est surjective. On obtient donc un isomorphisme :

$$\frac{\mathbb{H}^*}{\mathbb{R}^*} \simeq \text{SO}(\mathbb{R}^3).$$

On prolongera ces résultats dans l'exercice 23.10.

Les rotations de $\mathbb{R}^3$ ne sont pas des objets faciles à composer ; la paramétrisation par les angles d'Euler ne le permet pas en général. Un grand intérêt de cette section est de montrer comment faire, en paramétrant le groupe des rotations de $\mathbb{R}^3$ par les quaternions non nuls, à un facteur près. Pour déterminer la rotation composée $r \circ r'$, on peut donc chercher des quaternions x et x' tel que $\rho(x) = r$, $\rho(x') = r'$, puis déterminer la rotation $\rho(xx')$.

Cela est très utile en physique, en mécanique, en robotique, en cinéma d'animation...

23.10 RECHERCHES ARITHMÉTIQUES DE LAGRANGE

Pour terminer ce chapitre, nous présentons des résultats que Lagrange publia en 1773 dans les *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin* sous le simple titre : *Recherches arithmétiques* (pages 695-795 du volume III des Œuvres complètes). Voici comment Lagrange présente ses travaux.

Ces recherches ont pour objet les nombres qui peuvent être représentés par la formule

$$Bt^2 + Ctu + Du^2$$

où B, C, D sont supposés des nombres entiers donnés, et t, u des nombres aussi entiers, mais indéterminés. Je donnerai d'abord la manière de trouver toutes les différentes formes dont les diviseurs de ces sortes de nombres sont susceptibles ; je donnerai ensuite une méthode pour réduire ces formes au plus petit nombre possible ; je montrerai comment on en peut dresser des Tables pour la pratique, et je ferai voir l'usage de ces Tables dans la recherche des diviseurs des nombres. Je donnerai enfin la démonstration de plusieurs Théorèmes sur les nombres premiers de la même forme $Bt^2 + Ctu + Du^2$, dont quelques uns sont déjà connus, mais n'ont pas encore été démontrés, et dont les autres sont entièrement nouveaux.

Lagrange fait ici allusion à des affirmations de Fermat, à des travaux d'Euler ; il ajoute : *mais personne, que je sache, n'a encore traité cette matière d'une manière directe et générale ...*

À la différence des sections précédentes, les valeurs des variables sont ici des entiers et nous redonnerons les définitions qui nous seront utiles dans ce cadre. Nous avons suivi le texte de Lagrange et la présentation qui en est donnée dans le livre de Winfried Scharlau et Hans Opolka : *From Fermat to Minkowski*, édité en 1984 par Springer-Verlag.

Toutes les notations désigneront des entiers.

Formes quadratiques sur $\mathbb{Z}^2$. Une forme quadratique sur $\mathbb{Z}^2$ est une application $q : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ définie par une expression $q(x, y) = ax^2 + 2bxy + cy^2$ où $a, b, c \in \mathbb{Z}$. Le nombre $ac - b^2$ est appelé discriminant de la forme ; nous le supposons non nul.

Écriture matricielle. Si $q(x, y) = ax^2 + 2bxy + cy^2$, on peut écrire $q(x, y)$ sous la forme d'un produit de matrices ${}^tUMU$ avec $U = \begin{pmatrix} x \\ y \end{pmatrix}$ et $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. La matrice M est dite associée à q .

Ses coefficients sont des entiers. La forme est dite à matrice entière. La situation est plus simple que celle des formes quadratiques $ax^2 + bxy + cy^2$ à coefficients entiers dans laquelle la matrice associée peut comporter des termes $b/2$ non entiers.

Groupes $GL(2, \mathbb{Z})$ et $SL(2, \mathbb{Z})$. Dans l'anneau $M_2(\mathbb{Z})$ des matrices carrées d'ordre 2 à coefficients dans $\mathbb{Z}$, on considère le sous-groupe multiplicatif $GL(2, \mathbb{Z})$ des matrices inversibles. Le déterminant d'une telle matrice est égal à ± 1 , car si $M \in GL(2, \mathbb{Z})$, $\det(M)$ et $\det(M^{-1})$ sont des entiers et $\det(M)\det(M^{-1}) = 1$. Le sous-groupe des matrices de déterminant 1 de $GL(2, \mathbb{Z})$ est noté $SL(2, \mathbb{Z})$.

$$\text{Posons } T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ et } T' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} ; \text{ on a } TT'^{-1}T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$(TT'^{-1}T)^2 = -I \text{ et, pour tout } k \in \mathbb{Z}, T^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \text{ et } T'^k = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}.$$

$$\text{Soit } P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ une matrice de } SL(2, \mathbb{Z}).$$

$$\text{On a : } T^k P = \begin{pmatrix} a + kc & b + kd \\ c & d \end{pmatrix}$$

$$T'^k P = \begin{pmatrix} a & b \\ ka + c & kb + d \end{pmatrix}$$

$$TT'^{-1}TP = \begin{pmatrix} c & d \\ -a & -b \end{pmatrix}$$

Comme $ad - bc = 1$, a et c sont premiers entre eux, et des multiplications à gauche par des puissances de T ou de T' correspondants aux différentes étapes de l'algorithme de recherche de $\text{pgcd}(a, c)$ permettent d'obtenir (il faut le vérifier : si la division euclidienne de a par c s'écrit $a = qc + r$, on multiplie P à gauche par T^{-q} , etc.) une matrice de la forme $\begin{pmatrix} 1 & b' \\ 0 & 1 \end{pmatrix}$; comme cette dernière matrice est égale à $T^{b'}$, on voit que le groupe $SL(2, \mathbb{Z})$ est engendré par T et T' .

On en déduit que le groupe $GL(2, \mathbb{Z})$ est engendré par T , T' et $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Formes équivalentes. Soient q et q' deux formes quadratiques sur $\mathbb{Z}^2$ de matrices associées M et M' . Ces deux formes seront dites équivalentes s'il existe une matrice P de $GL(2, \mathbb{Z})$ telle que $M' = {}^tPMP$, proprement équivalentes s'il existe une matrice P de $SL(2, \mathbb{Z})$ telle que $M' = {}^tPMP$. Ces deux relations sont des relations d'équivalence.

Deux formes équivalentes ont même ensemble de valeurs. De plus, comme $\det(M') = \det({}^tPMP) = \det(M)\det(P)^2 = \det(M)$, elles ont des discriminants égaux.

Il existe donc une infinité de formes équivalentes de même discriminant.

Formes définies positives. La forme $q(x, y) = ax^2 + 2bxy + cy^2$ est dite définie positive si elle ne prend que des valeurs strictement positives pour x, y non nuls. On a donc $a, c > 0$ (prendre y ou x nul) et on voit que $ac - b^2 > 0$ en écrivant

$$ax^2 + 2bxy + cy^2 = a \left(x + \frac{b}{a}y \right)^2 + \frac{ac - b^2}{a}y^2.$$

Si $ac - b^2 < 0$, la forme est dite indéfinie ; elle peut prendre des valeurs positives, négatives ou nulles (prendre $x = kb, y = -ka$, etc.).

Représentation propre, diviseur d'une forme quadratique. Soit q une forme quadratique sur $\mathbb{Z}^2$ définie par $q(x, y) = ax^2 + 2bxy + cy^2$.

On dit qu'un entier $n \in \mathbb{Z}$ est représenté proprement par q s'il existe des entiers x_0, y_0 premiers entre eux tels que $q(x_0, y_0) = n$.

On dit qu'un entier $n \in \mathbb{Z}$ est un diviseur propre de q s'il existe un multiple de n représenté proprement par q .

Représentations propres, diviseurs propres et équivalences. Soient q et q' deux formes quadratiques équivalentes sur $\mathbb{Z}^2$ de matrices associées M et M' et $P \in GL(2, \mathbb{Z})$ telle que $M' = {}^tPMP$.

Si $n \in \mathbb{Z}$ est représenté proprement par q' , il existe x'_0, y'_0 premiers entre eux tels que $(x'_0 \ y'_0)M' \begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix} = n$. Par conséquent, n est représenté par q :

$$n = (x'_0 \ y'_0)^t PMP \begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix}.$$

Posons $P = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. On vérifie facilement qu'un nombre divisant

$x_0 = \alpha x'_0 + \beta y'_0$ et $y_0 = \gamma x'_0 + \delta y'_0$ divise x'_0 et y'_0 , donc est égal à ± 1 ; par conséquent, n est représenté proprement par q .

De même, un diviseur propre de q' est un diviseur propre de q .

Proposition 1. Soit $q(x, y) = ax^2 + 2bxy + cy^2$ une forme quadratique. Si n est un diviseur propre de q , il existe une forme quadratique $Q(x, y) = Ax^2 + 2Bxy + Cy^2$ de même discriminant que q et représentant proprement n .

Démonstration. Soient x_0, y_0 premiers entre eux et k tels que $ax_0^2 + 2bx_0y_0 + cy_0^2 = kn$. Soient $d = \text{pgcd}(y_0, k)$, X_0 et m tels que $y_0 = X_0d$, $k = md$. On a :

$$ax_0^2 + 2bx_0X_0d + cX_0^2d^2 = mdn$$

ce qui montre que d divise a , puisque d divisant y_0 est premier avec x_0 ; posons $a = a_1d$; on obtient :

$$a_1x_0^2 + 2bx_0X_0 + cX_0^2d = mn.$$

Comme $\text{pgcd}(X_0, m) = 1$, l'identité de Bézout donne, en multipliant par x_0 , des entiers r et Y_0 tels que $rX_0 + Y_0m = x_0$. Par conséquent :

$$a_1(rX_0 + Y_0m)^2 + 2b(rX_0 + Y_0m)X_0 + cX_0^2d = mn.$$

On en déduit :

$$(a_1r^2 + 2br + cd)X_0^2 + (2a_1rm + 2bm)X_0Y_0 + a_1m^2Y_0^2 = mn$$

Comme $\text{pgcd}(X_0, m) = 1$, le coefficient de X_0^2 est divisible par m . Posons $A = (a_1r^2 + 2br + cd)/m$, $B = a_1r + b$, $C = a_1m$. Un calcul facile montre que $AC - B^2 = ac - b^2$. D'autre part, un diviseur commun à X_0 et Y_0 doit diviser x_0 et y_0 ; ce ne peut être que 1. La forme $Q(x, y) = Ax^2 + 2Bxy + Cy^2$ répond aux conditions demandées. $\square$

Proposition 2. Soit $Q(x, y) = Ax^2 + 2Bxy + Cy^2$ une forme quadratique définie positive de discriminant $\Delta = AC - B^2$. Alors Q est proprement équivalente à une forme q , $q(x, y) = ax^2 + 2bxy + cy^2$, dite réduite, ce qui signifie que a, b, c vérifient la condition suivante :

$$\begin{cases} 0 < a < c \text{ et } -a < 2b \leq a \\ \text{ou} \\ 0 < a = c \text{ et } 0 \leq 2b \leq a \end{cases}$$

L'entier a est la plus petite valeur strictement positive de q .

La forme q vérifiant ces conditions est unique et on a :

$$a \leq 2\sqrt{\frac{\Delta}{3}}, \quad |b| \leq \sqrt{\frac{\Delta}{3}}.$$

Démonstration. Notons a la plus petite valeur strictement positive de q et soient X_0, Y_0 des entiers tels que $a = AX_0^2 + 2BX_0Y_0 + CY_0^2$.

1) Montrons qu'on peut trouver une forme Q' proprement équivalente à Q telle que $Q'(x, y) = ax^2 + \dots$

Si Y_0 est nul, la plus petite valeur strictement positive de Q est obtenue pour $X_0 = 1$; on a $a = A$ et il n'y a rien à faire. Si $X_0 = 0$, de manière analogue, on a

$a = C$; en effectuant le changement de base défini par $P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, on obtient

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} C & -B \\ -B & A \end{pmatrix}$$

qui définit une forme quadratique vérifiant la condition demandée.

Si X_0 et Y_0 sont tous deux non nuls, ils sont nécessairement premiers entre eux (sinon, on obtient une valeur de Q plus petite en divisant X_0 et Y_0 par leur pgcd) et il existe des entiers u, v tels que $uX_0 + vY_0 = 1$.

La matrice $P = \begin{pmatrix} X_0 & -v \\ Y_0 & u \end{pmatrix}$ appartient à $SL(2, \mathbb{Z})$; le changement de base

qu'elle définit transforme Q en $ax^2 + 2b_1xy + c_1y^2$ où b_1 et c_1 sont donnés par :

$${}^tPMP = \begin{pmatrix} X_0 & Y_0 \\ -v & u \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} X_0 & -v \\ Y_0 & u \end{pmatrix} = \begin{pmatrix} a & b_1 \\ b_1 & c_1 \end{pmatrix}$$

2) On utilise alors la matrice T pour faire un nouveau changement de base :

$${}^tT^k \begin{pmatrix} a & b_1 \\ b_1 & c_1 \end{pmatrix} T^k = \begin{pmatrix} a & b_1 + ka \\ b_1 + ka & c_2 \end{pmatrix}.$$

où c_2 dépend de k . On peut choisir k tel que $-a < 2b_1 + 2ka \leq a$. Cette valeur de k donne un entier $b = b_1 + ka$ et une valeur c de c_2 . On a donc $-a < 2b \leq a$ et $0 < a \leq c$ (car $c > 0$ puisque c'est une valeur de la forme quadratique en un vecteur non nul).

3) Si $c = a$ et si $b < 0$, on obtient les inégalités recherchées en effectuant le changement de base défini par $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} a & -b \\ -b & a \end{pmatrix}$$

on change ainsi b en $-b$ et on a $0 \leq 2b \leq a$.

La forme $q(x, y) = ax^2 + 2bxy + cy^2$ est une forme réduite proprement équivalente à Q satisfaisant les conditions de l'énoncé.

4) Comme $4b^2 \leq a^2$ et $c \geq a$, on a $4\Delta = 4(ac - b^2) \geq 3a^2$, donc : $2\sqrt{\frac{\Delta}{3}} \geq a$ et l'inégalité pour b s'en déduit.

5) Montrons maintenant l'unicité de la forme réduite obtenue.

Soit $q_1(x, y) = rx^2 + 2sxy + ty^2$ une forme réduite positive proprement équivalente à q ; on a donc :

$$0 < r < t \text{ et } -r < 2s \leq r,$$

ou :

$$0 < r = t \text{ et } 0 \leq 2s \leq r.$$

Commençons par une étude préliminaire. Déterminons pour quelles valeurs de (x, y) la forme q_1 atteint son minimum strictement positif (qui est a). On peut se limiter à une étude pour $x \geq 0$, puisque $q_1(-x, -y) = q_1(x, y)$. Si $0 < y \leq x$, $rx^2 + 2sxy = x(rx + 2sy) \geq (r + 2s)y^2 > 0$, donc $q_1(x, y) > ty^2 \geq t$. Si $0 < x \leq y$, $2sxy + ty^2 = y(2sx + ty) \geq (t + 2s)x^2 > 0$, donc $q_1(x, y) > rx^2 \geq r$. On raisonne de façon analogue si $y < 0$.

On en conclut que la valeur minimum de q_1 s'obtient en les deux points $(\pm 1, 0)$ si $t > r$ et en les quatre points $(\pm 1, 0)$ et $(0, \pm 1)$ si $r = t$. On a donc $r = a$.

Soit $P = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ un changement de variable de $SL(2, \mathbb{Z})$ transformant q en q_1 .

On a :

$$\begin{pmatrix} a & s \\ s & t \end{pmatrix} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

avec $a = a\alpha^2 + 2b\alpha\gamma + c\gamma^2 = q(\alpha, \gamma)$, $s = a\alpha\beta + b(\beta\gamma + \alpha\delta) + c\gamma\delta$.

Supposons d'abord $c > a$. La première égalité impose $\alpha = \pm 1$ et $\gamma = 0$ d'après l'étude préliminaire appliquée à q ; d'où $\delta = \alpha = \pm 1$. Alors $s = b \pm \beta a$; comme on a $-a < 2s \leq a$, on a $\beta = 0$, d'où $s = b$ puis $t = c$.

Supposons enfin $a = c$ et $0 \leq 2b \leq a$. La première égalité impose $\alpha = \pm 1$ et $\gamma = 0$ (et alors $\delta = \alpha = \pm 1$) ou $\alpha = 0$ et $\gamma = \pm 1$ (et alors $\beta = -\gamma = \mp 1$). Dans le premier cas, $s = b \pm \beta a$, d'où $\beta = 0$, $s = b$, $t = c$. Dans le second cas, $s = -b \pm a\delta$ et on ne peut trouver δ tel que $0 \leq 2s < a$. $\square$

Détermination des formes réduites. Soit $q(x, y) = ax^2 + 2bxy + cy^2$ une forme réduite. Les inégalités vérifiées par a et b montrent qu'il n'existe qu'un nombre fini de formes réduites de discriminant donné.

Cherchons, par exemple, les formes réduites de déterminant 11 en utilisant les inégalités de la proposition 2. Comme la partie entière de $\sqrt{11/3}$ est 1 et comme la partie entière de $2\sqrt{11/3}$ est 3, on doit chercher a, b, c tels que $0 < a < 3$, soit $a \in \{1, 2, 3\}$, que $|b| \leq 1$, soit $b \in \{-1, 0, 1\}$, que $-a < 2b \leq a$ ou $0 \leq 2b \leq a$ suivant la valeur de c , enfin qu'il existe c tel que $ac - b^2 = 11$. On étudie les différents cas.

$a = 1$ impose $b = 0$ et $c = 11$;

$a = 2$ impose $b = 0$ ou $b = 1$; $b = 0$ conduit à une impossibilité ; $b = 1$ donne $c = 6$;

$a = 3$ impose $b = 0$ ou $b = \pm 1$; $b = 0$ conduit à une impossibilité ; $b = \pm 1$ donnent $c = 4$.

Finalement, il existe quatre formes réduites de discriminant 11 :

$$\begin{aligned} &x^2 + 11y^2 ; \\ &2x^2 + 2xy + 6y^2 ; \\ &3x^2 \pm 2xy + 4y^2. \end{aligned}$$

Tout entier représenté par une forme de déterminant 11 peut donc s'écrire sous l'une des quatre formes précédentes.

Nous terminerons par une autre démonstration de la réciproque du théorème des deux carrés (voir proposition 21.4 et exercice 21.7).

Théorème. Soit $n = k^2r$ un nombre entier tel que k soit un entier non nul et $r > 1$ un entier sans facteurs carrés. Alors n est une somme de deux carrés si et seulement si les facteurs premiers de r sont, soit égaux à 2, soit congrus à 1 modulo 4.

Démonstration. La condition suffisante a été montrée en 21.4 et la condition nécessaire dans l'exercice 21.7. Montrons comment cette dernière peut être aussi obtenue à partir de la classification de Lagrange. Si $n = k^2r$ est une somme de deux carrés, il existe x_0 et y_0 tels que $k^2r = x_0^2 + y_0^2$; x_0 et y_0 sont non nuls et on peut se ramener au cas où ils sont premiers entre eux. Si p est un facteur premier de r , la proposition 1 ci-dessus montre que p peut être représenté par une forme de discriminant 1. Comme la seule forme réduite de discriminant 1 est $x^2 + y^2$ (c'est ce que donnent les inégalités de la proposition 2), p est proprement représenté par $x^2 + y^2$, donc est égal à 2, ou à 1 modulo 4. $\square$

L'exercice 23.11 donne des exemples d'applications des idées de Lagrange.

Exercices

23.1 Symétries orthogonales

Soit E un espace euclidien.

1) Le produit de deux symétries orthogonales par rapport à deux hyperplans de E est-il commutatif ?

2) Montrer qu'une involution de $GL(E)$ est une symétrie, qu'une involution de $O(E)$ est une symétrie orthogonale. Noter les cas particuliers.

3) Expression d'une symétrie

Soient F un hyperplan de E , v un vecteur non nul engendrant $F^\perp$. Montrer que la symétrie orthogonale $s : E \rightarrow E$ par rapport à F est définie par la formule :

$$s(u) = u - 2 \frac{\langle u, v \rangle}{\|v\|^2} v.$$

23.2 Changement de base

Soient $B = (e_1, e_2, e_3)$ une base de $\mathbb{R}^3$ et q la forme quadratique définie par rapport à cette base par $q(x, y, z) = x^2 + 6y^2 + 56z^2 - 4xy + 14xz - 36yz$.

- 1) Quelle est la matrice A de q par rapport à B ?
- 2) Quelle est la matrice A' de q par rapport à la base $B' = (e_1, 2e_1 + e_2, -3e_1 + 2e_2 + e_3)$? En déduire la signature de q .
- 3) Cette forme admet-elle des vecteurs isotropes ?

23.3 Réduction de formes quadratiques

- 1) Réduire la forme quadratique définie dans la base $B = (e_1, e_2, e_3)$ par $q(x, y, z) = x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz$ en utilisant la méthode de Gauss ; on précisera la nouvelle base $B' = (e'_1, e'_2, e'_3)$. Reprendre le même exercice en utilisant une diagonalisation.
- 2) Réduire en sommes de carrés les formes quadratiques suivantes ; donner leur signature, une base de leur noyau pour celles qui sont dégénérées.
 - a) $q(x, y, z) = xy + yz + zx$ (méthode de Gauss et diagonalisation) ;
 - b) $q(x, y, z, t) = xy + yz + zt + tx$ (méthode de Gauss) ;
 - c) $q(x, y, z, t, u) = xy + yz + zt + tu + ux$ (méthode de Gauss).

23.4 Diagonalisation simultanée de deux formes quadratiques

- 1) Existe-t-il une base simultanément orthogonale pour les deux formes quadratiques définies sur $\mathbb{R}^2$ par $q(x, y) = x^2 - y^2$ et $q'(x, y) = 2xy$?
- 2) Voici une autre démonstration de la proposition 2 de 23.7, en usage parmi les statisticiens qui font de l'analyse en composantes principales et sont, en général, plus tournés vers l'analyse que vers l'algèbre ; en s'intéressant aux extrêmes d'une fonction, elle masque la diagonalisation simultanée. On considère l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire habituel noté $\langle \cdot, \cdot \rangle$; la norme associée est notée $\| \cdot \|$. Soit q une forme quadratique sur E de forme polaire φ .
 - a) Montrer que la fonction $f : E - \{0\} \rightarrow \mathbb{R}$ définie par $f(x) = \frac{q(x)}{\|x\|^2}$ atteint son maximum en un point u_0 de la sphère unité S .
 - b) Calculer la différentielle $df(u)$ de f en un point u .
 - c) Montrer qu'il existe une base orthonormale de E pour le produit scalaire qui est orthogonale pour q .

3) Théorème d'Apollonius. Il s'agit d'Apollonius de Pergé (vers – 262 à – 190)¹, connu pour un magnifique traité sur les coniques dont subsiste sept des huit livres, deux en version arabe seulement. C'est dans ce traité que sont introduits les mots *ellipse*, *parabole*, *hyperbole* que nous utilisons depuis pour les coniques planes.

On se place dans l'espace euclidien $\mathbb{R}^n$ muni de la base canonique et du produit scalaire $\langle \cdot, \cdot \rangle$ usuel. On se donne une forme quadratique définie positive q sur $\mathbb{R}^n$.

a) Soient B et B' deux bases orthonormées pour q , A et A' les matrices du produit scalaire $\langle \cdot, \cdot \rangle$ par rapport à ces bases. Montrer que A et A' ont même polynôme caractéristique. En déduire que les coefficients du polynôme caractéristique de A ne dépendent que de q , et non de la base orthonormée pour q choisie.

b) Montrer que ces coefficients peuvent s'exprimer à l'aide de déterminants de Gram (voir exercice 16.14).

c) Dans $\mathbb{R}^2$, on considère l'ellipse d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ de centre O . Si

$M_1(x_1, y_1)$ et $M_2(x_2, y_2)$ sont des points de l'ellipse, on dit que OM_1 et OM_2 sont des axes conjugués de l'ellipse si (x_1, y_1) et (x_2, y_2) sont orthogonaux pour

la forme quadratique q définie par $q(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$. Montrer que $OM_1^2 + OM_2^2$

est une constante ainsi que l'aire du parallélogramme de sommets O, M_1, M_2 et le point $(x_1 + x_2, y_1 + y_2)$ (ces résultats sont ceux d'Apollonius) ; déterminer ces constantes.

23.5 Forme bilinéaire symétrique $\text{tr}(AA')$

On considère la fonction φ définie sur $M_n(\mathbb{R})$ par $\varphi(A, A') = \text{tr}(AA')$.

1) Vérifier que φ définit une forme bilinéaire symétrique sur $M_n(\mathbb{R})$.

2) Montrer que φ est non dégénérée et possède des vecteurs isotropes pour $n > 1$. Déterminer la signature de φ et une base orthogonale pour φ . (Traiter d'abord le cas $n = 2$.)

23.6 Sommes de carrés dans $K(X)$ et $\mathbb{R}(X_1, \dots, X_n)$

Cet exercice propose d'étudier des résultats de John Cassels (1922-2015) publiés en 1964 dans *Acta mathematica*.

1. On trouve parfois l'orthographe fautive avec deux p et un l ; Pergé était une ville grecque située sur la côte sud de la Turquie actuelle.

1) Soient K un corps, P un polynôme de $K[X]$. On suppose que P est la somme de n carrés de fractions rationnelles de $K(X)$. Montrer que P est la somme de n carrés de $K[X]$. On distinguera les cas suivants.

a) $n = 1$.

b) $\text{car}(K) = 2$.

c) $\text{car}(K) \neq 2$ et -1 est une somme de $n - 1$ carrés dans K (utiliser l'identité

$$x = \left(\frac{x+1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^2.$$

d) On n'est dans aucun des cas précédents. La démonstration demande du travail,

mais on va en être récompensé au **3)**. Si $P = \sum_{1 \leq i \leq n} \left(\frac{Q_i}{R_i}\right)^2$, on peut supposer

tous les R_i égaux à R , dénominateur commun aux différentes fractions et que R est de degré minimal parmi les polynômes permettant cette écriture de P . Si $\deg(R) = 0$, on a gagné. Sinon, des divisions euclidiennes donnent $Q_i = S_i R + T_i$ avec $\deg(T_i) < \deg(R)$; considérer la surface $\mathcal{H}$ de $(K(X))^{n+1}$ d'équation

$PZ^2 = \sum_{1 \leq i \leq n} Z_i^2$ dont un point est $(R, Q_1, \dots, Q_n)$; déterminer un point, noté $(r, q_1, \dots, q_n)$, de $\mathcal{H} \cap (K[X])^{n+1}$, égal, à un facteur près, au point d'intersection de $\mathcal{H}$ et de la droite joignant $(R, Q_1, \dots, Q_n)$ et $(1, S_1, \dots, S_n)$; aboutir à une contradiction.

2) Soient $n \geq 1$ un entier, K un corps de caractéristique différente de 2 et $d \in K$. Montrer que $X^2 + d$ est une somme de n carrés dans $K(X)$ si et seulement si :

- soit -1 est une somme de $n - 1$ carrés dans K ;
- soit d est une somme de $n - 1$ carrés dans K (avec $d = 0$ si $n = 1$).

3) Montrer que $X_1^2 + \dots + X_n^2$ n'est pas la somme de $n - 1$ carrés dans $\mathbb{R}(X_1, \dots, X_n)$.

23.7 Vecteurs isotropes

1) Montrer qu'une forme quadratique sur un $\mathbb{C}$ -espace vectoriel de dimension $n \geq 2$ possède un vecteur isotrope.

2) Montrer qu'une forme quadratique sur un $\mathbb{R}$ -espace vectoriel de dimension ≤ 2 qui n'est ni définie positive, ni définie négative, possède un vecteur isotrope.

3) Dans un espace quadratique (E, q) , une base orthogonale peut-elle posséder un vecteur isotrope ?

23.8 Orthogonal d'une somme et d'une intersection

Soient K un corps, E un K -espace vectoriel de dimension finie, $(F_i)_{1 \leq i \leq n}$ une famille finie de sous-espaces de E , φ une forme bilinéaire symétrique sur E .

1) Montrer que :

$$(F_1 + \dots + F_n)^\perp = \bigcap_{1 \leq i \leq n} (F_i^\perp).$$

2) Montrer que, si φ est non dégénérée, on a de plus :

$$(\bigcap_{1 \leq i \leq n} F_i)^\perp = F_1^\perp + \dots + F_n^\perp.$$

23.9 Groupe orthogonal d'un plan hyperbolique

Soient (E, q) un plan hyperbolique et $B = (e_1, e_2)$ une base de E telle que la matrice de q soit $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

1) Déterminer les matrices des isométries de (E, q) .

Le groupe $O(q)$ formé par ces matrices est appelé groupe orthogonal de (E, q) ; le sous-groupe $SO(q)$ des matrices de déterminant 1 de ce groupe est appelé groupe spécial orthogonal de (E, q) .

2) À quel groupe est isomorphe le groupe $SO(q)$?

3) Pour quel corps le groupe $O(q)$ est-il commutatif ? Quelle est alors la structure de ce groupe ?

23.10 Compléments sur les quaternions

On reprend les notations et résultats de la section 23.10.

1) **Une remarque de Legendre**

a) Montrer qu'un entier de la forme $8n + 7$ ne peut être la somme des carrés de trois entiers.

b) Soient A et B des sommes de carrés de trois entiers ; montrer que le produit $A \times B$ n'est pas nécessairement une somme de trois carrés.

Commentaire. Et si Hamilton avait pu tirer parti de cette remarque pour éviter des années d'efforts infructueux... (Si une formule de multiplication des triplets avait existé, elle aurait eu comme conséquence la stabilité des sommes de triplets par multiplication.)

2) **Paramétrage de $O(\mathbb{R}^3)$**

On munit $\mathbb{R}^3$ de l'orientation définie par la base canonique (e_1, e_2, e_3) (voir 14.10).

Soit $x = a + bi + cj + dk$ un quaternion non nul, $y = bi + cj + dk$. On pose $r = \rho(x)$.

a) Montrer que r est une rotation de $\mathbb{R}^3$, d'axe $u = (b, c, d)$.

b) Montrer que $r = \rho(\lambda x)$ pour tout λ réel non nul.

On suppose désormais $\|y\| = 1$.

c) Montrer qu'il existe un quaternion z tel que $\rho(zxz^{-1})$ soit une rotation de même angle que r et d'axe $(1, 0, 0)$.

d) On note θ l'angle de r . En calculant l'image de $(0, 1, 0)$ par $\rho(zxz^{-1})$, déterminer $\cos(\theta)$, $\sin(\theta)$; vérifier que $\tan(\theta/2) = 1/a$.

3) Composé de deux demi-tours

Soient deux demi-tours s et s' de $\mathbb{R}^3$ d'axes respectifs $u = (b, c, d)$ et $u' = (b', c', d')$. Déterminer leur composé à l'aide d'un produit de quaternions. Que se passe-t-il si les axes sont colinéaires ? orthogonaux ?

23.11 Résultats de Lagrange

1) Soient p un nombre premier tel que $p \equiv 3 \pmod{4}$ et a un entier non nul modulo p .

a) Montrer qu'on a l'alternative : soit a est un carré modulo p , soit $-a$ est un carré modulo p .

b) Montrer qu'on a l'alternative : soit p est un diviseur propre de la forme quadratique $x^2 - ay^2$, soit p est un diviseur propre de la forme $x^2 + ay^2$.

2) Déterminer les différentes formes quadratiques sur $\mathbb{Z}^2$ qui sont définies positives de discriminant 2, 3, 6 et réduites.

3) a) Déterminer les différentes formes quadratiques définies positives réduites de discriminant 15.

b) On note q_1 et q_2 celles de ces formes qui peuvent représenter des nombres impairs ; préciser ces deux formes. Montrer qu'un nombre premier p tel que $p \geq 7$ ne peut être représenté à la fois par q_1 et par q_2 .

c) On note E l'ensemble des nombres premiers $p \geq 6$ tels que $\left(\frac{-15}{p}\right) = 1$ et F l'ensemble des nombres premiers $p \geq 6$ tels que $\left(\frac{-15}{p}\right) = -1$. Déterminer E et F à l'aide de congruences modulo 15.

d) Soit $p \geq 6$ un nombre premier. Montrer que si $p \in E$, alors p est représenté proprement par l'une des deux formes q_1 ou q_2 .

e) Soit p un nombre premier tels que $p \equiv 1 \pmod{15}$. Montrer que $p \equiv 1 \pmod{30}$ et qu'il existe des entiers x, y tels que $x^2 + 15y^2 = p$. Déterminer les entiers x, y quand $p = 151$.

f) Donner des résultats analogues pour les autres valeurs de p dans E .

Solutions

23.1 1) La réponse est oui si la dimension n de l'espace est 1, non si $n \geq 2$. Dans le plan, si D_1 est une droite vectorielle et si $D_2 = r(D_1)$ avec r rotation d'angle α , le produit $s_{D_2} \circ s_{D_1}$ est une rotation d'angle 2α , alors que le produit $s_{D_1} \circ s_{D_2}$ est une rotation d'angle -2α . Ces deux rotations sont différentes si α n'est pas un multiple de $\pi/2$. Cet exemple, se généralise à deux réflexions dans un espace de dimension supérieure.

2) Soit $f \in GL(E)$ telle que $f^2 = \text{id}$. On pose $F = \{u, f(u) = u\}$ et $G = \{u, f(u) = -u\}$. On a $F \cap G = \{0\}$ et $F + G = E$ car

$$u = \frac{1}{2}(u + f(u)) + \frac{1}{2}(u - f(u))$$

avec $u + f(u) \in F$, $u - f(u) \in G$; d'où $E = F \oplus G$. Cette décomposition en somme directe peut aussi être vue comme une conséquence du théorème des noyaux (voir 15.8), puisque f est annulée par le polynôme scindé $X^2 - 1$. On en conclut que f est la symétrie par rapport à F parallèlement à G . Il reste à vérifier que, si $f \in O(E)$, F et G sont orthogonaux : si $u \in F$ et $v \in G$, on a

$$\langle u, v \rangle = \langle f(u), f(v) \rangle = \langle u, -v \rangle = -\langle u, v \rangle,$$

donc $\langle u, v \rangle = 0$.

Notons que, si $f = \text{id}_E$, l'espace G est réduit à 0 et que, si $f = -\text{id}_E$, l'espace F est réduit à 0.

3) La projection orthogonale de u sur $F^\perp = Kv$ est de la forme λv . Pour exprimer l'orthogonalité de v et $u - \lambda v$, on écrit $\langle v, u - \lambda v \rangle = 0$, ce qui donne λ . Comme $s(u) = u - 2\lambda v$, on obtient l'expression donnée dans l'énoncé.

$$\mathbf{23.2 \quad 1)} A = \begin{pmatrix} 1 & -2 & 7 \\ -2 & 6 & -18 \\ 7 & -18 & 56 \end{pmatrix}$$

$$\mathbf{2)} \text{ On pose } P = M(\text{id}, B', B) = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \text{ On a}$$

$$A' = {}^t P A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

d'où $q(x'_1, x'_2, x'_3) = x'^2_1 + 2x'^2_2 - x'^2_3$. La signature de q est $(2, 1)$.

3) La dernière expression de q montre que, par exemple, $e'_1 + e'_3 = -2e_1 + 2e_2 + e_3$ ou $e'_2 + \sqrt{2}e'_3 = (2 - 3\sqrt{2})e_1 + (1 + 2\sqrt{2})e_2 + \sqrt{2}e_3$ sont isotropes.

23.3 1) La matrice de q dans B est $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}$. En appliquant la méthode de Gauss, on trouve :

$$\begin{aligned} q(x, y, z) &= x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz \\ &= (x + 2y + z)^2 - 3y^2 + 2z^2 - 2yz \\ &= (x + 2y + z)^2 - 3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2 \end{aligned}$$

La signature de q est $(2, 1)$. On pose $Q = M(\text{id}, B, B') = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{pmatrix}$ et, en

résolvant le système donné par le changement de variable, on trouve

$$P = Q^{-1} = M(\text{id}, B', B) = \begin{pmatrix} 1 & -2 & -1/3 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix}$$

On a $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7/3 \end{pmatrix} = {}^tPAP$. Les vecteurs de la nouvelle base sont donnés

par les colonnes de P .

On remarque que ${}^tP \neq P^{-1}$: la matrice de changement de base n'est pas orthogonale.

Pour diagonaliser A , on calcule d'abord son polynôme caractéristique $\chi(\lambda) = -(\lambda - 1)(\lambda^2 - 6\lambda + 7)$. Pour $\lambda = -1$, on trouve le vecteur propre $(1, -1, 0)$; en le normant, on

obtient $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)$. On fait de même avec les deux autres valeurs propres $3 + \sqrt{2}$,

$3 - \sqrt{2}$ et on obtient les vecteurs propres associés : $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{\sqrt{2}}\right)$. La

matrice

$$P = M(\text{id}, B', B) = \begin{pmatrix} 1/\sqrt{2} & 1/2 & 1/2 \\ -1/\sqrt{2} & 1/2 & 1/2 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

définit un changement de base orthogonale. Dans la nouvelle base, la matrice de q est diagonale :

$${}^tPAP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3+\sqrt{2} & 0 \\ 0 & 0 & 3-\sqrt{2} \end{pmatrix}$$

Les vecteurs de la nouvelle base correspondent aux colonnes de P . Ils sont orthogonaux et définissent les axes principaux des hyperboloïdes définis par $q(x, y, z) = x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz = m$, avec m réel, dont l'équation dans la nouvelle base est

$$-x'^2 + (3+\sqrt{2})y'^2 + (3-\sqrt{2})z'^2 = m,$$

ce qui permet, par exemple, de connaître des axes orthogonaux (et de calculer leurs longueurs) des ellipses obtenues par intersection des hyperboloïdes avec des plans d'équation $x' = a$, a réel.

2) a) On a $A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. La méthode de Gauss donne :

$$q(x, y, z) = xy + yz + zx$$

$$= (x+z)(y+z) - z^2$$

$$= \frac{1}{4}(x+y+2z)^2 - \frac{1}{4}(x-y)^2 - z^2$$

La signature est $(1, 2)$.

On a l'égalité matricielle $D = {}^tPAP$ avec

$$Q = M(\text{id}, B, B') = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, P = Q^{-1}, D = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & -1/4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

La méthode de diagonalisation donne $\chi(A) = -(\lambda - 1)\left(\lambda + \frac{1}{2}\right)^2$ et

$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}$$

Détaillons ce dernier calcul. On trouve facilement un vecteur propre associé à la valeur propre 1 ; c'est le vecteur $(1, 1, 1)$ qui donne $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ en le normant. L'espace propre associé à la valeur propre $-1/2$ est défini par l'équation $x + y + z = 0$; on doit trouver une base orthonormée dans cet espace ; on prend comme premier vecteur, par exemple : $(1, -1, 0)$; le second vecteur de cette base doit de plus être orthogonal à $(1, -1, 0)$, ce qui donne les deux équations $x + y + z = 0$ et $x - y = 0$; on trouve, par exemple $(1, 1, -2)$; il reste à normer ces deux derniers vecteurs.

b) La méthode de Gauss donne

$$q(x, y, z, t) = (x + z)(y + t) = \frac{1}{4}(x + y + z + t)^2 - \frac{1}{4}(x - y + z - t)^2$$

et on trouve en complétant la base (ou le changement de variable), par exemple

$$Q = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La forme quadratique est dégénérée. Les deux dernières colonnes de la matrice $P = Q^{-1}$ donnent une base du noyau de q : $(-1, 0, 1, 0), (0, -1, 0, 1)$. La signature est $(1, 1)$.

$$\text{c) } q(x, y, z, t, u) = \frac{1}{4}(x + y + z + u)^2 - \frac{1}{4}(x - y + z - u)^2 - \frac{1}{4}(z - 2t - u)^2 + \frac{1}{4}(z - u)^2 + t^2.$$

La signature est $(3, 2)$.

23.4 1) Notons $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et $A' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ les matrices associées à q et q' dans la base canonique de $\mathbb{R}^2$. On trouve $A^{-1}A' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, matrice de la rotation d'angle

$-\pi/2$, qui n'est pas diagonalisable sur $\mathbb{R}$; d'après la proposition 1 de 23.7, la diagonalisation simultanée demandée est impossible.

Cet exercice donne l'occasion de noter que le produit de matrices symétriques n'est pas toujours symétrique.

2) a) On sait que S est un compact de $\mathbb{R}^n$. Par conséquent, la fonction continue f admet un maximum sur S et il existe un point u_0 de S où elle l'atteint. Comme f est constante sur les droites (privées de 0), ce maximum est global.

- b)** La différentielle est une application linéaire. Des résultats classiques d'analyse donnent la valeur de $df(u)$ pour un vecteur $h \in \mathbb{R}^n$:

$$df(u)(h) = \frac{2}{\langle u, u \rangle^2} (\varphi(u, h) \langle u, u \rangle - q(u) \langle u, h \rangle)$$

Pour obtenir ce résultat, on utilise la formule donnant la différentielle d'un quotient : $d\left(\frac{f}{g}\right) = \frac{g df - f dg}{g^2}$; la différentielle de q est donnée par la différentielle de la forme bilinéaire symétrique φ : on déduit de $\varphi(u+h, u+h) = \varphi(u, u) + 2\varphi(h, u) + \varphi(h, h)$ que $dq(u)(h) = 2\varphi(u, h)$; la formule pour le produit scalaire est analogue.

- c)** Comme $df(u_0) = 0$, on a, pour tout $h \in \mathbb{R}^n$,

$$\varphi(u_0, h) \|u_0\|^2 = q(u_0) \langle u_0, h \rangle ;$$

par conséquent, l'espace, de dimension $n - 1$, des vecteurs h orthogonaux à u_0 pour φ coïncide avec l'espace des vecteurs h orthogonaux à u_0 pour le produit scalaire. Une récurrence donne alors le résultat. On peut dire que la fonction f « compare les distances » définies par q et la norme euclidienne.

- 3) a)** Posons $P = M(\text{id}, B', B)$. On a $A' = {}^tPAP$. Comme P est une matrice orthogonale (voir 16.9, corollaire de la proposition 4), on a ${}^tP = P^{-1}$ et $A' = P^{-1}AP$, ce qui montre que A et A' ont même polynôme caractéristique et que les coefficients de ce polynôme ne dépendent pas de la base choisie.

- b)** Notons χ le polynôme caractéristique de A défini par

$$\chi(\lambda) = \det(A - \lambda I) = \sum_{0 \leq k \leq n} (-\lambda)^{n-k} M_k.$$

Le calcul du déterminant montre que M_k est la somme des déterminants des matrices extraites A_{jj} (notation de 14.8) où j est un sous-ensemble à k éléments de $\{1, \dots, n\}$. Il faut enfin noter que, si on note $e_1, \dots, e_n$ les vecteurs de B , la matrice A est la matrice dont les coefficients sont $\langle e_i, e_j \rangle$, ce qui montre que les M_k sont des sommes de déterminants de Gram.

- c)** Notons $B = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$. La base (ae_1, be_2) est orthonormale

pour q . Dans cette base, la matrice du produit scalaire est $\begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}$. Le

polynôme caractéristique de cette matrice est $\lambda^2 - (a^2 + b^2)\lambda + a^2b^2$. Dans toute autre base orthonormale pour q , on obtient le même polynôme caractéristique, en particulier dans la base définie par M_1 et M_2 . Dans cette base, la matrice du produit scalaire est

$$T = \begin{pmatrix} OM_1^2 & \langle \overrightarrow{OM_1}, \overrightarrow{OM_2} \rangle \\ \langle \overrightarrow{OM_1}, \overrightarrow{OM_2} \rangle & OM_2^2 \end{pmatrix}$$

Le polynôme caractéristique est :

$$\lambda^2 - (OM_1^2 + OM_2^2)\lambda + OM_1^2 OM_2^2 - \langle \overrightarrow{OM_1}, \overrightarrow{OM_2} \rangle^2.$$

On a donc $OM_1^2 + OM_2^2 = a^2 + b^2$. D'autre part, le terme constant de ce polynôme est égal à $a^2 b^2$; c'est le déterminant de Gram $\det(T)$; il représente (voir exercice 16.13) le carré de l'aire du parallélogramme indiqué dans l'énoncé, aire qui est donc constante.

23.5 Pour $i, j = 1, \dots, n$, notons E_{ij} la matrice carrée d'ordre n dont tous les coefficients sont nuls, excepté $a_{ij} = 1$. Les n^2 matrices E_{ij} , $i, j = 1, \dots, n$, forment donc une base de l'espace vectoriel $M_n(\mathbb{R})$.

1) Il suffit d'utiliser les propriétés de la trace, en particulier $\text{tr}(AA') = \text{tr}(A'A)$.

2) Étudions d'abord le cas $n = 2$.

Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$, on a $\varphi(A, A') = aa' + bc' + cb' + dd'$; la matrice

associée à φ dans la base $(E_{11}, E_{12}, E_{21}, E_{22})$ est :

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Cette matrice est visiblement de rang 4, donc φ est non dégénérée.

La forme quadratique q associée à φ est définie par $q(a, b, c, d) = a^2 + 2bc + d^2$. La méthode de Gauss donne la décomposition en somme de carrés :

$$q(a, b, c, d) = a^2 + \frac{1}{2}(b+c)^2 - \frac{1}{2}(b-c)^2 + d^2.$$

La signature est $(3, 1)$.

La décomposition de Gauss donne une matrice de passage ; les colonnes de l'inverse de cette matrice conduisent à la base orthogonale $(E_{11}, E_{12} + E_{21}, E_{12} - E_{21}, E_{22})$. Dans cette base, la matrice de φ est la matrice diagonale dont les coefficients diagonaux sont 1, 2, -2, 1.

Il est facile de trouver des vecteurs isotropes : par exemple, $(1, 1, -1, 1)$, ou en fixant a, d , en choisissant b non nul, on peut déterminer un c tel que $q(a, b, c, d) = 0$.

Dans le cas général, si $A = (a_{ij})$, on a :

$$q(A) = \sum_{1 \leq i \leq n} a_{ii}^2 + \frac{1}{2} \sum_{1 \leq i < j \leq n} [(a_{ij} + a_{ji})^2 - (a_{ij} - a_{ji})^2].$$

La matrice associée est de rang n et la signature est $\left(\frac{n(n+1)}{2}, \frac{n(n-1)}{2}\right)$. La base formée par les E_{ii} , $1 \leq i \leq n$ et les $E_{ij} + E_{ji}$, $E_{ij} - E_{ji}$ pour $1 \leq i < j \leq n$ est orthogonale pour φ .

23.6 1) a) Si $n = 1$, il existe $Q, R \in K[X]$, qu'on peut supposer premiers entre eux, tels que $P = \left(\frac{Q}{R}\right)^2$. On a $P R^2 = Q^2$; comme $K[X]$ est un anneau factoriel, il ne peut exister de polynôme irréductible divisant R , puisqu'alors il diviserait Q^2 donc Q ; par conséquent, $R \in K^*$. On en déduit que P est un carré dans $K[X]$.

b) Si P est une somme de n carrés de $K(X)$, il existe des polynômes $Q_i, R_i \in K[X]$, $1 \leq i \leq n$, de $K[X]$ tels que

$$P = \sum_{1 \leq i \leq n} \left(\frac{Q_i}{R_i}\right)^2.$$

Comme K est de caractéristique 2, on a

$$P = \left(\sum_{1 \leq i \leq n} \frac{Q_i}{R_i}\right)^2.$$

La somme $\sum_{1 \leq i \leq n} \frac{Q_i}{R_i}$ peut s'écrire comme une fraction rationnelle $\frac{Q}{R}$ de $K(X)$, ce qui ramène au cas précédent.

c) Posons $-1 = \sum_{1 \leq i \leq n-1} a_i^2$ avec $a_i \in K$ pour $1 \leq i \leq n-1$. Dans ce cas, tout polynôme de $K[X]$ s'écrit comme une somme de n carrés, sans qu'il soit besoin de supposer que P soit une somme de n carrés de $K(X)$. En effet, comme $\text{car}(K) \neq 2$, on a :

$$P = \left(\frac{P+1}{2}\right)^2 + \sum_{1 \leq i \leq n-1} a_i^2 \left(\frac{P-1}{2}\right)^2$$

d) Un point de la droite indiquée est de la forme

$$(R, Q_1, \dots, Q_n) + \lambda(1, S_1, \dots, S_n)$$

avec $\lambda \in K^*$. En écrivant qu'il vérifie l'équation de $\mathcal{H}$, on trouve

$$P(R + \lambda)^2 = \sum_{1 \leq i \leq n} (Q_i + \lambda S_i)^2; \text{ en simplifiant, on obtient :}$$

$$\lambda \left(\sum_{1 \leq i \leq n} S_i^2 - P \right) = 2 \left(PR - \sum_{1 \leq i \leq n} S_i Q_i \right).$$

On peut prendre λ égal à la valeur donnée par cette relation et multiplier les composantes du point obtenu par le dénominateur $\sum_{1 \leq i \leq n} S_i^2 - P$. On trouve alors

$$r = R \left(\sum_{1 \leq i \leq n} S_i^2 - P \right) + 2 \left(PR - \sum_{1 \leq i \leq n} S_i Q_i \right) = \sum_{1 \leq i \leq n} T_i^2 / R$$

avec $\deg(r) < \deg(R)$ (les vérifications demandent de bien conduire les calculs), ce qui donne la contradiction cherchée. On notera que r n'est pas nul car $\sum_{1 \leq i \leq n} T_i^2$ ne l'est pas puisqu'on a exclu la possibilité du c).

Cette démonstration donne une procédure pour écrire un polynôme dont on connaît une expression comme somme de n carrés de fractions rationnelles de $K(X)$ comme somme de n carrés de polynômes de $K[X]$.

2) Si $d = \sum_{1 \leq i \leq n-1} a_i^2$, $X^2 + d$ est bien une somme de n carrés dans $K(X)$. Si $-1 = \sum_{1 \leq i \leq n-1} a_i^2$, on applique le c) de la première question. Dans ces deux cas, on obtient même que $X^2 + d$ est une somme de n carrés dans $K[X]$.

Réciproquement, si $X^2 + d$ est une somme de n carrés dans $K(X)$, c'est une somme de n carrés dans $K[X]$ d'après le 1) et on peut écrire $X^2 + d = \sum_{1 \leq i \leq n} Q_i^2$ avec $Q_i \in K[X]$ pour $1 \leq i \leq n$. Notons q le degré maximum des Q_i , $1 \leq i \leq n$, a_i le coefficient du terme de degré q de Q_i pour $i = 1, \dots, n$ (éventuellement, $a_i = 0$). On a $q \geq 1$ et le coefficient du terme de degré q de $\sum_{1 \leq i \leq n} Q_i^2$ est $\sum_{1 \leq i \leq n} a_i^2$, une somme de n carrés non tous nuls de K .

On doit montrer que -1 ou d est une somme de $n - 1$ carrés de K . Pour cela, on suppose que -1 ne vérifie pas cette condition et on va montrer que d la satisfait.

Si $q > 1$, l'égalité entre les coefficients des termes de degré $2q$ donne $a = 0$; comme l'un des a_i , disons a_1 , est non nul, on obtient, en divisant par a_1^2 , que -1 est une somme de $n - 1$ carrés dans K , ce qu'on vient d'exclure.

Si $q = 1$, les Q_i s'écrivent sous la forme $Q_i = a_i X + b_i$ pour $1 \leq i \leq n$; on a donc

$$d = \sum_{1 \leq i \leq n} (a_i x + b_i)^2 - x^2 \text{ pour tout } x \in K.$$

On peut choisir x de façon que $(a_n x + b_n)^2 - x^2 = 0$; en effet, cette équation équivaut à $(a_n \pm 1)x + b_n = 0$; si $a_n \neq 1$, $x = b_n(1 - a_n)^{-1}$ est solution ; si $a_n = 1$, $x = -b_n/2$ est solution puisque $\text{car}(K) \neq 2$. Cette valeur de x permet d'écrire d comme somme de $n - 1$ carrés dans K .

3) Ce résultat est assez frappant et justifie, espérons-le, l'étude précédente. On raisonne par récurrence sur $n \geq 1$. Comme X_1^2 n'est pas 0, la base de la récurrence est vraie. On suppose alors le résultat vrai pour tout $k < n$, avec $n > 1$, et on pose

$$K = \mathbb{R}(X_1, \dots, X_{n-1}), d = X_1^2 + \dots + X_{n-1}^2;$$

on voit que $d \in K$. Si $X_1^2 + \dots + X_n^2 = X_n^2 + d$ était la somme de $n - 1$ carrés dans $\mathbb{R}(X_1, \dots, X_n) = K(X_n)$, le **2)** montre que $d = X_1^2 + \dots + X_{n-1}^2$ serait la somme de $n - 2$ carrés dans K , ce qui contredit l'hypothèse de récurrence et permet de conclure.

23.7 1) On sait qu'il existe une base B telle que q s'écrive comme une somme de carrés dans cette base : $q(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2$. En prenant x_1 non nul, $x_2 = \iota x_1$, $x_k = 0$ pour $k \geq 3$, on obtient un vecteur isotrope.

2) Notons φ la forme polaire de q . Il existe des vecteurs u et v tels que $q(u) > 0$ et $q(v) < 0$. On a

$$q(u + \lambda v) = q(u) + 2\lambda\varphi(u, v) + \lambda^2 q(v).$$

Comme $q(u)q(v) < 0$, l'équation $q(u + \lambda v) = 0$ a un discriminant strictement positif ; elle a deux racines réelles λ_1 et λ_2 définissant des vecteurs isotropes pour q : $u + \lambda_1 v$ et $u + \lambda_2 v$. On peut également remarquer que, dans la décomposition de q en sommes de carrés, il existe au moins un coefficient strictement positif et un coefficient strictement négatif et construire un vecteur isotrope à l'aide de ces deux termes.

3) Oui, par exemple, la base canonique de $\mathbb{R}^3$ est orthogonale pour la forme quadratique définie par $q(x, y, z) = x^2 + y^2$ et le vecteur $(0, 0, 1)$ est isotrope.

23.8 1) Si x est orthogonal à la somme des F_i , il est orthogonal à chacun d'entre eux, donc appartient à l'intersection de $F_i^\perp$. La réciproque est aussi évidente.

2) En remplaçant les F_i par les $F_i^\perp$ dans la relation précédente, on a $(F_i^\perp + \dots + F_n^\perp)^\perp = \cap_{1 \leq i \leq n} ((F_i^\perp)^\perp)$. Comme φ est non dégénérée, $(F_i^\perp)^\perp = F_i$. En prenant les orthogonaux des deux membres, on obtient la relation.

23.9 1) Une isométrie de $O(E, q)$ est définie par une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que ${}^tMAM = A$ ce qui donne les équations :

$$\begin{cases} ac = 0 \\ bd = 0 \\ ad + bc = 1 \end{cases}$$

On a donc deux possibilités : $a = d = 0$ et $bc = 1$ ou $b = c = 0$ et $ad = 1$; les matrices

des isométries de $O(q)$ sont de l'une des deux formes $M_a = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ ou $N_b = \begin{pmatrix} 0 & b \\ b^{-1} & 0 \end{pmatrix}$.

2) Le groupe $SO(q)$ est formé des matrices M_a ; il est isomorphe à K^* et commutatif.

3) On a $M_a N_b = N_{ab}$ et $N_b M_a = N_{a^{-1}b}$. Ces deux matrices sont différentes, sauf si $a^2 = 1$. Dans un corps, cette condition impose $a = \pm 1$; le seul corps où cette condition est vérifiée est le corps $\mathbb{F}_3$.

Par conséquent, le groupe orthogonal d'un plan hyperbolique n'est pas commutatif, sauf dans le cas particulier où le corps de base est $\mathbb{F}_3$. Dans ce dernier cas, le groupe orthogonal d'un plan hyperbolique possède quatre éléments, l'identité et trois éléments d'ordre 2 (–id et deux symétries orthogonales) ; les matrices de ces transformations sont :

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

Le groupe $O(E, q)$ est dans ce cas isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.

23.10 1) a) Comme les carrés modulo 8 sont 0, 1, 4, la somme de trois d'entre eux ne vaut jamais 7 mod 8.

- b)** Prendre $A = 3 = 1^2 + 1^2 + 1^2$ et $B = 5 = 0^2 + 1^2 + 2^2$. Leur produit 15 n'est pas une somme de trois carrés car $15 = 7 \pmod{8}$ (une vérification directe est possible puisque les seuls carrés inférieurs à 15 sont 1, 4, 9). Par contre, 15 est une somme de quatre carrés : $15 = 9 + 4 + 1 + 1$.

2) a) On sait que r est une rotation. Comme

$$\rho(x)(y) = xyx^{-1} = (a+y)y(a-y)/N(x) = y,$$

puisque y commute avec a et y , l'axe de r est $u = \tau(y) = (b, c, d)$.

b) On a vu que le noyau de ρ est $\mathbb{R}^*$, donc

$$\rho(\lambda x) = \rho(\lambda)\rho(x) = \rho(x) = r.$$

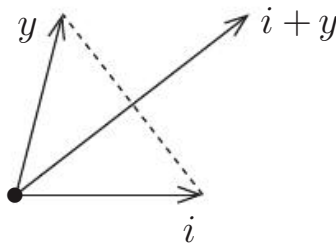
c) D'après l'exercice 16.5 3), il suffit de choisir z tel que $\rho(z)$ soit une rotation transformant u en $(1, 0, 0)$. Prenons $z = i + y$; on a

$$N(z) = (b+1)^2 + c^2 + d^2 = 2b+2.$$

On sait que $\rho(z)$ est un demi-tour et on a

$$\begin{aligned}\rho(z)(y) &= (i+y)y(-i-y)/(2b+2) \\ &= -(iy-1)(i+y)/(2b+2) = \dots = i;\end{aligned}$$

la figure suivante illustre la situation.



d) Les vecteurs correspondants aux quaternions purs i, j, k forment une base orthogonale de $\mathbb{R}^3$; comme i est l'axe de la rotation $\rho(zxz^{-1})$, on trouve $\cos(\theta)$, $\sin(\theta)$ en calculant l'image de j . On a $zxz^{-1} = zaxz^{-1} + zyz^{-1} = a + i$, puis

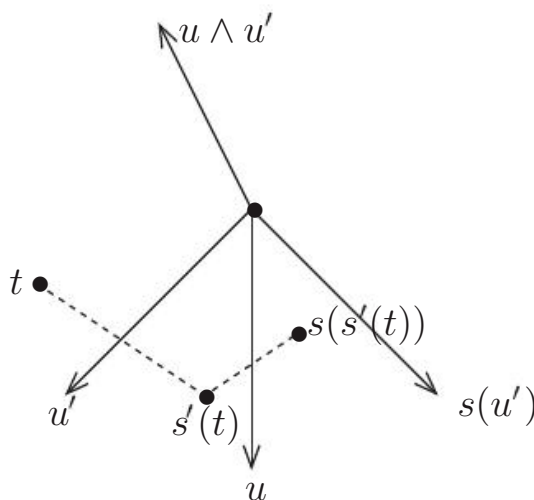
$$(a+i)j(a-i)/(a^2+1) = \dots = \frac{a^2-1}{a^2+1}j + \frac{2a}{a^2+1}k;$$

on en déduit $\cos(\theta) = \frac{a^2-1}{a^2+1}$, $\sin(\theta) = \frac{2a}{a^2+1}$ et, pour $a \neq 0$:

$$\tan(\theta/2) = \frac{\sin(\theta)}{1+\cos(\theta)} = 1/a.$$

Ces calculs permettent aussi de trouver un quaternion correspondant à une rotation donnée de $\mathbb{R}^3$.

3) Posons $y = bi + cj + dk$ et $y' = b'i + c'j + d'k$. Le composé $s \circ s'$ est défini par $\rho(yy')$. Si u et u' sont colinéaires, yy' est un réel et $\rho(yy') = \text{id}$. Sinon, on a $yy' = -(bb' + cc' + dd') + (cd' - c'd)i + (b'd - bd')j + (bc' - b'c)k$. Le quaternion pur $(cd' - c'd)i + (b'd - bd')j + (bc' - b'c)k$ s'identifie au produit vectoriel $u \wedge u'$. Par conséquent, $\rho(yy')$ est une rotation d'axe $u \wedge u'$. L'angle de $s \circ s'$ s'obtient en remarquant que $s(s'(u')) = s(u')$ est le symétrique de u' par rapport à u , ce que montre $\rho(yy')(y') = (yy')y'(yy')^{-1} = yy'y^{-1}$.



Si les axes sont orthogonaux, yy' est un quaternion pur, donc $\rho(yy')$ est un demitour.

23.11 1) a) En posant $p = 4k + 3$, la loi de réciprocité quadratique donne

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = (-1)^{2k+1} = -1.$$

Comme $\left(\frac{-a}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{a}{p}\right) = -\left(\frac{a}{p}\right)$, l'alternative indiquée est vraie.

b) Si p est un diviseur propre de la forme $x^2 - ay^2$, il existe x_0, y_0 premiers entre eux et k tels que $x_0^2 - ay_0^2 = kp$. Comme $y_0 \not\equiv 0 \pmod{p}$, a est le carré de $x_0 y_0^{-1}$ dans $\mathbb{F}_p$ et $\left(\frac{a}{p}\right) = 1$. Réciproquement, si a est un carré modulo p , il existe x_0 tel que $x_0^2 - a = kp$, ce qui montre que p est un diviseur propre de la forme $x^2 - ay^2$. On montre de même que p est un diviseur propre de la forme $x^2 + ay^2$ si et seulement si $\left(\frac{-a}{p}\right) = 1$. Le **a)** permet de conclure.

2) Les majorations de la proposition 2 de 23.11 conduisent aux formes $x^2 + 2y^2$, $x^2 + 3y^2$, $2x^2 + 2xy + 2y^2$, $x^2 + 6y^2$, $2x^2 + 3y^2$.

3) a) La condition $a \leq 2\sqrt{\frac{\Delta}{3}}$ impose $0 < a \leq 4$.

Si $a = 1$, alors $b = 0$ d'où $c = 15$; on obtient la forme $q_1(x, y) = x^2 + 15y^2$.

Si $a = 2$, $b = 0$ conduit à $2c = 15$, une impossibilité ; $b = 1$ conduit à $2c = 16$ et à la forme $2x^2 + 2xy + 8y^2$ dont les valeurs sont paires.

Si $a = 3$, $b = 0$ conduit à $q_2(x, y) = 3x^2 + 5y^2$; $b = \pm 1$ conduit à $3c = 16$, une impossibilité.

Si $a = 4$, $b = 0$ conduit à $4c = 15$, une impossibilité ; $b = \pm 1$ conduit à $4c = 16$, d'où $c = 4$ et $b = 1$ et à la forme $4x^2 + 2xy + 4y^2$ dont les valeurs sont paires ; $b = 2$ conduit à $4c = 19$, une impossibilité.

b) Si p est représenté par q_1 et q_2 , il existe x, y, u, v tels que $p = x^2 + 15y^2 = 3u^2 + 5v^2$. Comme p est premier, v n'est pas divisible par 3, donc on a modulo 3 : $v^2 = 1$ donc $x^2 = 5 = 2$, ce qui est impossible.

$$\begin{aligned} \text{c) On a } \left(\frac{-15}{p} \right) &= \left(\frac{-1}{p} \right) \left(\frac{3}{p} \right) \left(\frac{5}{p} \right) = \\ &= (-1)^{(p-1)/2} \times (-1)^{(p-1)/2} \left(\frac{p}{3} \right) \times \left(\frac{p}{5} \right) = \\ &= \left(\frac{p}{3} \right) \times \left(\frac{p}{5} \right). \end{aligned}$$

On raisonne modulo 15 ; il y a $\varphi(15) = 8$ valeurs possibles de p : 1, 2, 4, 7, 8, 11, 13, 14. On construit le tableau suivant.

p	-1	2	4	7	8	11	13	14
$\left(\frac{p}{3} \right)$	1	-1	1	1	-1	-1	1	-1
$\left(\frac{p}{5} \right)$	1	-1	1	-1	-1	1	-1	1
$\left(\frac{p}{3} \right) \left(\frac{p}{5} \right)$	1	1	1	-1	1	-1	-1	-1

On en déduit E est l'ensemble des nombres premiers égaux à 1, 2, 4, 8 modulo 15 et que F est l'ensemble des nombres premiers égaux à 7, 11, 13, 14 modulo 15.

d) Soit $p \in E$, avec $p \geq 7$. Il existe x, k tels que $x^2 = -15 + kp$; par conséquent p est un diviseur propre de la forme $x^2 + 15y^2$; d'après les résultats de Lagrange, p est représenté proprement par une forme de même discriminant, donc par q_1 ou q_2 .

e) Si $p = 1 + k \times 15$, alors k est nécessairement pair (si k est impair, alors $1 + 15k$ est pair et ne peut être premier) donc $p = 1 \pmod{30}$. Si p est représenté par q_2 , il existe u, v tels que $3u^2 + 5v^2 = p = 1 + 15k$. On a donc $5v^2 = 1 \pmod{3}$, d'où $v^2 = 2 \pmod{3}$, ce qui est impossible. Par conséquent, un nombre premier tels que $p = 1 \pmod{30}$ est représenté par q_1 et il existe des entiers x, y tels que $x^2 + 15y^2 = p$.

Un programme Excel qui calcule $p - 15y^2$ et regarde si la racine est entière conduit très vite aux solutions, au moins pour de petites valeurs de p ; on peut évidemment changer les signes de x et y et je ne donne pas toutes les solutions.

On trouve :

$31 = 16 + 15$, $61 = 1 + 15.4$, $151 = 16 + 15.9$, $181 = 121 + 15.4$,
 $211 = 196 + 15$, $241 = 1 + 15.16$, $271 = 256 + 15$, $301 = 7 \times 43$ n'est pas premier,
 $331 = 196 + 15.9$, $361 = 121 + 15.16$, $391 = 16 + 15.25$, $421 = 361 + 15.4$, etc.

f) Si $p = 2 \pmod{15}$ ou si $p = 8 \pmod{15}$, c'est-à-dire si $p = 17 \pmod{30}$ ou si $p = 23 \pmod{30}$, on ne peut avoir $x^2 + 15y^2 = 2 + 15k$ (ou $8 + 15k$) qui est impossible modulo 3, donc p est représenté par $3x^2 + 5y^2$; par exemple, $17 = 3 \times 4 + 5$, $23 = 3 + 5 \times 4$, etc.

Si $p = 4 \pmod{15}$, c'est-à-dire si $p = 19 \pmod{30}$, on ne peut avoir $3x^2 + 5y^2 = \dots 5k$ qui est impossible modulo 3, donc p est représenté par $x^2 + 15y^2$; par exemple, $19 = 4 + 15$, $79 = 64 + 15 \times 4$, $109 = 49 + 15 \times 4$, etc.

ANNEXES POUR LE PLAISIR, AVEC BEAUCOUP DE PRÉSENT

24

J'ai pensé que ceux et celles qui utilisent un livre d'algèbre pour leurs études auraient plaisir à lire, pour compléter leurs connaissances, quelques pages sur des problèmes actuels. Il ne s'agit pas ici de donner les détails mathématiques qu'ils renferment, ce qui pourrait être fort long et d'une grande difficulté, mais juste de présenter quelques très beaux résultats et d'essayer d'en donner une idée (tous mériteraient des prix de beauté).

J'ai déjà parlé d'applications utilisant essentiellement de l'algèbre, comme la théorie des codes et la cryptographie (voir chapitre 22), l'analyse des données (voir 16.8), la robotique (voir 23.9 et exercice 23.10). D'autres parties des mathématiques utilisent aussi beaucoup d'algèbre (la géométrie, l'analyse, l'analyse numérique...), certaines parties de la physique également (la relativité, la physique quantique...). J'ai choisi de parler d'un domaine foisonnant, la théorie des nombres, plein de questions non résolues et plus facilement accessibles aux non-spécialistes. J'ai choisi de parler de travaux récents de Manjul Bhargava, de conjectures importantes et de travaux anciens de Dirichlet et Riemann qui sont toujours au centre des mathématiques d'aujourd'hui. Vous trouverez dans les pages qui suivent des indications que j'ai cherchées à rendre le plus compréhensible possible sur des travaux qui sont menés aujourd'hui dans ce domaine, des questions récentes qui se posent. Mon livre d'histoire [ESC3] peut compléter votre lecture. J'espère vous intéresser, réussir à vous faire apercevoir ces domaines de recherches, et vous donner de beaux rêves... Ne doutez pas qu'il existe encore plein d'autres parties des mathématiques qui contiennent aussi de belles questions non résolues !

24.1 MANJUL BHARGAVA (NÉ EN 1974)

Courte biographie. Manjul Bhargava a obtenu la médaille Fields en 2014, au congrès de Séoul. Il est né au Canada, a grandi aux États-Unis ; vers deux-trois ans, il aimait multiplier de grands nombres.

Manjul Bhargava se sent de culture indienne ; on parlait hindi à la maison. Il connaît bien la poésie sanskrite, grâce à son grand-père, directeur du département de sanskrit de l'université du Rajasthan. Sa mère, mathématicienne, enseignait à l'université privée Hofstra à New-York et son père était chimiste.

C'est aussi un remarquable joueur de tablas, instrument qu'il a étudié, entre autres, avec le célèbre Zakir Hussain. Il en joue tous les jours, en particulier quand un problème mathématique lui résiste et cela lui permet souvent, explique-t-il, de trouver de nouvelles idées pour le résoudre. Il dit avoir appris dès l'enfance combien les liens entre les facultés artistiques et scientifiques du cerveau permettent de progresser dans les deux domaines. Pour lui, les mathématiciens font des mathématiques pour des raisons très proches de celles d'un artiste pratiquant son art (un musicien jouant de la musique, etc.) : il s'agit de se comprendre soi-même et de comprendre le monde qui nous entoure.

Manjul Bhargava a soutenu sa thèse, dirigée par Andrew Wiles (voir 19.14), en 2001. Il est professeur à l'université de Princeton depuis 2003.

Les domaines de recherches de Manjul Bhargava tournent autour de la théorie des nombres. On raconte que c'est en manipulant le Rubik's cube qu'il aurait compris comment créer de nouvelles lois sur l'ensemble des formes quadratiques.

Représentations des entiers par des sommes de carrés. En 23.9, nous avons énoncé sans démonstration le théorème des quatre carrés : la forme quadratique définie positive $q(x, y, z, t) = x^2 + y^2 + z^2 + t^2$ est universelle au sens où elle peut prendre toutes les valeurs entières quand on donne à x, y, z, t des valeurs entières. En 1916, le génial Srinivasa Ramanujan (1887-1920) proposa une généralisation de ce théorème. Il donna une liste de 55 formes quadratiques diagonales $q(x, y, z, t) = ax^2 + by^2 + cz^2 + dt^2$ qui étaient universelles. Si on note ces formes par le quadruplet de leurs coefficients entre crochets, voici le début de cette liste :

[1,1,1,1] [1,1,1,2] [1,1,1,3] [1,1,1,4] [1,1,1,5] [1,1,1,6] [1,1,1,7]

[1,1,2,2] [1,1,2,3] [1,1,2,4] [1,1,2,5] [1,1,2,6] [1,1,2,7] [1,1,2,8]

La seule forme sur laquelle Ramanujan s'est trompé est la forme [1,2,5,5], ce qu'on vérifie immédiatement puisqu'elle ne représente pas 15 (vérifier !).

En 1993, en donnant un cours à Princeton sur les formes quadratiques, John Conway (né en 1937), pensa qu'il devait exister une constante C telle que toute forme quadratique à matrice entière (les coefficients des termes $x_i x_j$ sont pairs, voir 23.10, écriture matricielle) représentant tous les entiers jusqu'à C devait être universelle. Il pensait aussi qu'il existait une autre constante C' , plus

grande, pour les formes quadratiques à coefficients entiers (leurs matrices peuvent comporter des $1/2$). Il n'avait d'abord pas d'idée sur la valeur de ces constantes, mais des discussions le conduisirent rapidement à penser que $C = 15$, $C' = 290$, des valeurs bien inférieures à celles qu'il avait d'abord envisagées, puis à donner dans les semaines suivantes une démonstration, compliquée, avec des collègues, pour $C = 15$.

Manjul Bhargava trouva des démonstrations du théorème des 15 en 2000 et du théorème des 290 en 2005 en collaboration avec Jonathan Hanke. Les différents résultats obtenus par Bhargava sont des énoncés simples et ne nécessitent pas de gros bagage mathématique. Les démonstrations sont très techniques, virtuoses, et étudient beaucoup de cas. L'utilisation de journées d'ordinateurs est nécessaire.

Le théorème des 15. 1) Si une forme quadratique à coefficients entiers ayant une matrice à coefficients entiers représente les entiers de 1 à 15, alors elle représente tous les entiers naturels.

2) En fait, il suffit que la forme représente les 9 entiers 1, 2, 3, 5, 6, 7, 10, 14 et 15.

Le théorème des 290. 1) Si une forme quadratique à coefficients entiers représente les entiers de 1 à 290, alors elle représente tous les entiers naturels.

2) En fait, il suffit que la forme représente les 29 entiers 1, 2, 3, 5, 6, 7, 10, 13, 14, 15, 17, 19, 21, 22, 23, 26, 29, 30, 31, 34, 35, 37, 42, 58, 93, 110, 145, 203, 290.

Autres théorèmes sur le même sujet. Plus généralement, si S est un sous-ensemble de l'ensemble $\mathbb{N}$ des entiers, il existe un sous-ensemble $T \subset S$ minimal (pour la relation d'inclusion) parmi tous les sous-ensembles tels que toute forme quadratique à coefficients entiers qui représente les entiers de T représente tous les entiers de S . Par exemple, si S est l'ensemble des nombres premiers, l'ensemble T est formé des nombres premiers 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 67, 73.

Commentaires. Ces résultats simples et spectaculaires sont vraiment superbes !

1) Le théorème des quatre carrés de Lagrange (voir 23.9) ne résulte pas du théorème des 15, car il sert à sa démonstration.

2) Comme application, il suffit de vérifier que la forme quadratique $x^2 + y^2 + 2z^2 + 5t^2$ représente les 9 entiers 1, 2, 3, 5, 6, 7, 10, 14, 15, par exemple $10 = 1 + 4 + 5$, $15 = 1 + 9 + 5$, pour affirmer qu'il est possible fabriquer tous les entiers avec cette forme.

24.2 GROUPE DES POINTS D'UNE COURBE ELLIPTIQUE

Nous avons présenté les courbes elliptiques en 22.13 dans un chapitre sur la cryptographie et nous avons parlé de courbes définies sur les corps finis. Nous nous intéressons ici aux courbes définies sur le corps $\mathbb{Q}$ des nombres rationnels.

Soit (E) une courbe elliptique d'équation :

$$y^2 = x^3 + ax + b$$

avec a, b rationnels et $x^3 + ax + b$ n'ayant que des racines simples.

La courbe (E) est un ensemble de couples de points réels (x, y) vérifiant l'équation donnée ; elle est symétrique par rapport à l'axe des x . On sait qu'on peut définir par une construction géométrique une addition sur les points de (E) (voir les formules donnant $P_1 + P_2$ dans 22.13). Cette addition fournit une loi de groupe abélien sur les points de (E) . On peut restreindre cette addition aux points rationnels de (E) (points dont les deux coordonnées sont rationnelles) pour obtenir le groupe $E(\mathbb{Q})$. La détermination de la structure de ce groupe a conduit à des résultats très beaux et très profonds.

Structure du groupe d'une courbe elliptique. Un résultat classique, qui n'a pas été évoqué au chapitre 18, est qu'un groupe abélien G engendré par un nombre fini d'éléments (on dit que G est de type fini) est une somme directe d'un sous-groupe T dit de torsion, formé des points d'ordre fini de G et d'un sous-groupe de G isomorphe à un produit $\mathbb{Z}^r$:

$$G \sim T \oplus \mathbb{Z}^r.$$

Poincaré a indiqué en 1901 l'intérêt de montrer que le groupe $E(\mathbb{Q})$ est de type fini ; il a conjecturé qu'il en était ainsi et a souligné l'intérêt de chercher le r minimal.

La conjecture de Poincaré est démontrée vingt ans plus tard par Louis Mordell (1888-1972) dans son article : On the rational solutions of the indeterminate equations of the 3rd and 4th degrees, *Proceedings of the Cambridge Philosophical Society*, vol. 21, 1922, p. 179-192.

Théorème de Mordell-Weil. Étant donnée une courbe elliptique (E) , le groupe $E(\mathbb{Q})$ est abélien de type fini ; en particulier, le sous-groupe T des points de torsion est fini.

Autrement dit, il existe des points $P_1, \dots, P_r, Q$ de la courbe elliptique tels que tout point P de la courbe s'écrive :

$$\sum_{1 \leq i \leq r} n_i P_i + Q,$$

où Q est un point de torsion et où les P_i sont linéairement indépendants dans le groupe. L'entier r est appelé rang de la courbe elliptique. Quand $r = 0$, le groupe de la courbe est égal à son groupe de torsion et est fini.

Une des idées de la démonstration met en œuvre la méthode de descente infinie de Fermat.

Dans sa thèse publiée en 1928, André Weil généralise le théorème de Mordell en simplifiant la démonstration.

De nombreux résultats ont été obtenus depuis, mais on ne dispose pas d'algorithme général pour calculer le rang (on ne dispose d'algorithmes que pour certains types d'équations).

On a supposé que le rang pouvait prendre des valeurs aussi grandes qu'on le souhaite, mais on n'a réussi à calculer que des rangs de petites valeurs. En 1938, on connaissait quelques courbes de rang ≥ 3 , en 1997 de rang ≥ 12 , en 1994 de rang ≥ 21 . Le record actuel est une courbe de rang $r \geq 28$ (la valeur précise n'est pas connue), construite par Noam Elkies (né en 1966) en 2006. L'équation de la courbe d'Elkies comporte de grands nombres :

$$y^2 + xy + y = x^3 - x^2 - \dots$$

$$20067762415575526585033208209338542750930230312178956502x + \dots$$

$$3448161179503055646703298569039072037485594435\dots$$

$$9\,319180361266008296291939448732243429.$$

Elkies donne ensuite les coordonnées de 28 points indépendants, toutes aussi énormes !

En n'ayant qu'un seul exemple de rang ≥ 28 , on est loin d'avoir des exemples de grande taille permettant de penser que le rang peut être aussi grand qu'on veut, mais en théorie des nombres, cela ne prouve rien. Pensons à l'exemple des suites arithmétiques finies faites de nombres premiers, par exemple, la suite de raison 6 formée de 5, 11, 17, 23, 29. La plus longue suite arithmétique de nombres premiers qu'on connaisse aujourd'hui à une longueur $L = 26$, ce qui semble fort peu. Cependant, Ben Green (né en 1977) et Terence Tao (né en 1975, médaille Fields 2006) ont montré en 2004 qu'on pouvait trouver des suites arithmétiques de nombres premiers avec L aussi grand qu'on veut. Difficile de penser ce résultat exact quand on ne connaît que des exemples avec des L très petits.

Revenons aux courbes elliptiques. Ont-elles ou non des rangs aussi grands qu'on veut ? Certains le croient, d'autres non et viennent de publier leur conviction que le rang des courbes elliptiques serait en fait borné. L'avenir dira qui a raison.

Pour la partie de torsion, le résultat suivant de Barry Mazur (né en 1937) démontré en 1977 permet de bien la décrire (on a des exemples de chacun des cas indiqués).

Théorème de Mazur. Le groupe de torsion T d'une courbe elliptique $E(\mathbb{Q})$ est isomorphe :

- soit à l'un des 11 groupes cycliques $\mathbb{Z}/n\mathbb{Z}$ avec $1 \leq n \leq 10$ ou $n = 12$;
- soit à l'un des quatre groupes produits $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z}$ avec $1 \leq n \leq 4$.

Exemple. Déterminons le groupe de la courbe elliptique (E) d'équation

$$y^2 = x^3 - x$$

en suivant la démarche décrite dans le livre de Henry McKean et Victor Moll : *Elliptic curves* (1999). Le graphe de cette courbe a été donné en 22.13.

La courbe possède trois points d'ordre 2, les points à tangente verticale : $(-1,0)$, $(0,0)$, $(1,0)$. Le point O , élément neutre du groupe est d'ordre 1.

On va montrer que la courbe n'a pas d'autre point rationnel, autrement dit que $E(\mathbb{Q}) \approx \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Pour cela, on va appliquer la méthode de descente infinie de Fermat.

On suppose qu'il existe un autre point rationnel $(u,v) = (a/b, c/d)$ de la courbe (les deux fractions sont mises sous forme irréductible) ; on a $u \neq -1, 0, 1$ et $v^2 = u^3 - u$.

1) Si $-1 < u < 0$, on obtient un autre point de la courbe en considérant le point : $(u', v') = (-1/u, v/u^2)$ avec $u' > 1$. C'est bien un point de la courbe puisque $v'^2 = u'^3 - u'$ (vérification immédiate). En changeant éventuellement le signe de v' , on voit qu'on obtient un point de la courbe qu'on note encore $(u,v) = (a/b, c/d)$ avec $a, b, c, d \geq 1$. Continuons le raisonnement à partir de l'existence de ce dernier point.

2) On a $(c/d)^2 = (a/b)^3 - a/b$, donc $b^3 c^2 = ad^2(a+b)(a-b)$. On en déduit $b^3 = d^2$, d'où $b = B^2$ avec $B \geq 1$ et $c^2 = a(a+b)(a-b)$.

3) Soit p un diviseur premier de a . Comme a et b sont premiers entre eux, p ne peut diviser les deux autres facteurs. Comme p divise c^2 , on en déduit que $p^2 \mid a$, donc $a = A^2$ avec $A \geq 1$.

4) On a alors $A^4 - B^4 = a^2 - b^2 = c^2/a = c^2/A^2$, ce qui montre que $A^4 - B^4 = C^2$ avec $C = c/A \geq 1$ entier.

La seconde partie de la démonstration commence. Notre supposition initiale entraîne qu'il existe des nombres entiers $A, B, C \geq 1$ premiers entre eux tels que $A^4 - B^4 = C^2$ (c'est très proche de la situation rencontrée dans le cas $n = 4$ du théorème de Fermat). On va montrer, en utilisant la méthode de descente infinie de Fermat que cette dernière égalité est impossible.

5) Supposons B pair. Alors B^4 est divisible par 16. On a :

$$B^4 = A^4 - C^2 = (A^2 - C)(A^2 + C).$$

Le pgcd des deux termes du produit de droite doit diviser leur somme $2A^2$ et leur différence $2C$; il est donc égal à 2. On a donc $B^4 = 2p \times 8q$ avec p et q premiers entre eux. Donc pq est une puissance quatrième et il en est de même de p et q : $p = r^4$, $q = s^4$.

On a donc : $A^2 + C = 2r^4$ et $A^2 - C = 8s^4$ ou $A^2 + C = 8s^4$ et $A^2 - C = 2r^4$ avec r, s premiers entre eux.

Dans les deux cas, on a $B = 2rs$, $A^2 = r^4 + 4s^4$, soit $(A - r^2)(A + r^2) = 4s^4$.

On voit alors que $A - r^2$ et $A + r^2$, ayant 2 pour pgcd, sont des doubles de puissances quatrièmes d'entiers u et v , d'où : $2r^2 = (A + r^2) - (A - r^2) = 2v^4 - 2u^4$, soit $v^4 - u^4 = r^2$. On retrouve la relation entre A, B, C avec $v < 2rs = B < A$ et la conclusion en suivant la méthode de descente infinie.

6) Supposons B impair. On a $B^2 \equiv 1 \pmod{4}$; on sait que $A^2 \equiv 0 \pmod{4}$ ou $A^2 \equiv 1 \pmod{4}$. Dans le premier cas, on a donc $C^2 = A^4 - B^4 \equiv -1 \pmod{4}$, ce qui est impossible. Donc $A^2 \equiv 1 \pmod{4}$, $C^2 \equiv 0 \pmod{4}$ et C est pair.

On a $A^4 - B^4 = (A^2 + B^2)(A^2 - B^2)$ multiple de 4 et, comme ci-dessus, on en déduit que $A^2 + B^2 = 2p^2$, $A^2 - B^2 = 2q^2$ et $C = 2pq$, avec p, q premiers entre eux. On a $p^4 - q^4 = (AB)^2$ et on peut conclure par descente infinie puisque $p < A$.

Remarque. Dans d'autres cas, la méthode de descente infinie ne donne pas de contradiction, mais conduit à au moins un point de la courbe. En reprenant les calculs dans l'autre sens, on voit que ce point est un générateur d'un sous groupe isomorphe à $\mathbb{Z}$ du groupe de la courbe ; cette courbe est donc de rang au moins 1. Cette démarche est utile dans la pratique : pour déterminer le rang d'une courbe elliptique, on tâtonne un peu en utilisant la méthode de Fermat ; mais cela ne donne bien souvent qu'une idée du rang. On sent bien après ces indications que le rang n'est vraiment pas quelque chose de facile à calculer.

Cependant, tout n'est pas perdu et on peut aborder l'étude du rang d'une autre façon. Nous retrouvons Manjul Bhargava.

Résultats de Manjul Bhargava et Arul Shankar sur le rang des courbes elliptiques. À défaut de pouvoir calculer aisément le rang d'une courbe elliptique à partir de son équation, une autre idée s'est révélée intéressante : calculer la moyenne du nombre de courbes elliptiques de rang donné. Il faut préciser. Pour que cette question ait un sens, on définit ce qu'on appelle hauteur d'une courbe elliptique E : c'est un nombre $H(E)$ qui est fonction de la taille des coefficients de l'équation de la courbe ; pour Bhargava, si l'équation de la courbe s'écrit $y^2 = x^3 + ax + b$ (on suppose qu'il n'existe pas de nombre premier p tel que p^4 divise b et p^6 divise a pour choisir une équation particulière de la courbe), la hauteur est $H(E) = 4 |a|^3 + 27b^2$.

Notons $R(E)$ le rang de la courbe elliptique E . On définit la proportion de courbes elliptiques de rang r par

$$P(r) = \lim_{N \rightarrow +\infty} \frac{\#\{E, H(E) \leq N \text{ et } R(E) = r\}}{\#\{E, H(E) \leq N\}},$$

la notation $\#$ indiquant le cardinal d'un ensemble.

Les premiers calculs des années 1980 ont conduit à conjecturer que $P(0) = P(1) = 1/2$ et que les $P(r)$ pour $r \geq 2$ étaient nuls, autrement dit qu'il y avait autant de courbes elliptiques de rang 0 que de rang 1 et que les courbes de rang supérieur formaient un ensemble négligeable ; on a conjecturé que le rang moyen r_m des courbes elliptiques est $1/2$.

Conjecture de Birch et de Swinnerton-Dyer (notée BSD). Mais la difficulté de cette question était redoutable. Les premiers à obtenir des résultats durent faire appel à des conjectures toujours non démontrées aujourd'hui, les conjectures de Riemann

(voir 24.5) et de Birch et Swinnerton-Dyer. Ainsi, on montra d'abord que $r_m \leq 2,3$, puis que $r_m \leq 2$, puis que $r_m \leq 25/14 \approx 1,79$.

En 2010, le résultat de Manjul Bhargava et Arul Shankar fit sensation en montrant, sans s'appuyer sur aucune conjecture, que r_m était fini et que $r_m \leq 1,5$ (résultat amélioré l'année suivante : $r_m < 1$).

La conjecture BSD (le cinquième problème du millénaire, doté d'un prix de un million de dollars par la fondation Clay) est une conjecture sur les courbes elliptiques et des fonctions de variable complexe associées à ces courbes, appelées fonction L . Il y a différents types de fonctions L , comme les fonctions définies à partir de caractères dont je parle plus loin, ce qui donne beaucoup de travail et de questions aux mathématicien(ne)s. La conjecture BSD relie le rang de la courbe elliptique et le comportement au voisinage de 1 de la fonction L associée : le rang serait égal à l'ordre du zéro de la fonction L en 1. Bryan Birch (né en 1931) et Peter Swinnerton-Dyer (1927-2018, un très grand joueur de bridge) ont formulé cette conjecture en 1963 à partir de quelques calculs menés sur l'EDSAC à Cambridge en Angleterre, un des tout premiers ordinateurs construit en 1949, utilisant des tubes à vides.

Manjul Bhargava a continué à travailler et à améliorer ses résultats. Le 18 juillet 2014, avec Christopher Skinner (né en 1972) et Wei Zhang (né en 1981), il put affirmer que :

- 66,48% des courbes elliptiques satisfaisaient la conjecture BSD ;
- $P(r = 0) \geq 16,5\%$;
- $P(r = 1) \geq 20,68\%$;
- $r_m \geq 0,2068$.

L'impact du travail de Bhargava est véritablement énorme en théorie des nombres, écrit Étienne Ghys (né en 1954), actuel secrétaire perpétuel de l'Académie des sciences pour la division mathématique.

24.3 LA CONJECTURE *abc*

Je voudrais maintenant présenter une autre conjecture, formulée indépendamment en 1985 par David Masser (né en 1948) et Joseph Oesterlé (né en 1954). Certains la considèrent comme le plus important problème actuel de la théorie des nombres.

Nous avons présenté, dans l'exercice 20.10, le théorème de Mason-Stothers, une relation nouvelle découverte en 1981-1984 dans l'anneau des polynômes et qui permet de démontrer le théorème de Fermat dans ce cadre. Il s'agissait de polynômes A, B, C premiers entre eux tels que $A + B + C = 0$. La relation découverte par Mason et Stothers mêle les additions et les multiplications de polynômes. La conjecture *abc* est une conjecture où l'addition et la multiplication des entiers sont aussi mêlées et dit quelque chose sur les facteurs premiers de a et b et de leur somme.

Intéressons-nous donc aux triplets (a, b, c) d'entiers naturels premiers entre eux tels que

$$a + b = c.$$

Introduisons une définition : le radical d'un nombre n , que nous noterons $\text{rad}(n)$, est le produit des différents facteurs premiers de n .

Exemples. 1) On a $\text{rad}(xyp^r) = \text{rad}(xyp)$ pour tout nombre premier p . Par exemple, $\text{rad}(120) = \text{rad}(60) = \text{rad}(30) = 30$.

2) Comparons l'entier c et le radical de abc .

$a + b = c$	$\text{rad}(abc)$
$5 + 7 = 12$	210
$11 + 13 = 24$	858
$16 + 9 = 25$	30

Quelques essais nous montrent qu'en général le radical du produit abc est nettement plus grand que les nombres a, b et leur somme c .

3) Il est clair que la condition a, b, c premiers entre eux deux à deux est nécessaire : si on prend $a = b = p^r$, on a $c = 2p^r$ avec $p > 3$ premier et $r > 1$ qui est plus grand que $\text{rad}(abc) = 2p$.

4) En cherchant, on découvre des cas où l'inégalité n'est pas vérifiée ; pour les nombres a, b plus petits que 50, on trouve les cas suivants.

$a + b = c$	$\text{rad}(abc)$
$1 + 8 = 9$	6
$5 + 27 = 32$	30
$1 + 48 = 49$	42
$32 + 49 = 81$	42

5) Comme il semble que de tels cas soient rares, on pourrait penser qu'il n'en existe qu'un nombre fini et formuler la conjecture : il existe une constante K telle que, pour tout triplet (a, b, c) tel que $a + b = c$, on ait :

$$c \leq K \text{rad}(abc),$$

6) L'exemple suivant montre que la conjecture précédente est fausse. On part de $(a, b, c) = (1, 8, 9)$; on a $\text{rad}(abc) = 6 < 9$. On généralise l'égalité précédente en

prenant $a = 1, b = 3^{2^n} - 1, c = 3^{2^n}$ avec $n \geq 1$. On voit par récurrence que 2^{n+2} divise b (c'est vrai pour $n = 1$ et si c'est vrai pour un entier $n \geq 1$, c'est vrai pour l'entier $n + 1$, car $3^{2^{n+1}} - 1 = (3^{2^n} - 1)(3^{2^n} + 1)$ où le second facteur est pair). On a donc :

$$\text{rad}(abc) = \text{rad}[(b/2^{n+2}) \times 2^{n+2} \times 3^{2^n}] \leq (b/2^{n+2}) \times 2 \times 3 < 3c/2^{n+1},$$

donc $\frac{c}{\text{rad}(abc)} \geq 2^{n+1}/3$. Le quotient $\frac{c}{\text{rad}(abc)}$ peut donc prendre des valeurs plus grandes que toute constante K fixée ; notre conjecture avec $\varepsilon = 1$ est mise en défaut.

Énoncé de la conjecture. Oesterlé et Masser ont donc proposé la conjecture suivante :

pour tout $\varepsilon > 0$, il existe une constante $K(\varepsilon)$ telle que, pour tout triplet (a, b, c) d'entiers premiers entre eux tels que $a + b = c$, on a :

$$c \leq K(\varepsilon) \text{rad}(abc)^{1+\varepsilon}.$$

On voit que $K(\varepsilon)$ doit être plus grand quand ε est plus petit.

Une autre façon d'énoncer la conjecture est la suivante : pour tout $\varepsilon > 0$, le nombre de triplets (a, b, c) d'entiers premiers entre eux tels que $a + b = c$ et :

$$c > \text{rad}(abc)^{1+\varepsilon}$$

est fini.

En fait, on ne connaît pas de contre-exemple à l'inégalité avec $K = 1$ et $\varepsilon = 1$ et on conjecture que, pour tout triplet (a, b, c) tel que $a + b = c$, on a :

$$c \leq \text{rad}(abc)^2.$$

La conjecture abc est très puissante. Elle permet d'obtenir très simplement des résultats dont les démonstrations ont été considérées comme des tours de force (voir aussi [LAN], pages 204-208).

Donnons juste l'exemple du théorème de Fermat pour les entiers, le grand exploit de Wiles (voir 19.14).

Théorème de Fermat. Rappelons que le théorème de Fermat dit qu'il n'existe pas, pour $n \geq 3$ d'entiers x, y, z positifs et non nuls tels que :

$$x^n + y^n = z^n.$$

Supposons la conjecture abc démontrée pour $\varepsilon = 1$. Il existe donc une constante K telle que pour tout triplet (a, b, c) de nombres premiers entre eux tels que $a + b = c$, on a $c \leq K \text{rad}(abc)^2$. En prenant $a = x^n, b = y^n$ et $c = z^n$, on a donc

$$z^n < K \text{rad}((xyz)^n)^2 = K \text{rad}(xyz)^2 \leq K(xyz)^2 \leq Kz^6.$$

Cela implique que $z^{n-6} \leq K$. Si on peut prendre $K = 1$, on obtient que le théorème de Fermat est vrai pour $N \geq 6$. Il ne reste plus qu'à démontrer le théorème dans les cas $n = 3, 4, 5$. Comme ces démonstrations sont connues (voir 19.14 pour le cas $n = 3$ et l'exercice 12.11 pour le cas $n = 4$), on peut en conclure que le théorème de Fermat est vrai pour tout $n \geq 3$. C'est une démonstration, complètement différente de l'approche de Wiles !!!

24.4 LES FONCTIONS L DE DIRICHLET

Le mystère de la répartition des nombres premiers. Euler, dans l'introduction de son article : *Découverte d'une loi tout extraordinaire des nombres, par rapport à la somme de leurs diviseurs*, publié à Berlin en 1751 (mais lu en 1747 à l'Académie) avait bien exposé le problème :

Les mathématiciens ont tâché jusqu'ici en vain à découvrir quelque ordre dans la suite des nombres premiers et on a lieu de croire que c'est un mystère auquel l'esprit humain ne saurait jamais pénétrer. Pour s'en convaincre, on n'a qu'à jeter les yeux sur les tables des nombres premiers que quelques-uns se sont donné la peine de continuer au-delà de cent mille et on s'apercevra d'abord qu'il n'y règne aucun ordre ni règle.

Donnons des exemples de ces irrégularités.

1) Prenons un entier N . La suite des nombres $N! + k$ avec $2 \leq k \leq N$ est une suite de nombres sans facteurs premiers puisque $N! + k$ est divisible par k . Ainsi, on connaît des sous-ensembles d'entiers consécutifs sans nombre premier de longueur aussi grande qu'on veut.

2) Inversement, on peut avoir des nombres très rapprochés comme les nombres premiers jumeaux, dont la différence est 2, comme 29 et 31. Hilbert demandait de montrer qu'il existait une infinité de paires de tels nombres, mais Viggo Brun (1885-1978) montra en 1919 que la série de leurs inverses était convergente (un résultat qui serait trivial s'il n'y avait qu'un nombre fini de nombres premiers jumeaux !). D'énormes progrès sur ce problème ont été réalisés ces dernières années. Notons A_s l'ensemble des paires de nombres premiers différents d'un entier s , pair bien sûr. On a montré qu'il existe un nombre N tel que, si on prend la famille des ensembles A_s pour $s \leq N$, l'un au moins de ces ensembles est infini. En 2009, Ytang Zhang (né en 1955 et inconnu jusque là), avait fait sensation en annonçant ce résultat avec $N = 70\,000\,000$. La valeur de N a été depuis considérablement abaissée par paliers ; on peut même prendre $N = 6$, en admettant une certaine conjecture, d'après des travaux de James Maynard (né en 1987) et de Tao.

3) Une petite bizarrerie : Tchebycheff avait remarqué en 1853 un *biais* : les nombres premiers égaux à 3 modulo 4 sont, en général, plus nombreux que les nombres premiers égaux à 1 modulo 4 dans l'ensemble des nombres premiers de 1 à N ; on a constaté depuis que ce n'est pas toujours le cas, comme quand N est entre 616 877 et 617 011.

Comment arriver à dire quelque chose sur la répartition des nombres premiers ?
Nous allons expliquer que l'étude de la fonction $\zeta : \mathbb{C} \setminus \{1\} \rightarrow \mathbb{C}$ définie par

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}$$

et des zéros dits non triviaux où elle s'annule est au cœur des difficultés.

Mais avant de parler de la fonction ζ , commençons par évoquer quelques résultats plus anciens sur cette fonction et les travaux de Dirichlet.

Les résultats d'Euler. Euler lisait ses travaux devant l'Académie où il se trouvait, Saint-Petersbourg ou Berlin suivant l'époque, et ses textes étaient publiés quelques années plus tard ; ce sont ces deux dates que j'indique avec le numéro Exxx du texte dans l'index établi par Gustaf Eneström (1852-1923).

Le problème posé par Pietro Mengoli (1626/27-1686) en 1644 est de calculer $\zeta(2) = \sum_{n \geq 1} 1/n^2$, la somme des inverses des carrés des entiers. Personne ne réussit avant Euler (E41, 1735-1740)), plus de 90 ans plus tard. Il réussit à établir, avec de jolies astuces de calcul, que :

$$\sum_{n \geq 1} \frac{1}{n^2} = \pi^2 / 6.$$

Le résultat était une surprise : que venait faire le nombre π (qu'Euler note p) ici ?

Euler trouve en même temps les sommes des inverses des puissances paires d'entiers, $\zeta(2s)$ pour s entier, $s = 2, 4, 6, \dots$: $\zeta(4) = \pi^4/90$, $\zeta(6) = \pi^6/945$, etc., toujours avec l'apparition de π .

C'est dans E72 (1737-1744) qu'Euler donne sa magnifique formule (Th. 8) que nous écrivons :

$$\sum_{n \geq 1} \frac{1}{n^k} = \prod_{p \in \mathcal{P}} \frac{1}{1 - 1/p^k}.$$

Theorema 8.

Si ex serie numerorum primorum sequens formetur expressio

$$\frac{2^n \cdot 3^n \cdot 5^n \cdot 7^n \cdot 11^n \cdot \text{etc.}}{(2^n - 1)(3^n - 1)(5^n - 1)(7^n - 1)(11^n - 1) \text{ etc.}}$$

erit eius valor aequalis summae huius seriei

$$1 + \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \frac{1}{5^n} + \frac{1}{6^n} + \frac{1}{7^n} + \text{etc.}$$

Cette formule est fondamentale pour tous les développements futurs. Euler exprime la somme du membre de gauche comme un produit infini. Il la donne pour les entiers $n \geq 1$ entier. Il commence par $n = 1$ (Th. 7), où le membre de gauche est une série divergente, résultat déjà connu de Nicole Oresme (vers 1315-1382).

On peut s'étonner que les nombres premiers, absents du membre de gauche apparaissent dans le membre de droite ; le lien entre les deux membres est donné par la factorisation unique des entiers en produit de nombres premiers (nous avons rencontré cette idée dans l'exercice 12.9). Pour le voir, développons en séries entières les fractions $1/(1-1/p^k)$ en utilisant le développement $1/(1-x) = 1 + x + x^2 + x^3 + \dots$ valable pour $|x| < 1$. On a, pour tout $p \in \mathcal{P}$ (où $\mathcal{P}$ note l'ensemble des nombres premiers) :

$$1/(1-1/p^k) = 1 + 1/p^k + 1/p^{2k} + 1/p^{3k} + \dots$$

Si n est un entier dont le développement en facteurs premiers est $p_1^{t_1} \dots p_r^{t_r}$, on voit apparaître dans le produit du second membre le terme $1/n^k$ comme produit des $1/(p_i^{n_i})^k$ pris dans les développements des $1/(1-1/p_i^k)$.

La démonstration d'Euler dans E72 se présente différemment. Euler pose $x = 1 + 1/2^n + 1/3^n + 1/4^n + 1/5^n + 1/6^n + \dots$, calcule $1/2^n x = 1/2^n + 1/4^n + 1/6^n + 1/8^n + \dots$ (Euler indique un terme $1/3^n$ qui doit être supprimé), en déduit $(2^n - 1)/2^n x = 1 + 1/3^n + 1/5^n + 1/7^n + 1/9^n + \dots$, où n'apparaît plus de nombres pairs, puis multiplie par $(3^n - 1)/3^n$ pour faire disparaître les multiples de 3 dans la somme ; en continuant ainsi, il obtient 1 et conclut. Euler montre aussi que la série des inverses des nombres premiers est divergente (Th. 19), encore un moyen de montrer que la suite des nombres premiers est infinie.

Le résultat d'Euler de E352 (1749-1768), écrit en français : *Remarques sur un beau rapport entre les séries de puissances tant directes que réciproques*, est tout aussi remarquable et annonce les formules ultérieures sur la fonction ζ . Euler ajoute une note sur les séries $\zeta(k)$ avec k impair : *toutes mes recherches pour en trouver les sommes ont été inutiles jusqu'ici*. Le problème est effectivement difficile. Roger Apéry (1916-1994) a surpris tout le monde en montrant que $\zeta(3)$ était irrationnel en 1977 ; on n'en sait pas beaucoup plus aujourd'hui.

Théorème de la suite arithmétique. Les idées d'Euler et de Gauss sont à la base des travaux de Dirichlet publiés en 1837 (ils sont traduits en français dès 1839 par Olry Terquem (1782-1862) dans le Journal de Joseph Liouville (1809-1882), tome 4, page 393). Dirichlet écrit : *La méthode que j'emploie me paraît surtout mériter quelque attention par la liaison qu'elle établit entre l'Analyse infinitésimale et l'Arithmétique transcendante...*

Dirichlet fait donc un lien qui va s'avérer essentiel entre ce que nous appelons maintenant l'Analyse et la Théorie des nombres et Dirichlet utilise des séries de fonctions de la variable complexe là où on ne les attendait pas. Il montre le théorème de la suite arithmétique.

Théorème. Soient $a, m \geq 1$ des entiers premiers entre eux.

Il existe une infinité de nombres premiers de la forme $p = a \bmod m$; autrement dit, il existe une infinité de nombres premiers dans la suite arithmétique $(km + a)_{k \geq 0}$.

Par exemple, il existe une infinité de nombres premiers de la forme $(9k + 4)$.

Legendre avait eu besoin de ce résultat pour montrer la loi de réciprocité quadratique. Dirichlet explique que la démonstration de Legendre comporte une faille et qu'il a dû suivre un autre chemin que lui.

Je vais décrire la démarche de Dirichlet sur le cas des suites $(4k + 1)$ et $(4k + 3)$. Mais auparavant, il faut introduire la théorie des caractères ; Dirichlet n'avait pas le langage de la théorie des groupes pour l'expliquer.

Caractères d'un groupe. Si G est un groupe abélien fini noté multiplicativement, on appelle caractère de G tout homomorphisme $\chi : G \rightarrow \mathbb{C}^*$ (χ est la lettre grecque *khi* et $\mathbb{C}^*$ est le groupe multiplicatif des complexes non nuls). Les valeurs d'un caractère sont des racines de l'unité puisque les éléments de G sont d'ordre fini et leurs images dans $\mathbb{C}^*$ le sont aussi.

Les caractères forment un groupe pour la multiplication, le caractère χ_0 défini par $\chi_0(x) = 1$ pour tout $x \in G$ en étant l'élément neutre.

On va utiliser les caractères du groupe des unités $U(m)$ de l'anneau $\mathbb{Z}/m\mathbb{Z}$ (l'étude générale de $U(m)$ a été faite dans l'exercice 21.5).

Par exemple, on peut définir les caractères du groupe $U(7)$, isomorphe au groupe cyclique $\mathbb{Z}/6\mathbb{Z}$ à six caractères, par l'image du générateur 3 de $U(7)$. Ces caractères prennent pour valeurs des racines sixièmes de l'unité. Les caractères χ_1 et χ_2 , définis par $\chi_1(3) = w = e^{i\pi/3}$, $\chi_2(3) = w^2 = e^{2i\pi/3}$ prennent les valeurs suivantes.

x	$1 = 3^0$	$3 = 3^1$	$2 = 3^2$	$-1 = 3^3$	$4 = 3^4$	$5 = 3^5$
$\chi_1(x)$	1	w	w^2	$w^3 = -1$	$w^4 = -w$	$w^5 = -w^2$
$\chi_2(x)$	1	w^2	w^4	$w^6 = 1$	w^2	w^4

Un caractère $\chi : U(m) \rightarrow \mathbb{C}^*$ définit une application de l'ensemble des entiers premiers à m dans $\mathbb{C}^*$. On prolonge cette fonction à une fonction $\mathbb{Z} \rightarrow \mathbb{C}$ en lui donnant la valeur 0 pour tout entier non premier avec m . On appelle caractère de Dirichlet modulo m ou D-caractère cette fonction. Le D-caractère χ_0 modulo m est défini par $\chi_0(k) = 1$ si k est premier avec m et 0 sinon ; attention, ce caractère dépend de m .

Un D-caractère χ est une fonction complètement multiplicative, c'est-à-dire qu'on a, pour tous entiers a, b , $\chi(ab) = \chi(a)\chi(b)$.

Esquisse de démonstration du théorème de la suite arithmétique pour les suites $1+4k$ et $3+4k$. On a $U(2) = (1, 3)$ et deux D-caractères modulo 4 : le caractère χ_0 et le caractère χ définis par $\chi_0(2k) = 0$, $\chi_0(2k+1) = 1$ et par $\chi(2k) = 0$, $\chi(4k+1) = 1$ et $\chi(4k+3) = -1$.

À ces caractères, Dirichlet associe des séries de variable complexe, qu'il note L , définies par

$$L(s, \chi_0) = \sum_{n \geq 1} \frac{\chi_0(n)}{n^s} = \sum_{k \geq 1} \frac{1}{(2k+1)^s} = 1 + 1/3^s + 1/5^s + 1/7^s + 1/9^s \dots;$$

$$L(s, \chi) = \sum_{n \geq 1} \frac{\chi(n)}{n^s} = \sum_{k \geq 1} \frac{(-1)^k}{(2k+1)^s} = 1 - 1/3^s + 1/5^s - 1/7^s + 1/9^s \dots$$

avec $s = \sigma + i\tau$.

Ces séries sont absolument convergentes si $\sigma > 1$ puisque $|1/n^s| \leq 1/n^\sigma$.

Sur le modèle de la formule d'Euler, Dirichlet montre que :

$$L(s, \chi_0) = \prod_{p \in \mathcal{P}, p \geq 2} \left(\frac{1}{1 - \chi_0(p)/p^s} \right) = \prod_{p \in \mathcal{P}, p \geq 2} \left(\frac{1}{1 - 1/p^s} \right).$$

$$L(s, \chi) = \prod_{p \in \mathcal{P}, p \geq 2} \left(\frac{1}{1 - \chi(p)/p^s} \right).$$

La fonction de variable complexe $z \mapsto e^z$ prend les mêmes valeurs pour des complexes différents de $2ik\pi$. Pour tourner cette difficulté, sa fonction réciproque peut être définie sur $\mathbb{C} \setminus]-\infty, 0]$; on la note $\log$ dans ce qui suit ; quand x est réel, avec $x > 0$, $\log x$ est bien le logarithme népérien de x . Pour $|z| < 1$, on a $\log(1/(1-z)) = \sum_{n \geq 0} z^n / n$.

Dans ce qui suit, on note B toute fonction de la variable complexe bornée au voisinage de $s = 1$.

En développant les logarithmes des produits donnant $L(s, \chi_0)$ et $L(s, \chi)$, on montre que :

$$\log L(s, \chi_0) = \sum_{p \in \mathcal{P}, p \geq 2} \frac{1}{p^s} + B(s),$$

$$\log L(s, \chi) = \sum_{p \in \mathcal{P}, p \geq 2} \frac{\chi(p)}{p^s} + B(s).$$

On voit maintenant l'utilité des caractères dans les simplifications des deux calculs suivants : d'une part :

$$\log(L(s, \chi_0)) + \log(L(s, \chi)) = \sum_{p \in \mathcal{P}, p > 2} \frac{1 + \chi(p)}{p^s} + B(s) = 2 \sum_{p \in \mathcal{P}, p \equiv 1 \pmod{4}} \frac{1}{p^s} + B(s)$$

et, d'autre part :

$$\log(L(s, \chi_0)) - \log(L(s, \chi)) = \sum_{p \in \mathcal{P}, p > 2} \frac{1 - \chi(p)}{p^s} + B(s) = 2 \sum_{p \in \mathcal{P}, p \equiv 3 \pmod{4}} \frac{1}{p^s} + B(s)$$

On divise par $-\log(s-1)$ les deux égalités précédentes et on calcule les limites des quotients quand $s \rightarrow 1^+$, donc quand $-\log(s-1)$ tend vers l'infini.

En omettant les difficultés de la démonstration, on arrive aux égalités :

$$\frac{\sum_{p \in \mathcal{P}, p \equiv 1 \pmod{4}} 1/p^s}{-\log(s-1)} \xrightarrow{s \rightarrow 1} 1/2$$

$$\frac{\sum_{p \in \mathcal{P}, p \equiv 3 \pmod{4}} 1/p^s}{-\log(s-1)} \xrightarrow{s \rightarrow 1} 1/2$$

Par conséquent, l'ensemble des nombres premiers p tels que $p \equiv 1 \pmod{4}$ et l'ensemble des nombres premiers p tels que $p \equiv 3 \pmod{4}$ sont infinis et ils ont la même densité de Dirichlet, cette notion étant définie pour tout sous-ensemble $A \subset \mathcal{P}$ par :

$$d(A) = \lim_{s \rightarrow 1} \frac{\sum_{p \in A} 1/p^s}{-\log(s-1)}.$$

Cas général. La démonstration est nettement plus délicate dans le cas général. On peut encore montrer que les nombres premiers de la suite des $km + a$ pour chaque a premier avec m ont la même densité de Dirichlet.

Par exemple, il y aura à *peu près* autant de nombres premiers dans les suites arithmétiques, $(9k+1), (9k+2), (9k+4), (9k+5), (9k+7), (9k+8)$.

24.5 AUTOUR DE L'HYPOTHÈSE DE RIEMANN (HR)



Bernhard Riemann en 1863
(1826-1866)

Je voudrais aussi vous parler de l'hypothèse de Riemann (que j'abrège souvent en HR), ce sera la conclusion de cette édition. Sa résolution faisait partie de l'énoncé du huitième des 23 problèmes proposés par Hilbert en 1900. Dans son huitième problème, Hilbert avait réuni plusieurs problèmes de théorie des nombres que nous allons évoquer dans ce qui suit :

- l'hypothèse de Riemann ;
- la conjecture de Golbach ;
- la conjecture des nombres premiers jumeaux ... et d'autres.

Cent ans plus tard, aucun de ces problèmes n'est encore résolu et l'hypothèse de Riemann est le premier des sept problèmes du millénaire. C'est un domaine immense, parcouru par de nombreux(ses) mathématicien(ne)s. De nouvelles idées apparaissent, les résultats s'accumulent, tant les fonctions de ce domaine entretiennent de connexions avec le reste des mathématiques.

Le génie de Riemann 1. L'Académie des sciences de Berlin accueille Riemann en 1859. En guise de remerciements, Riemann offre à l'Académie, en novembre, l'étude : *Sur le nombre de nombres premiers inférieurs à une grandeur donnée* (*Über die Anzahl der Primzahlen unter einer gegebenen Grösse*).

Je ne crois pouvoir mieux exprimer mes remerciements à l'Académie pour la distinction à laquelle elle m'a fait participer en m'admettant au nombre de ses Correspondants qu'en faisant immédiatement usage du privilège attaché à ce titre pour lui communiquer une étude sur la fréquence des nombres premiers. C'est un sujet qui, par l'intérêt que Gauss et Dirichlet lui ont voué pendant de longues années, ne me semble peut-être pas indigne de faire l'objet d'une telle Communication.

FIGURE 24.1 – Riemann, début du mémoire de 1859, traduction de Laugel, 1898.

C'est un texte génial, plein d'idées nouvelles. Tout est concentré en une dizaine de pages seulement ; on pense que Riemann avait encore d'autres idées, qu'il n'a pas tout dit de ce qu'il avait calculé et compris faute de place ; quelques notes manuscrites éparses retrouvées en témoignent. Un manuscrit présenté comme étant de la main de Riemann est, en réalité, une copie de la main d'Alfred Clebsch (1833-1872).

1) Le but de Riemann est de trouver une formule donnant le nombre exact, noté $\pi(x)$ ¹, de nombres premiers inférieurs ou égal à x (par exemple : $\pi(16,5) = 6$, $\pi(17) = 7$).

2) Riemann part des séries $\sum_{n \geq 1} 1/n^k$ définies par Euler. Il étend leur définition aux nombres complexes sur lesquels de grands progrès ont été faits les années précédentes, principalement par Cauchy (intégrale d'une fonction de variable complexe le long d'un contour, théorème des résidus). Il définit la fonction :

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s},$$

lui donnant son nom : *zêta*, avec la lettre grecque ζ . Ici, n est un entier et $s = \sigma + i\tau$ est un nombre complexe de partie réelle σ et de partie imaginaire τ ; le nombre complexe $n^s = n^\sigma \times n^{i\tau}$ est de module n^σ et d'argument $\tau \log n$.

1. C'est la notation usuelle, la même que celle qui désigne le nombre $\pi = 3,141592\dots$. Riemann utilise la lettre F .

3) Riemann étend la formule d'Euler aux exposants complexes :

$$\sum_{n \geq 1} \frac{1}{n^s} = \prod_{p \in \mathcal{P}} \frac{1}{1 - 1/p^s}.$$

4) La somme $\sum_{n \geq 1} \frac{1}{n^s}$ n'est définie que pour les nombres complexes $s = \sigma + i\tau$ tels que $|\sigma| > 1$ puisque la série ne converge pas sinon. Quand $s \rightarrow 1$ avec $\sigma > 1$, elle tend vers l'infini.

Riemann montre comment prolonger la fonction ζ en une fonction holomorphe (dérivable au sens complexe) notée encore ζ définie sur $\mathbb{C} \setminus \{1\}$.

Ne soyez pas surpris de tels prolongements ; quand on considère la série :

$$f(z) = \sum_{n \geq 0} z^n,$$

le membre de droite est la somme d'une série et n'est défini que dans le domaine de convergence de cette série : $|z| < 1$. Mais comme la somme de cette série pour $|z| < 1$ est la fonction $z \mapsto f(z) = 1/(1-z)$, la fonction f peut être prolongée (de manière unique) à $\mathbb{C} \setminus \{1\}$ par $f(z) = 1/(1-z)$.

5) Notons Γ la fonction de variable complexe vérifiant $\Gamma(z+1) = z\Gamma(z)$. Cette fonction s'appelle la fonction gamma ; on peut la définir par $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$ pour $\operatorname{Re}(z) > 0$ et la prolonger au plan complexe, sauf aux entiers négatifs ou nuls. Elle étend la notion de factorielle pour les entiers ; pour tout entier $n \geq 2$, on a $\Gamma(n) = (n-1)!$. Une des belles idées de Riemann est de voir que la fonction ξ définie sur $\mathbb{C} \setminus \{1\}$ par :

$$\xi(s) = \frac{s(s-1)}{2} \pi^{-s/2} \Gamma(s/2) \zeta(s)$$

vérifie l'équation fonctionnelle :

$$\xi(s) = \xi(1-s).$$

L'hypothèse de Riemann (HR). Riemann montre que la fonction ζ s'annule dans le demi-plan complémentaire de $\{z, z \geq 0\}$ en les seuls points situés sur l'axe réel aux abscisses $-2n$ avec $n > 0$. Ces points sont appelés zéros triviaux de ζ .

On appelle bande critique la bande $]0, 1[\times \mathbb{R}$.

Les zéros de ζ situés dans la bande critique sont appelés zéros non triviaux de ζ . On les note avec la lettre ρ .

L'hypothèse de Riemann (que je note HR) est que les zéros non triviaux de ζ sont de partie réelle $1/2$.

Dans son article, Riemann s'exprime un peu différemment, indiquant que toutes les racines de la fonction $\tau \mapsto \zeta(1/2 + i\tau)$ sont réelles :

... et il est très probable que toutes les racines soient réelles.

Il serait à désirer, sans doute, que l'on eût une démonstration rigoureuse de cette proposition ; néanmoins, j'ai laissé cette recherche de côté pour le moment après quelques rapides essais infructueux, car elle paraissait superflue pour le but immédiat de mon étude.

Pour une telle hypothèse, donnons le texte allemand :

... es ist sehr wahrscheinlich, dass alle Wurzeln reell sind.

Hiervon wäre allerdings ein strenger Beweis zu wünschen ; ich habe indess die Aufsuchung desselben nach einigen flüchtigen vergeblichen Versuchen vorläufig bei Seite gelassen, da er für den nächsten Zweck meiner Untersuchung entbehrlich schien².

Nous allons essayer d'expliquer cette hypothèse et de dire pourquoi elle est si importante. Je renvoie, en particulier, à l'article de Michel Balazard : Un siècle et demi de recherches sur l'hypothèse de Riemann, publié en octobre 2010 dans la *Gazette des mathématiciens*. Mais bien d'autres textes abordent le même sujet.

Le génie de Riemann 2. 1) Riemann introduit une fonction que nous noterons J (Riemann utilise la lettre f , trop banale aujourd'hui), définie par :

$$J(x) = \sum_{k \geq 1} \frac{1}{k} \pi(x^{1/k}).$$

Cette somme est une somme finie, car $x^{1/k}$ tend vers 1 quand $k \rightarrow +\infty$, donc $\pi(x^{1/k})$ devient nul. Par exemple, pour $x = 100$, on a :

k	$x^{1/k}$	$\pi(x^{1/k})$
1	100	25
2	10	4
3	4,64...	2
4	3,16...	2
5	2,51	1
6	2,15...	1
7	1,93...	0

2) Une fonction qu'on rencontre souvent en théorie des nombres est la fonction d'August Möbius (1790-1868) notée $\mu : \mathbb{N} \rightarrow \mathbb{Z}$ et définie par :

$$\mu(1) = 1 ;$$

$$\mu(n) = 0 \text{ si } n \text{ possède un facteur carré ;}$$

2. Wikipedia donne une version française médiocre, probablement retraduite de l'anglais ou de l'allemand, et non la version française de Léonce Laugel (1859-1936).

$\mu(n) = (-1)^r$ si n est un produit de r nombres premiers distincts.

Une formule, appelée formule d'inversion de Möbius, permet de retrouver la fonction π à partir de la fonction J par :

$$\pi(x) = \sum_{k \geq 1} \frac{\mu(k)}{k} J(x^{1/k}).$$

Écrivons les premiers termes de la somme :

$$\pi(x) = J(x) - \frac{1}{2} J(x^{1/2}) - \frac{1}{3} J(x^{1/3}) - \frac{1}{5} J(x^{1/5}) + \frac{1}{6} J(x^{1/6}) + \dots$$

compte-tenu de $\mu(2) = \mu(3) = \mu(5) = -1$, $\mu(4) = 0$ car 4 est un carré, $\mu(6) = 1$ car 6 est le produit de 2 et de 3.

3) La fonction $: [0, \rightarrow +\infty[\rightarrow \mathbb{R}$, appelée logarithme intégral, est définie par :

$$\text{li}(x) = \int_0^x \frac{dt}{\log t}.$$

Gauss avait déjà remarqué la proximité de $\text{li}(x)$ et de $\pi(x)$. Une curiosité de cette fonction est que tous les calculs numériques montrent que $\pi(x) < \text{li}(x)$ quels que soient les nombres qu'on considère (par exemple, pour $x = 10^{14}$, $\text{li}(x) - \pi(x) = 314\,890$). En fait, John Littlewood (1885-1977) a montré, en 1914, que la différence $\text{li}(x) - \pi(x)$ change de signe une infinité de fois, mais pour des valeurs extrêmement grandes, ce que Riemann n'avait pas soupçonné. En 2000, l'estimation du premier de ces changements donnait un nombre de l'ordre de $6,7 \times 10^{316}$. Sur cet exemple, on voit encore qu'on ne peut inférer le comportement de certains objets de la théorie des nombres des expérimentations numériques qu'on peut faire.

4) Riemann a calculé trois zéros non triviaux de la fonction ζ ; il trouva le premier de partie imaginaire 14, 1386 (légèrement trop grand). Sa méthode, excellente, a été retrouvée dans ses papiers à Göttingen par Carl Siegel (1896-1981) et publiée en 1932.

La formule de Riemann. La formule

$$\frac{\log \zeta(s)}{s} = \int_1^{+\infty} \frac{J(x)}{x^{s+1}} dx$$

permet à Riemann d'arriver à ce qui doit être encadré en grand :

$$J(x) = \text{li}(x) - \sum_{\rho} \text{li}(x^{\rho}) - \log 2 + \int_x^{+\infty} \frac{1}{t(t^2 - 1)\log t} dt.$$

3. Il y a une difficulté pour $x = 1$; on prend en fait l'intégrale $\int_0^{1-\varepsilon} + \int_{1+\varepsilon}^0$. Mais on pourrait aussi ajouter une constante convenable (1,045...) à la fonction $x \mapsto \int_2^x \frac{dt}{\log t}$.

Les deux derniers termes de cette formule ont de petites valeurs : $\log 2 \approx 0,69\dots$ et l'intégrale a une valeur inférieure à l'intégrale pour $x = 2$ qui vaut approximativement 0,14.

On calcule $\text{li}(x)$ par des méthodes numériques.

Il reste cette somme $\sum_{\rho} \text{li}(x^{\rho})$. C'est là qu'interviennent les fameux zéros ρ non triviaux de la fonction ζ , ceux de partie réelle $1/2$. La somme est calculée en additionnant suivant les valeurs croissantes de la partie imaginaire de ρ à partir de 0, les valeurs en ρ et en son conjugué $\bar{\rho} = 1 - \rho$ étant additionnées ensemble, ce qui conduit à des nombres réels.

Répetons que cette formule extraordinaire permet le calcul de valeurs exactes de π (avec pas mal de travail).

Le théorème des nombres premiers (TNP). Compter les nombres premiers jusqu'à un entier x , autrement dit déterminer $\pi(x)$, a été entrepris indépendamment par Legendre et Gauss.

Les valeurs de $\pi(x)$ conduisent Legendre à penser que, quand x devient grand, on a :

$$\pi(x) \simeq \frac{x}{A \log(x) + B},$$

avec des constantes A, B à déterminer (*Essai sur la théorie des nombres*, 1798). En 1808, Legendre précise : $A = 1$ et $B = 1,08366$. Gauss explique dans une lettre de 1849 qu'il avait fait des calculs en 1792-1793 (il avait alors 15-16 ans), comptant le nombre de nombres premiers dans des séquences de 1000 nombres (des chiliades) : *Je perçus bientôt que, malgré les fluctuations, la fréquence est, en moyenne, plus ou moins inversement proportionnelle au logarithme* ; Gauss n'est pas d'accord avec la formule proposée par Legendre. Il ajoute : *J'ai très souvent passé un quart d'heure ici ou là à compter une autre chiliade, mais j'ai finalement abandonné avant de parvenir à un million. Seriez-vous capable d'en faire autant en un quart d'heure, chronométrez-vous !*

Les études de Legendre et Gauss amènent à étudier le comportement asymptotique de la fonction π au voisinage de $+\infty$. Le résultat célèbre est le théorème démontré en 1896 par Jacques Hadamard (1865-1963) et Charles de la Vallée Poussin (1866-1962), indépendamment l'un de l'autre :

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x / \log(x)} = \lim_{x \rightarrow \infty} \frac{\pi(x)}{\text{li}(x)} = 1.$$

C'était une démonstration difficile (elle a été simplifiée depuis). Montrer que la fonction ζ n'a pas de zéro autre que 1 sur la droite $\sigma = 1$ était un des éléments importants de la preuve.

Cette démonstration du TNP était vraiment étonnante pour tout le monde : la théorie des fonctions de la variable complexe permettait de résoudre une question purement arithmétique !

Il fallut attendre 1949 pour que Paul Erdős (1913-1996) et Atle Selberg (1917-2007) parviennent à construire une démonstration purement arithmétique du TNP (voir le livre très clair de Michel Balazard : *Le théorème des nombres premiers*, publié en 2016 chez Calvage et Mounet).

Le calcul des zéros non triviaux. Riemann indique une valeur approchée du nombre de zéros non triviaux de partie imaginaire inférieure à une valeur T donnée :

$$\frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} = \frac{T}{2\pi} \log\left(\frac{T}{2\pi e}\right).$$

Le calcul des zéros (de partie imaginaire positive) a commencé, nous l'avons dit, avec Riemann. On en connaissait 15 en 1903, 25 000 en 1956, 10^{13} en 2004. Ce dernier chiffre semble énorme, mais il ne montre pas que HR est vérifiée et il suffirait de trouver un zéro, un seul ! dans la bande critique dont la partie réelle ne soit pas $1/2$ pour invalider HR.

On a cherché à restreindre la partie de la bande critique dans laquelle se situe les zéros non triviaux. Les calculateurs jouent leur rôle dans l'amélioration des résultats. On a cherché des zones rectangulaires de la forme $[1/2 + \alpha, 1] \times [1, T]$ ne contenant pas de 0, des zones limitées par une courbe comme $\sigma \geq 1 - \frac{c}{\log|\tau|}$ et $|\tau| \geq T$.

La meilleure zone connue a été donnée en 1958. Ces travaux ont des conséquences sur différentes valeurs numériques dépendant des résultats précis sur les zéros de la fonction ζ .

Un autre résultat spectaculaire, datant de 1989 : au moins 40% des zéros non triviaux de ζ sont de partie réelle $1/2$.

Les formes équivalentes. On connaît de nombreuses formes équivalentes de HR. Démontrer l'une d'entre elles démontrerait HR elle-même. En voici de curieuses.

1) Pour $x \geq 2657$, on a :

$$|\pi(x) - \text{li}(x)| \leq \frac{1}{8\pi} \sqrt{x} \log x.$$

2) Pour $N \geq 74$, le ppcm des entiers k , $2 \leq k \leq N$ est inférieur à :

$$e^{N + (1/8\pi)\sqrt{N} \log^2 N}.$$

Les conséquences en supposant HR vraie. L'énoncé HR a été généralisé (en 1884) aux prolongements des fonctions L de Dirichlet au plan complexe et à leurs zéros dans la zone critique : ils devraient avoir $1/2$ pour partie réelle. On a depuis étendu HR à différentes situations.

On disposerait, si HR était démontrée, d'un théorème puissant. Un certain nombre de résultats ont été démontrés en supposant HR vraie.

1) On sait que les groupes multiplicatifs des corps finis sont cycliques (Voir 21.1). Emil Artin a conjecturé (vers 1925) que, pour tout entier $g > 0$ non carré, il existe une infinité de nombres premiers p tels que la classe de g modulo p engendre le groupe $(\mathbb{Z}/p\mathbb{Z})^*$. Les expériences numériques le confirme. En supposant HR vraie, on en a trouvé une démonstration en 1967. Sans HR, on a seulement montré que la conjecture est vraie pour une infinité de g .

2) La conjecture de Goldbach est énoncée dans une lettre de Christian Goldbach (1690-1764) à Euler de 1742 : tout entier pair supérieur ou égal à 4 est somme de deux nombres premiers (par exemple : $818 = 691 + 127$). La conjecture faible de Goldbach est : tout entier impair supérieur ou égal à 7 est somme de trois nombres premiers. L'HR a été utilisée en 1997 pour montrer cette seconde conjecture, mais des progrès successifs ont été réalisés depuis sans supposer l'HR vraie et Harald Helfgott (né en 1977) a donné en 2013 une démonstration sans condition de la conjecture faible.

3) Le test de primalité de Miller-Rabin a été présenté en 22.11.1. Si un nombre n n'est pas premier, on aimerait connaître l'ordre de grandeur du plus petit entier a pour lequel n n'est pas fpp. En supposant HR, on a montré que $a \leq 2(\log n)^2$. Ceci prouve qu'on pourrait déterminer en temps polynomial si n est ou non premier : il suffirait de faire le test avec des entiers inférieurs à la borne précédente. Mais le test de Rabin-Miller est bien plus rapide (on possède aussi, depuis 2003, le test AKS, voir 22.11.2).

4) Si π et li sont des fonctions équivalentes en $+\infty$, j'ai donné un exemple des valeurs différentes qu'elles peuvent prendre. En supposant HR, on montre qu'il existe des nombres A et X tels que, pour $x \geq X$, l'inégalité suivante donne une borne de ces différences :

$$|\pi(x) - \text{li}(x)| \leq A\sqrt{x \log(x)}$$

(on écrit souvent : $\pi(x) = \text{li}(x) + O(\sqrt{x \log(x)})$).

5) Mais on ne sait pas encore, je crois, répondre aux questions suivantes.

a) Existe-t-il une infinité de nombres premiers de la forme $n^2 + 1$?

b) Existe-t-il toujours un nombre premier entre n^2 et $(n+1)^2$? C'est une question de Legendre. Le postulat de Bertrand (1822-1900) demandait s'il existait toujours un nombre premier entre n et $2n$. Il a été démontré par Tchebycheff en 1850 et Paul Erdős en a donné une nouvelle et très belle démonstration en 1931, en allemand ; il avait 19 ans, c'était son premier article !

c) Et on a bien d'autres questions...

En attendant la suite... De nouveaux articles paraissent chaque année pour proposer une nouvelle approche de HR et de ses généralisations, pour reprendre une ancienne approche avec des idées différentes, pour énoncer un nouveau résultat

partiel, une amélioration de majoration ; on pourrait, dit-on, utiliser des résultats anciens de Weil, la géométrie non commutative d'Alain Connes (né en 1947, médaille Fields 1982), etc.

La fécondation des mathématiques par la fonction ζ est incroyable, on découvre des connexions avec de nombreux domaines de recherche distincts, jusqu'aux probabilités et à la mécanique quantique. D'autres conjectures proches passionnent aussi les mathématicien(ne)s actuellement, comme les conjectures de Langlands de 1965 qui établiraient des liens profonds entre la théorie des nombres, la théorie des groupes, la théorie analytique des nombres.

Quand comprendra-t-on vraiment tout cela ?

Que vous restera-t-il de tout cela, dites-le moi, dites-le moi !

...ainsi que vos critiques, suggestions, en remerciant ceux qui l'ont fait pour les premières éditions.

jean-pierre.escofier@univ-rennes1.fr

BIBLIOGRAPHIE

Lire plusieurs approches différentes d'un même sujet favorise la compréhension et l'approfondissement. Que conseiller pour compléter la lecture de ce livre ? Des livres destinés aux classes préparatoires ne traitent pas tous les thèmes abordés ici. Voici quelques propositions par ordre alphabétique.

[ART] Artin Michael, *Algebra*, Prentice Hall, 1991 et rééditions.

Le livre couvre le programme d'algèbre des années d'université et l'algèbre linéaire est rapidement traitée. La suite du livre introduit à différents domaines de l'algèbre actuelle ; c'est toujours très bien expliqué, très agréable à lire, en anglais.

[BER] Berger Marcel, *Géométrie*, Cédic, Fernand Nathan, 1990.

[BER] Berger Marcel, *Géométrie vivante ou l'échelle de Jacob*, Cassini, 2009.

De grandes références de livres de géométrie ; de très nombreux résultats, très beaux, très ouverts sur les problèmes actuels ; la rédaction est rapide.

[BRA] BRAISE, Base RAISONnée d'Exercices présentée sur le site :

<http://braise.univ-rennes1.fr/braise.cgi>

La base est conçue par un groupe d'enseignants de mathématiques de l'université de Rennes ; elle comporte actuellement plus de 2 000 pages d'exercices sur les espaces vectoriels, les matrices, les polynômes, la réduction des endomorphismes ainsi que différentes parties sur l'analyse (suites, fonctions, primitives, équations différentielles) et sur les probabilités ; elle est enrichie régulièrement. On peut choisir le thème, le niveau de difficulté, le type de tâche ; pour chaque exercice sont donnés son énoncé, des éléments de cours et des aides appropriées, des éléments de solution, des idées à retenir.

[EPI] Épistémon Léonard, *algèbre vol. 1*, rédigé par Jean-Louis Ovaert et Jean-Luc Verley, Cédic, Fernand Nathan, 1981.

C'est un livre d'exercices et de problèmes, étudiant par grands thèmes l'algèbre linéaire des chapitres 1 à 10, avec des solutions et des commentaires ; les auteurs ont rédigé des notes historiques passionnantes et donné une bibliographie extrêmement riche. Vous pourrez le trouver en bibliothèque.

[ESC1] Escofier Jean-Pierre, *Théorie de Galois*, Dunod, 2^e édition, 2004 (une version en anglais de 2001 est disponible chez Springer).

Un livre sur les très belles idées de Galois, élaborées pour mieux comprendre la résolution des équations algébriques. L'exposé est fait pour être accessible au plus grand nombre.

[ESC2] Escofier Jean-Pierre, *Toute l'analyse de la licence*, Dunod, 2020.

Bibliographie

Le correspondant en analyse du livre que vous avez dans les mains, avec la même recherche d'éclairages historiques significatifs.

[ESC3] Escofier Jean-Pierre, *Petites histoires des mathématiques*, Dunod, 2018.

Le double de pages d'un livre de 2008 dont le but est de donner des aperçus historiques sur les mathématiques depuis les origines jusqu'à nos jours. Mais il y aurait tellement de choses à dire...

[GHY] Ghys Étienne, *A singular mathematical promenade*, ENS Éditions, 2017.

Étienne Ghys nous invite à visiter une trentaine de sujets mathématiques, chaque arrêt étant plein de charme et d'intérêt.

[GOD] Godement Roger : *Cours d'algèbre*, Hermann, 1963.

Le livre de Godement (1921-2016) apportait en 1963 un ton nouveau et enthousiaste ; il était ouvert sur l'air du temps, aussi bien sur les mathématiques bourbakistes que sur les liens entre les mathématiques, le monde militaire, les luttes politiques ; il contient près de 200 pages de textes d'exercices et de problèmes. Godement a beaucoup écrit sur les rapports entre les sciences et l'armement ; la fameuse phrase du mathématicien Godfrey Hardy affirmant que les mathématiques n'auraient jamais aucune application militaire est fausse.

[LAN] Lang Serge, *Algèbre*, troisième édition revisitée, Dunod, 2004.

Traduit enfin en français par Christos Grammatikas, le livre de Serge Lang est un des plus célèbres traités d'algèbre de ces dernières années et de nombreux mathématiciens du monde entier se sont formés en l'étudiant. Serge Lang avait tenu à ajouter des compléments pour la version française. Serge Lang est mort le 12 septembre 2005 à 78 ans. Il avait une énergie extraordinaire.

[PER] Perrin Daniel, *Cours d'algèbre*, Ellipses, 1996.

Rédigé pour les élèves de l'École normale supérieure, ce livre traite des chapitres 18 à 22 et va souvent plus loin.

[PIC] Pickover Clifford A., *Le beau livre des MATHS, de Pythagore aux fractales, 250 découvertes qui ont changé le monde*, Dunod, nouv. éd. 2019.

Un livre conçu sur le principe : un événement mathématique expliqué sur une page, avec une belle illustration sur l'autre page ; 250 événements et beaucoup d'informations sur les mathématiques dans leur diversité. J'ai travaillé sur les différentes éditions, particulièrement la dernière, pour laquelle j'ai rédigé une douzaine d'événements.

Vous trouverez dans les bibliothèques différents livres accessibles sur l'histoire des mathématiques. Si vous cherchez à lire en version papier, tels qu'ils ont été publiés, des livres anciens (Bernoulli, Cramer, Euler, Gauss, etc.), vous pouvez aller dans les bibliothèques universitaires ou dans les grandes bibliothèques municipales. Elles conservent des trésors. Aujourd'hui, le réseau apporte des possibilités nouvelles extraordinaires et vous aurez accès aux livres anciens sur des sites comme celui de la BNF ou d'autres ; on a peine à imaginer que cela n'existait pas il y a 30 ans...

INDEX

A

A-algèbre, 499
action, 416
 de groupe, 416
 transitive, 417
Adelman Len, 601
Agrawal Manindra, 606
Alberti Leon Battista, 598
Alembert Jean Le Rond d', 254, 383, 417, 508
algorithme
 de Berlekamp, 566
 de Cantor-Zassenhaus, 567
 de Karatsuba, 157
 de Strassen, 157
 d'Euclide, 226, 529
Al Haytham Ibn (Alhazen), 234
Al Khwarizmi Mohammed, 245
analyse des données, 347
angle de vecteurs, 566
anneau
 à division, 555
 commutatif unitaire, 219
 de fractions, 482
 euclidien, 471
 factoriel, 472
 intègre, 223
 nul, 220
 principal, 224
 quotient, 449
Apéry Roger, 691
Apollonius de Pergé, 660
application, 5
 définie par un vecteur, 114
 diagonalisable, 299
 linéaire, 110
Archimède, 533
Artin Emil, 475, 645
Aryabhata, 464

associativité, 194
automorphisme
 d'espace vectoriel, 121
 intérieur, 420, 424
axe, 81
 conjugués, 660
axiome de choix, 400
Ayoub Joseph, 483

B

Babylone, 53, 215, 236, 338
Bachet de Méziriac Claude Gaspard, 228, 614, 649
Banach Stefan, 75, 109, 177
Bartels Martin, 380, 381, 393
base, 93
 canonique de $\mathbb{R}^n$, 34
 de facteurs, 606
 directe, 289
 duale, 180
 indirecte, 289
 orthogonale, 339
 orthonormée, 339
Baudelaire Charles, 597
Baudot Émile, 581
Benze Dorothea, 380
Benzécri Jean-Paul, 347
Berlekamp Elwyn, 512, 567
Bernoulli Jean, 3, 4, 6, 109, 381, 508
Bernoulli Nicolas, 6
Bertrand Joseph, 702
Bessel Friedrich, 385, 396
Bézout Étienne, 228, 272
Bhagarva Manjul, 679-681, 685, 686
Bhaskara, 464
bidual d'un espace vectoriel, 178
bilinéaire, 278
bilinéarité, 278

Index

binary digit, 581
Binet Jacques, 284
Birch Bryan, 686
Birkhoff Garrett, 172
bit, 581
 de parité, 584
 de redondance, 583
Bolyai Janos, 386-388, 393, 395
Bonaparte Jérôme, 387
Bougainville Louis Antoine de, 383
Bourbaki Nicolas, 75, 473, 646
Brahmagupta, 464
Brunswick duc de, 379, 381, 383, 387
Brun Viggo, 689
Bühler Walter, 396
Bunyakowski Viktor, 337
Burnside William, 421
but, 3
Büttner J. G., 380

C

can, 34
Cantor Georg, 400
Cantor Moritz, 383
caractère d'un groupe, 692
Carcavi Pierre de, 565
Cardan Jérôme, 538
Carmichael Robert, 614
Carnot Lazare, 78
carré magique, 103
Cartan Élie, 646
Cartan Henri, 646
Cassels John, 660
Catalan Eugène, 277, 293
Cauchy Augustin, 274, 275, 337, 419, 420, 517, 696
Cayley Arthur, 135, 184, 272, 313
centralisateur, 420
centre d'un groupe, 207
Cervantès Miguel, 542
César Jules, 597
changement de base, 147
 pour le produit scalaire, 334
 pour une application linéaire, 148
 pour un vecteur, 147
changement de signe, 515
Chevalley Claude, 322
Cicéron, 383
classe
 à droite, 405
 à gauche, 405
 de conjugaison, 427
Clebsch Alfred, 696
clôture algébrique, 504
codage systématique, 583
code
 BCH, 582, 592
 correcteur, 581
 cyclique, 589
 de Golay, 582
 de Goppa, 582
 de Hamming, 587
 de Reed-Müller, 582
 de Reed-Solomon, 582
 linéaire, 586
 parfait, 586
coefficient dominant, 248
cofacteur, 284
Colbert Jean-Baptiste, 525
colinéaire, 33, 83
combinaison linéaire, 82
 de fonctions, 9
 de suites, 18
commutateur, 424
commutativité, 195
complexification, 623
complexifié, 624
congruence, 216
conjecture (s)
 abc, 686
 BSD (Birch-Swinnerton-Dyer), 685
 de Golbach, 695
 de Langlands, 703
 de Shimura-Taniyama-Weil, 476
conjugaison, 420, 649
conjugués, 196
Connes Alain, 703
contenu, 505
Conway John, 680
Cooper Fenimore, 394
coordonnées, 94

corps, 78, 219
 de décomposition, 504
 de rupture, 503
 des fractions, 480
 gauche, 555
 non commutatif, 555
 cosinus d'un angle de vecteurs, 331
 Coste Michel, 517
 couple, 30
 Courant Richard, 171
 courbe elliptique, 609, 682
 Cramer Gabriel, 55, 268, 271-273, 277, 280, 285
 critère d'Eisenstein, 513
 cycle, 202

D

Daemen Joan, 599
 Davout Louis Nicolas, 387
 décomposition de Dunford, 321
 décomposition unique en produit de facteurs irréductibles, 472
 Dedekind Richard, 78, 79, 219, 394
 degré
 d'une extension, 447
 d'un polynôme, 248, 519
 total, 520
 demi-tour, 355
 dénominateur, 480
 densité de Dirichlet, 694
 dérivation, 254
 Descartes René, 27-29, 245, 517, 535
 descente infinie, 236, 539
 déterminant
 de Gram, 370
 de Vandermonde, 294
 d'une matrice carrée, 274
 d'un endomorphisme, 284
 développement en fraction continue, 526, 530
 diagonalisable, 299, 303
 diagonalisation, 299
 diagonaliser, 299
 Dickson Leonard, 555
 différence finie, 112

Diffie Whitfield, 600
 dimension, 94
 Diophante, 222, 649
 Dirichlet Gustav Lejeune, 171, 333, 465, 486, 679, 690-693
 discriminant, 653
 distance
 à un sous-espace, 345
 de Hamming, 585
 minimale d'un code, 585
 prescrite, 593
 diviseur, 465
 de 0, 223
 d'une forme quadratique, 654
 divisibilité, 225, 465
 division euclidienne, 215, 248, 471
 Dreyfus Alfred, 597
 droite vectorielle, 31
 dual d'un espace vectoriel, 178
 dualité, 177
 Dubost Antonin, 646
 Dumouriez Charles-François, 379
 Dunford Nelson, 321
 Dunnington Waldo, 396

E

Eilenberg Samuel, 171, 190
 Einstein Albert, 390
 élément (s), 3
 associés, 465
 étrangers, 459
 extrémal, 467
 inverse, 194
 irréductible, 257, 467
 neutre, 194
 nilpotent, 543
 premier, 470
 premiers entre eux, 466
 réductible, 467
 simple
 dans $R(X)$, 511
 dans $C(X)$, 509
 Elkies Noam, 683
 endomorphisme
 adjoint, 356

Index

- autoadjoin, 357
 - d'espace vectoriel, 111
 - diagonalisable, 303
 - direct, 290
 - indirect, 290
 - symétrique, 357
 - Eneström Gustaf, 690
 - ensemble
 - d'arrivée, 3
 - de départ, 3
 - ordonné inductif, 470
 - quotient, 401
 - entiers de Gauss, 452
 - épimorphisme, 484, 489
 - équation
 - de Pell-Fermat, 464
 - des classes, 419, 420
 - équation caractéristique, 8, 19
 - équation différentielle, 21
 - équation différentielle linéaire, 10
 - à coefficients constants, 6
 - avec second membre, 6
 - du premier ordre, 6
 - du second ordre, 6
 - homogène, 6
 - sans second membre, 6
 - Erdős Paul, 701
 - Ernst-August, 393
 - espace
 - engendré, 9, 32, 82
 - euclidien, 333
 - isotrope, 641
 - préhilbertien réel, 332
 - propre, 302
 - quadratique, 634, 642
 - régulier, 642
 - vectorel, 31, 76, 121
 - Euclide, 27, 215, 226, 236, 393
 - Euler Leonhard, 4, 6-9, 109, 231, 234, 254, 381, 426, 464, 476, 526, 530, 561, 570, 572, 649, 652, 689-691, 693, 696, 697
 - évaluation, 112, 248, 259
 - Ewald Heinrich, 392, 393, 395
 - extension
 - de corps, 446
 - d'un corps, 447
 - extrémal, 467
- ## F
- factorisation, 606
 - famille
 - de Vandermonde, 100
 - échelonnée, 36, 97
 - finie, 31
 - génératrice, 90
 - libre, 90
 - liée, 91
 - orthogonale, 339
 - triangulaire, 36, 97
 - Farkas Wolfgang von Bolyai, 381
 - Fermat Pierre de, 27, 28, 231, 383, 464, 476, 540, 566, 572, 606, 649, 652
 - Fibonacci, 21, 22
 - fixer, 417
 - Foch Ferdinand, 598
 - Foncenex François Daviet, 383
 - fonction, 5
 - constante, 5
 - polynomiale, 246
 - polynomiale associée, 259
 - vide, 5
 - forme bilinéaire, 278, 628
 - symétrique, 628
 - forme linéaire, 178
 - forme multilinéaire, 278
 - forme polaire, 633
 - forme(s) quadratique(s), 335, 632, 633
 - anisotrope, 641
 - associée, 632
 - définie positive, 654
 - dégénérée, 641
 - équivalentes, 654
 - non dégénérée, 641
 - proprement équivalentes, 654
 - réduite, 657
 - formule
 - de Cramer, 55, 69, 267, 273
 - de Newton, 541
 - des classes, 418
 - des orbites, 418

des trois niveaux, 187
 d'inversion de Möbius, 699
 du volume, 283
 Fourier Joseph, 4, 75, 517
 fraction
 irréductible, 480
 rationnelle, 482
 rationnelle symétrique, 525
 Fréchet Maurice, 4
 Fredholm Ivar, 4
 Frobenius Georg, 287, 647

G

Gabriel Pierre, 396
 Galilée, 542
 Galois Évariste, 74, 193, 403, 555, 581
 Gardner Martin, 602
 Gauss Charles Frederick, 393, 396
 Gauss Eugen, 391, 395, 396
 Gauss-Ewald Minna, 391, 392
 Gauss Joseph, 388, 390, 395
 Gauss Carl Friedrich, 55, 135, 171, 216,
 229, 254, 274, 347, 365, 379-396,
 452, 453, 476, 505, 517, 557, 558,
 634, 691, 699, 700
 Gauss Louis, 388
 Gauss, mesure d'induction
 magnétique, 392
 Gauss Therese, 389, 392
 Gauss Gebhard, 380
 Gauss Wilhelm, 389
 G-ensemble, 416
 George II, 381
 Germain Sophie, 385, 386, 474
 Ghys Étienne, 686
 Girard Albert, 254, 542, 565
 Goethe, 394
 Golay Marcel, 582
 Goldbach Christian, 702
 Goppa Valeri, 582
 Gram Jorgen, 340
 graphe d'une fonction, 3
 Grasmann Hermann, 74, 96
 Green Ben, 683
 Grothendieck Alexander, 483, 536

groupe
 abélien 195, 196
 cyclique, 416
 commutatif, 195
 de permutations, 201
 dérivé, 424
 diédral, 421
 monogène, 415
 orthogonal, 348
 quotient, 406
 spécial linéaire, 423
 spécial orthogonal, 349
 Guillaume IV, 393
 Guillou Louis, 604

H

Hadamard Jacques, 700
 Hamilton William, 79, 646, 647, 662
 Hamming Richard, 582
 Hanke Jonathan, 681
 Harding Ludwig, 388
 Hasse Helmut, 611
 Heisenberg Werner, 137
 Helfgott Harald, 702
 héliotrope, 390
 Hellegouarch Yves, 476
 Hellman Martin, 600
 Hemachandra Acharya, 21
 Hermite Charles, 363, 517
 Heyne Christian Gottlob, 382
 Hilbert David, 171, 219, 369, 475,
 689, 695
 Hölder Otto, 403
 homomorphisme
 d'anneaux, 221
 de A-algèbres, 500
 de groupes, 199
 de Frobenius, 448
 de K-algèbres, 259
 homothétie, 111
 Humboldt Alexandre de, 387, 391, 393, 395
 Hussain Zakir, 680
 Huygens Christian, 526, 533
 hyperplan, 96
 hypothèse de Riemann (HR), 695, 697

Index

I

idéal (aux), 223
 d'une algèbre, 500
 engendré, 224
 étrangers, 459
 maximal, 468
 premier, 468
 principal, 224
 strict, 468
identité
 de Bézout, 227
 de Lagrange, 292
image
 d'une application linéaire, 116
 d'un élément, 3
 d'un homomorphisme de groupe, 199
indéterminée, 247
indicateur d'Euler, 423
inégalité
 de Cauchy-Bunyakovski-Schwarz, 336
 triangulaire, 338
inverse
 d'un élément, 194, 223
 d'une matrice, 143
isométrie, 348
isomorphisme
 d'espaces vectoriels, 121
 de groupes, 200

J

Jacobi Karl, 289, 291, 394
jauge euclidienne, 471
Jean-Paul (Richter), 394
jeu de taquin, 211
Jordan Camille, 28-30, 170, 193, 308, 321,
 339, 349, 632

K

K-algèbre, 258
K-algèbre commutative, 248
Karatsuba Anatole, 157
Kästner Abraham, 382
Kayal Neeraj, 606
Klein Félix, 171, 386

Koblitz Neal, 607
Kontsevitch Maxim, 483
Kronecker Leopold, 632
Kummer Eduard, 476

L

Lagrange Joseph Louis, 381, 383, 385, 464,
 523, 526, 531, 561, 649, 652, 681
Laguerre Edmond, 137
Lambert Heinrich, 381, 393, 530
Laplace Pierre Simon de, 254, 274, 275,
 359
Laugel Léonce, 698
Lavinde Gabriel de, 597
Legendre Adrien-Marie, 359, 363, 383, 385,
 389, 561, 662, 692, 700
Leibniz Gottfried von, 3, 55, 231, 234, 483,
 508
lemme de Gauss, 229, 252
Léonard de Pise, 21, 22
Lesage Alain-René, 381
L'Hospital Guillaume de, 55
Le Verrier Urbain, 321
Libri Guillaume, 22
Lichtenberg Georg Christoph, 382
Lindenau Bernhard, 394
linéairement
 dépendants, 91
 indépendants, 91
Liouville Joseph, 691
Littlewood John, 699
Liu Hui, 53, 54, 135
Lobatchevski Nikolai, 380, 382
logarithme discret, 600
logarithme intégral, 699
longueur
 d'un code, 583
 d'un cycle, 202
Louis XVIII, 379
Lucas Édouard, 571

M

Mac Kay James, 419
Mac Lane Saunders, 171, 172,
 190, 411

Mac Laurin Colin, 542
 Mason R. C., 540, 686
 Masser David, 686, 688
 matrice, 143
 antisymétrique, 291
 associée à une forme bilinéaire, 629
 associée à un produit scalaire, 334
 carrée, 137
 colonne, 138
 de contrôle, 587
 de passage, 148
 de transposition, 139
 de transvection, 138
 diagonale, 137
 diagonalisable, 299
 d'une application linéaire, 131
 d'un vecteur, 132
 ligne, 138
 nilpotente, 154
 semblable, 152
 symétrique, 139
 triangulaire inférieure, 138
 triangulaire supérieure, 138
 unité, 137
 Maynard James, 689
 Mazur Barry, 683
 Mengoli Pietro, 690
 Mersenne Marin, 604
 Merzbach Uta, 396
 message, 583
 mesure d'induction
 magnétique, 392
 méthode des moindres carrés, 347, 385,
 389, 394
 Miller Gary, 605, 702
 Minitel, 612
 Minkowski Hermann, 171
 Möbius August, 698
 monôme, 247, 519
 Mordell Louis, 682
 morphisme de G-ensembles, 417
 Morse Samuel, 581
 mot du code, 583
 Müller David, 582
 multiple, 465
 multiplicité d'une racine, 252

N

Napoléon, 387, 389
 Neugebauer Otto, 236
 Newton Isaac, 381, 541, 542, 608
 nilpotent, 154, 220
 Noether Emmy, 171, 475
 Noether Max, 475
 nombre
 algébrique sur $\mathbb{Q}$, 483
 composé, 225
 d'or, 23
 premier, 225
 premiers entre eux, 228
 norme
 d'un élément de $\mathbb{Q}[\sqrt{n}]$, 454
 d'un vecteur, 330, 336
 noyau
 d'un homomorphisme de
 groupes, 200
 d'une application linéaire, 115
 d'une forme bilinéaire symétrique, 641
 numérateur, 480
 n -uplet, 30

O

Oesterlé Joseph, 686, 688
 Olbers Heinrich Wilhelm, 387-389, 392,
 394
 opération, 416
 opérer, 416
 orbite, 417
 ordre
 d'un élément, 198
 d'un groupe, 198
 d'une matrice, 137
 d'une racine, 252
 Oresme Nicole, 691
 orientation, 289
 orthogonal d'un sous-espace, 181, 342
 orthogonalité, 641
 de deux vecteurs, 330
 d'un vecteur et d'une forme
 linéaire, 181
 Osthoff-Gauss Johanna, 386-389

P

- Painvin Georges, 598
- partie
 - entière, 529
 - multiplicative, 482
- partition
 - d'un ensemble, 400
 - d'un entier, 425
- Peano Giuseppe, 74, 75, 109, 329
- Pell John, 464
- Pépin Théophile, 572
- Perec Georges, 597
- période, 483
- permutation, 201
 - impaire 206
 - paire 206
- Pernety Joseph Marie, 386
- Pfaff Johann Friedrich, 383
- pgcd, 226, 466
- Piazzzi Giuseppe, 347, 385, 388
- Pincherle Salvatore, 74
- pivot de Gauss, 39, 60, 67, 68
- plan
 - artinien, 645
 - hyperbolique, 644
 - vectorel, 31
- Plimpton George, 235
- plus grand commun diviseur, 466
- plus petit commun multiple, 466
- Poe Edgar Allan, 597
- poids, 585
- Poincaré Henri, 682
- point fixe, 417
- Pollard John M., 608
- polynôme, 246
 - annulateur, 309
 - à plusieurs indéterminées, 518
 - caractéristique, 302
 - cyclotomique, 512
 - de Bernstein, 104
 - de Jacobi, 362
 - de Laguerre, 363
 - de Legendre, 359, 363, 364
 - dérivé, 255
 - de Tchebycheff, 363, 370
 - d'Hermite, 363
 - d'interpolation de Lagrange, 103
 - générateur d'un code cyclique, 590
 - homogène, 520
 - minimal, 309
 - primitif, 505
 - scindé, 254
 - symétrique, 520
 - unitaire, 248
- polynômes premiers entre eux, 251
- Pouchkine, 394
- Pouillet-Delisle A. C. M., 383
- ppcm, 466
- préhilbertien, 332
- principe de superposition, 12, 15, 119
- procédé d'orthogonalisation de Gram-Schmidt, 340
- produit
 - d'anneaux, 456
 - de fonctions, 5
 - de groupes, 411
 - de polynômes, 247
 - d'espaces vectoriels, 166, 169
 - de suites, 17
 - d'idéaux, 458
 - d'une application linéaire par un scalaire, 111
 - d'une fonction par le scalaire, 5
 - d'une suite par un scalaire, 18
 - d'un vecteur par un scalaire, 31, 76
- produit scalaire, 330, 332
- projecteur, 169
- projection, 112, 113, 169
 - canonique, 401
 - orthogonale, 331, 344
- pseudopremier, 604
- p-sous-groupe de Sylow, 427
- Puvis de Chavannes Pierre, 29
- Pythagore, 338
- Python, 12

Q

- quaternion, 646
 - pur, 649
- quotient d'espace vectoriel, 170

de division euclidienne, 215, 249, 471
 d'un anneau, 449
 d'un ensemble, 401
 d'un groupe, 406

R

Rabin Mickael, 605, 702
 Rabin-Miller, 605
 racine d'un polynôme, 252
 radical, 687
 Ramanujan Srinivasa, 680
 rang
 d'une application linéaire, 117
 d'une famille de vecteurs, 44
 d'une forme bilinéaire symétrique, 630
 d'une matrice, 152
 du système linéaire, 63
 redondance, 583
 réduction, 299
 réduire, 299
 réduite (fraction continue), 531
 Reed Irving, 582
 réflexion, 622
 règle de Sarrus, 279
 relation d'équivalence, 217
 relation de récurrence linéaire
 avec second membre, 18
 sans second membre, 18
 représentant, 217, 401
 représentation propre, 652
 résidu quadratique, 561
 reste d'une division euclidienne, 215, 249, 471
 retournement, 622
 Ribet Ken, 476
 Richardson Samuel, 381
 Riemann Bernhard, 171, 390, 395, 679, 695, 696-698, 701
 Rijmen Vincent, 599
 Rivest Ron, 601
 Rodrigues Olinde, 363
 Rossum Guido (van), 12
 rotation, 349
 Rousseau Jean-Jacques, 381, 387
 Roy Marie-Françoise, 517

RSA, 601
 Rubik's cube, 680

S

Saccheri, 393
 Sachs Abraham (1914-1983), 236
 Sage, 12, 147
 Sarrus Pierre, 279
 Sartorius Wolfgang, 395
 Saxena Nitin, 606
 Serre Jean-Pierre, 476
 scalaire, 31
 Schiller, 394
 Schmidt Erhard, 170, 333, 340, 369
 Schumacher Heinrich Christian, 389, 393, 396
 Schwarz Herman, 337
 Scott Walter, 394
 section, 400
 Selberg Atle, 701
 Serre Jean-Pierre, 476
 Shakespeare William, 542
 Shamir Adi, 601
 Shankar Arul, 685
 Shannon Claude, 582, 587
 Shimura Goro, 476
 Siegel Carl, 699
 signature d'une forme quadratique sur $\mathbb{R}$, 638
 signature d'une permutation, 204
 Skinner Christophe, 686
 Snyder Noah, 540
 Solomon Gustave, 582
 solution générale
 d'une équation différentielle, 10
 d'une relation de récurrence linéaire, 21
 solution particulière
 d'une équation différentielle, 10
 d'une relation de récurrence linéaire, 21
 somme
 d'applications linéaires, 111
 de fonctions, 3
 de suites, 17
 de vecteurs, 76
 de vecteurs de $\mathbb{R}^n$, 31
 d'idéaux, 458

Index

- directe, 164, 166
- orthogonale, 644
- source, 3
- sous-anneau, 445
 - engendré, 446
- sous-corps engendré, 483
- sous-espace stable, 116
- sous-espace vectoriel, 79
- sous-groupe, 404
 - conjugué, 405
 - distingué, 404
 - engendré, 197
 - normal, 404
- stabilisateur, 417
- stabilité
 - par produit par un scalaire, 32
 - par somme, 32
- stathme, 471
- Stein William, 12
- Stothers Wilson, 686
- Strassen Volker, 157
- structure de groupe, 399
- Sturm Charles-François, 517
- Suétone, 597
- suite, 17, 18, 474
 - de Fibonacci, 318
 - de Sturm, 515
 - stationnaire, 474
- Swinnerton-Dyer Peter, 686
- Sylow Ludwig, 419
- Sylvester James, 135
- symétrie, 113
- symétrie orthogonale, 349
- symétriques, 628
- système de représentants, 400, 401, 467
- système linéaire
 - de Cramer, 285
 - sans second membre, 56
- systèmes linéaires équivalents, 56

T

- taille d'une matrice, 137
- Taniyama Yukata, 476
- Tao Terence, 683, 689
- Tarski Alfred, 517

- Taurinus F.A., 393
- taux d'information d'un code, 583
- Taylor Richard, 477
- Tchebycheff Pafnuty, 362, 371, 689, 702
- Terquem Olry, 691
- test AKS, 606
- test de Rabin-Miller, 234, 605
- théorème
 - chinois, 459
 - de Cauchy, 419
 - de d'Alembert, 254
 - de Fermat, 231, 233, 688
 - de Hamilton-Cayley, 313
 - de Lagrange, 198
 - de Mason-Stothers, 540
 - de Pythagore, 338
 - des noyaux, 310
 - du losange, 424, 485
- trace
 - d'une application linéaire, 153
 - d'une matrice, 152
- transformation
 - orthogonale, 348
 - orthogonale directe, 349
 - orthogonale indirecte, 349
- transitive, 417
- transposée
 - d'une application linéaire, 183
 - d'une matrice, 184
- transposition, 202
- triangulaire, 312
- triangularisation, 311
- trigonaliser, 299
- triplet, 30
 - pythagoricien, 235
- Tukey John, 581
- Turing Alan, 598
- type du code, 585

U

- unité, 463

V

- valeur propre, 301
- Vallée Poussin Charles de la, 700

valuation, 473
 Vandermonde Alexandre, 274, 275, 293, 294, 523
 Van der Waerden Bartel Leendert, 75, 475
 Vect, 18, 32, 82
 vecteur, 76, 647
 de $\mathbb{R}^n$, 31
 isotrope, 641
 nul, 31
 propre, 302
 vecteurs orthogonaux, 330, 333, 641
 Victoria reine, 393
 Viète François, 55, 245
 Vigénère Blaise de, 598
 Volterra Vito, 4, 75
 volume, 287

W

Wagner Rudolf, 395
 Waldeck-Gauss Minna, 389-391
 Waring Edward, 523

Weber Wilhem, 392, 394
 Wedderburn Joseph, 555
 Weierstrass Karl, 632
 Weil André, 5, 465, 476, 611, 682, 703
 Weyl Hermann, 171
 Wiles Andrew, 477, 608, 680, 688, 689
 Wilson John, 234
 Witt Ernst, 646

Y

Yoccoz Jean-Christophe, 526

Z

Zach Franck Xaver von, 385
 Zagier Don, 483, 573
 Zermelo Ernst, 400
 zéro d'un polynôme, 252
 Zhang Wei, 686
 Zhang Ytang, 689
 Zimmermann Ebherhard, 381, 385
 Zorn Max, 470