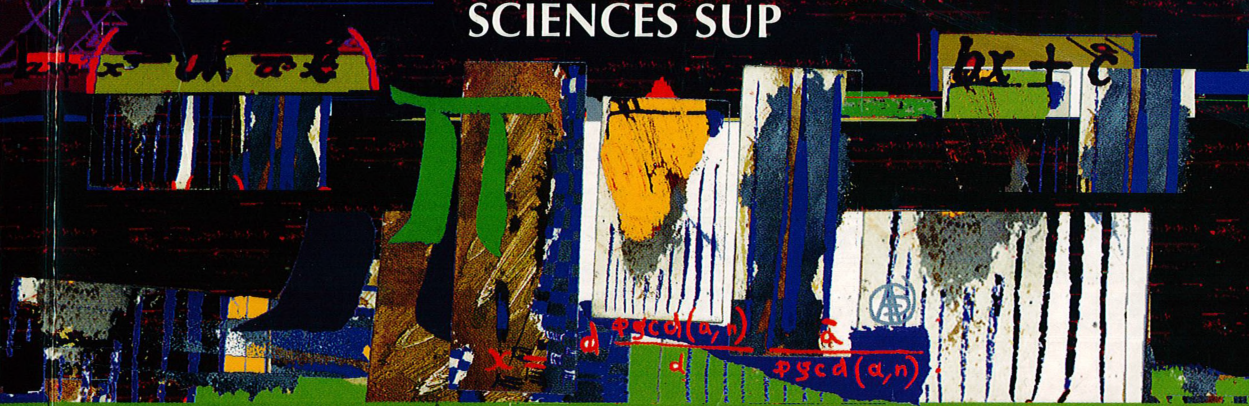


SCIENCES SUP



Cours et exercices avec solutions

Licence 3^e année

ALGÈBRE

3^e ANNÉE

2^e édition

Lionel Schwartz

DUNOD

ALGÈBRE

3^e ANNÉE

Consultez nos catalogues
sur le Web

<http://www.dunod.com>

SCIENCES ET TECHNIQUES

INFORMATIQUE
Dunod & Microsoft Press

GESTION MANAGEMENT

SCIENCES HUMAINES

Acheter en ligne

Nouveautés

DUNOD
Microsoft Press
Ediscience ETSF
InterÉditions

Compléments en ligne

Magazine et interviews d'auteurs

ALGÈBRE

3^e ANNÉE

Cours et exercices avec solutions

Lionel Schwartz

Professeur à l'université de Paris-Nord
(Paris 13 – Villetanneuse)

Préfacé par

Michel Zisman

Professeur émérite de l'université Paris 7 - Denis Diderot

2^e édition

DUNOD

• **Cours de Mathématiques, filières MIAS, MASS et SM (4 volumes)**

Analyse, 1^{re} année

Algèbre, 1^{re} année

Analyse, 2^e année

Algèbre et géométrie, 2^e année

François Liret et Dominique Martinais

• **Cours de Mathématiques pour la 3^e année de licence (4 volumes)**

Algèbre, Lionel Schwartz

Fonctions analytiques, Pierre Vogel

Topologie et analyse, Georges Skandalis

Calcul différentiel et calcul intégral, Marc Chaperon

Illustration de couverture : *Lionel Auvergne*

Ce pictogramme mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du **photocopillage**.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les

établissements d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation du Centre français d'exploitation du droit de copie (**CFC**, 20 rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2003

ISBN 2 10 007057 6

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (Art L 122-4) et constitue une contrefaçon réprimée par le Code pénal. • Seules sont autorisées (Art L 122-5) les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L 122-10 à L 122-12 du même Code, relatives à la reproduction par reprographie.

Préface

C'est avec ce cours d'Algèbre que les éditions Dunod ont initié une série d'ouvrages destinés aux étudiantes et étudiants préparant la LICENCE DE MATHÉMATIQUES. Cette série, de même que le cours des deux premières années de LICENCE de François Liret et Dominique Martinais dans la même collection, a pour vocation de présenter à ses lecteurs les théories classiques qui constituent le fond usuel des programmes des premières années d'études universitaires fécondées par l'esprit des idées et des connaissances provenant des recherches contemporaines. Les théories les plus abstraites sont riches d'applications concrètes importantes récemment découvertes que l'on peut maintenant présenter parfois dans un cours de ce niveau, non seulement pour illustrer sur des exemples les méthodes et les théorèmes généraux, mais surtout pour montrer explicitement aux étudiants et étudiantes la profonde unité de la science et son impact sur la vie moderne.

Le cours d'Algèbre de Lionel Schwartz, premier de la série, se devait de prouver que cet objectif pouvait être atteint. La richesse et la densité de son texte, sa présentation de la cryptographie à clés publiques, aujourd'hui couramment utilisée comme application de la théorie des corps finis, ainsi que d'autres applications aussi intéressantes sinon aussi spectaculaires, dans un ouvrage dont la vocation reste élémentaire et qui distingue soigneusement le niveau LICENCE de celui du MASTER, tout cela montre qu'il y a parfaitement réussi.

Michel Zisman

Professeur émérite de l'université Paris 7 - Denis Diderot

Table des matières

Chapitre 0.

Rappels

1. Rappel de théorie des ensembles	1
2. Rappels sur l'anneau $\mathbb{Z}$ des entiers relatifs et son corps des fractions $\mathbb{Q}$	2
3. Topologie de $\mathbb{R}^n$	4
4. Le corps des nombres complexes $\mathbb{C}$	4
5. Espaces vectoriels et algèbre linéaire	5
6. Matrices, déterminants, systèmes linéaires	8

Chapitre 1.

Groupes. Groupes quotients Groupes abéliens de type fini

1. Groupes, générateurs, groupes monogènes	12
2. Relations d'équivalence dans un groupe, groupes quotients	22
3. Décomposition primaire des groupes abéliens finis	33
4. Structure des groupes abéliens de type fini, définitions, énoncés, structure des groupes libres	36
5. Structure des groupes abéliens de type fini de torsion	41
Exercices	44

Chapitre 2.

Actions de groupes Groupes symétriques

1. Actions de groupes, exemples	51
2. Orbites, ensembles transitifs, décomposition en orbites	56
3. Théorèmes de Sylow	61
4. Groupes symétriques, classes de conjugaison	64
5. Groupes symétriques, générateurs, simplicité du groupe alterné	72
Exercices	77

Chapitre 3.

Anneaux, idéaux, polynômes et séries formelles

1. Anneaux, idéaux, anneaux quotients	85
2. Anneaux principaux	96
3. L'anneau $\mathbb{Z}/n\mathbb{Z}$	103
4. Cryptographie à clé révélée	105

5. Anneaux de polynômes et de séries formelles	110
6. Polynômes en plusieurs indéterminées, polynômes symétriques	116
7. Anneaux factoriels et applications	122
8. Zéros des polynômes	127
9. Compléments sur les séries formelles	133
Exercices	135

Chapitre 4. Extensions des corps. Applications

1. Définitions, caractéristique d'un corps, applications	150
2. Éléments algébriques dans un corps et extensions algébriques des corps	153
3. Corps de rupture d'un polynôme	158
4. Corps algébriquement clos, clôture algébrique	160
5. Structure et classification des corps finis	162
6. Applications arithmétiques	167
7. Racines de l'unité, polynômes cyclotomiques	169
8. Théorème de Wedderburn	172
9. L'algorithme de Berlekamp	174
10. Les codes correcteurs d'erreurs	178
Exercices	191

Chapitre 5. Réductions des endomorphismes Structure du groupe linéaire

1. Polynôme minimal d'un endomorphisme	204
2. Théorème de Cayley-Hamilton	210
3. Triangularisation des matrices et réduction de Jordan	214
4. Applications topologiques de la diagonalisation et de la triangularisation	220
5. Générateurs du groupe linéaire	225
Exercices	234

Chapitre 6. Formes bilinéaires et sesquilineaires Groupes orthogonaux et unitaires

1. Formes bilinéaires et sesquilineaires	245
2. Structure du groupe orthogonal euclidien	255
3. Les quaternions, les groupes $SO(3)$ et $SO(4)$	264
4. Structure du groupe unitaire	273
Exercices	276

Bibliographie	283

Index	285

Avant-Propos

Vous avez entre les mains la seconde édition d'un cours d'algèbre de LICENCE. Cette édition arrive au moment où les cursus universitaires sont en cours de refonte, pour passer du système « DEUG, LICENCE, MAÎTRISE, DEA » au système dit « LICENCE, MASTER, DOCTORAT ». Ce livre s'adresse aux étudiants de troisième année, c'est-à-dire l'actuelle, et la fin de la future, LICENCE. Comme il se place à un moment charnière dans les études de mathématiques, il a vocation à intéresser au-delà de ce public, les étudiants qui, après la LICENCE, iront préparer le CAPES, ou ceux qui poursuivront en MASTER (quelle qu'en soit l'orthographe !), puis passeront l'AGRÉGATION.

Le programme officiel de la LICENCE DE MATHÉMATIQUES est très vague, et les traditions sont variées d'une université à l'autre. Nous avons voulu rester dans un équilibre raisonnable, en tenant compte du fait que l'algèbre est une discipline fondamentale du programme de troisième année, mais que certaines théories (telle que la théorie de Galois, par exemple) relèvent plutôt du programme de MASTER. Nous ne prétendons certainement pas que tout le contenu du livre doit être fait en LICENCE mais que l'on peut y faire des choix en fonction de ce qu'on souhaite faire au-delà.

L'algèbre, tout comme la géométrie, sont des matières que les programmes des lycées ont une fâcheuse tendance à négliger (quand bien même on assiste à un modeste retour). Ceci entraîne des retards qu'il est important de corriger. Les prérequis consistent ici en de solides bases d'algèbre linéaire et de géométrie euclidienne, qui relèvent des deux premières années de la future LICENCE. Un premier contact avec les groupes est, lui aussi, souhaitable.

Il y a trois notions centrales dans ce cours. D'abord, la structure des groupes et les actions de groupes, en insistant particulièrement sur les groupes symétriques. Puis viennent les anneaux, avec la question de la divisibilité sous ses divers aspects. Enfin, les extensions de corps, sans la théorie de Galois, en s'attardant plus spécialement sur les corps finis.

Certains énoncés ne sont pas centraux et peuvent être réservés à une seconde lecture. Par exemple, les théorèmes de Sylow, ou bien la démonstration des théorèmes de structure des groupes abéliens de type fini.

Les quatre premiers chapitres sont consacrés à la présentation de ces différentes notions. Les deux derniers sont, quant à eux, des compléments sur la réduction des endomorphismes et sur les formes quadratiques. Un certain nombre de théorèmes plus avancés sont abordés dans les exercices (deux étoiles « ** » après un numéro d'exercice indiquent une difficulté plus grande). On accorde aussi une place privilégiée aux propriétés topologiques des groupes classiques.

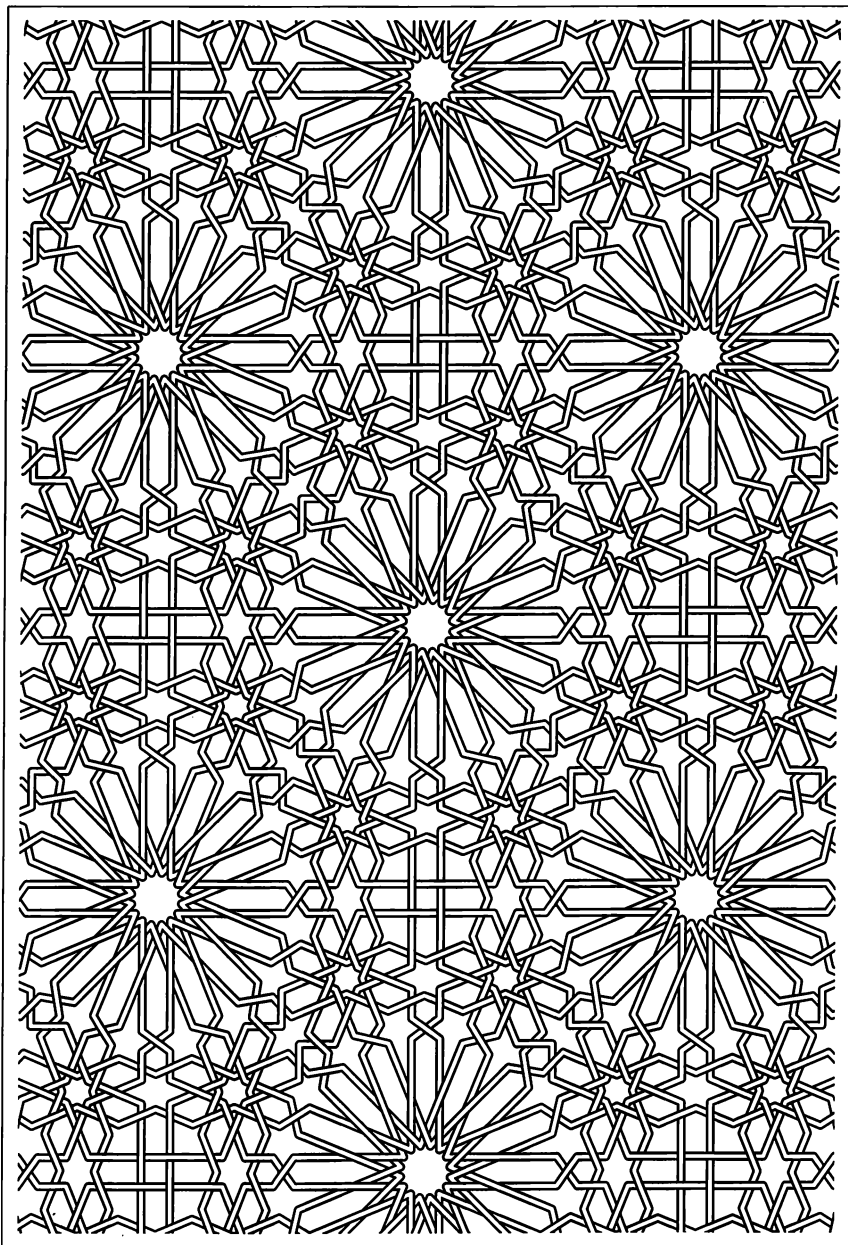
Enfin, l'ouvrage introduit des applications classiques de ces mathématiques (parmi d'autres !) : la cryptographie et les codes correcteurs d'erreurs, qui sont deux prolongements naturels de ce cours. Si les techniques utilisées dans ce cadre ne sont pas plus difficiles, voire sont plus simples, que les autres développées dans ce livre, leur enchaînement rend l'approche du problème plus délicate.

Dans la mesure où ce livre reflète tout à la fois des choix personnels, associés à des contraintes de temps et d'espace, mentionnons un certain nombre de sujets qui auraient pu être abordés. Le choix de faire une étude des actions de groupes en liaison beaucoup plus étroite avec la géométrie n'a pas été retenu ici. Le lecteur pourra se référer à [Be] sur ce point. D'autre part, l'étude des séries formelles en liaison avec la combinatoire pouvait fournir d'autres développements, qui sont exposés plus précisément dans [Co]. Ces sujets peuvent trouver, évidemment, leur place en MASTER.

Ce cours peut déboucher, ainsi qu'on l'a dit plus haut, sur des applications « concrètes » ; ou tout au moins sur d'autres cours qui, eux, y donnent accès. Les mathématiques ont toujours eu des liens étroits, explicites ou implicites, avec d'autres activités, même si ces liens ne sont pas toujours apparents. La figure ci-dessous est, un peu simplifiée, la trame d'un zellige des tombeaux saadiens à Marrakech. Cet entrelacs présente des symétries internes par rapport à des groupes diédraux d'ordres 32 et 16, et est également invariant par diverses actions de translation. On y retrouve le savoir-faire des artisans et des artistes arabes. Et ici, on se souviendra du rôle crucial des savants arabes dans la transmission et le développement des mathématiques.

Il convient enfin de dire deux mots sur les différences par rapport à la première édition. D'abord, on a apporté des précisions et des corrections, je tiens ici à remercier les collègues et tout particulièrement les étudiants qui m'ont signalé tel ou tel point à corriger ou préciser. Ensuite, quelques ajouts ont été faits : en cryptographie, codes correcteurs et sur les sous-groupes finis $SO(3)$. Enfin, on a ajouté des exercices et complété des corrections.

Je remercie encore une fois Michel Zisman de m'avoir proposé d'écrire ce cours, Alberto Arabia, pour son soin, sa superbe mise en pages, et les figures qu'il a faites. Je remercie enfin Anne Bourguignon, pour son suivi dans l'édition de cet ouvrage.



Principales notations utilisées

$\mathbb{N}$	ensemble des entiers naturels
$\mathbb{Z}$	ensemble des entiers relatifs
$\mathbb{Q}$	ensemble des nombres rationnels
$\mathbb{R}$	ensemble des nombres réels
$\mathbb{C}$	ensemble des nombres complexes
$\mathbb{F}_{p^n}$	corps fini à p^n éléments
$\mathbb{Z}/m\mathbb{Z}$	groupe des classes de congruences modulo m
$\mathcal{P}(E)$	groupe des permutations de l'ensemble E
$\mathfrak{S}_n$	groupe des permutations de l'intervalle d'entiers naturels $[1, n]$
$\mathrm{GL}(E), \mathrm{GL}_n(k)$	groupes linéaires
$\mathrm{SL}(E), \mathrm{SL}_n(k)$	groupes spéciaux linéaires
$\mathrm{O}(E), \mathrm{O}(n)$	groupes orthogonaux
$\mathrm{U}(E)$	groupe unitaire
$\mathbf{Z}(G)$	centre d'un groupe G
$\mathrm{C}_G(x)$	centralisateur d'un élément x du groupe G
$\varphi(d)$	indicateur d'Euler
$k[X]$	anneau des polynômes sur k
$k[[X]]$	anneau des séries formelles sur k
ϕ_n	polynôme cyclotomique
μ_n	racines n -ièmes de l'unité
$[E : F]$	degré d'une extension
Id_E	application identité de l'ensemble E
$\# A, \mathrm{Card} A$	cardinal de l'ensemble fini A

Chapitre 0

Rappels

On rappelle ici quelques notions nécessaires à la lecture du cours. Elles sont supposées avoir été vues durant les deux premières années de LICENCE ou des CLASSES PRÉPARATOIRES. Il s'agit d'abord de rappels élémentaires de théorie des ensembles, puis sur l'anneau $\mathbb{Z}$ et le corps $\mathbb{Q}$, ensuite de notions de topologie de $\mathbb{R}^n$, et enfin des bases d'algèbre linéaire.

Ces rappels sont brefs et doivent être lus avec le soutien d'un manuel de première et seconde année. Ils seront l'occasion de fixer un certain nombre de notations.

1. Rappel de théorie des ensembles

Dans ce livre, on utilisera souvent le résultat de factorisation des applications exposé ci-après.

Rappelons d'abord que, si X désigne un ensemble et $\mathcal{R}$ une relation d'équivalence sur X , l'ensemble des classes d'équivalence pour la relation considérée constitue une partition de X . C'est-à-dire que la réunion de ces classes est égale à l'ensemble X tout entier, et que deux classes disjointes sont d'intersection vide.

L'ensemble des classes d'équivalence est appelé l'ensemble quotient de X par la relation $\mathcal{R}$, et est noté $X/\mathcal{R}$. On appellera l'application π , qui à un élément associe sa classe d'équivalence, application canonique.

Théorème. Soit $f : X \rightarrow Y$ une application telle que

$$x \mathcal{R} y \implies f(x) = f(y).$$

Il existe alors une unique application $\bar{f} : X/\mathcal{R} \rightarrow Y$ telle que $f = \bar{f} \circ \pi$.

Démonstration. Soit c une classe d'équivalence dans X . L'application $\bar{f}$ est définie par $\bar{f}(c) = f(x)$, où x est un élément quelconque de la classe c . ■

Remarque

On peut choisir comme exemple de relation d'équivalence celle définie par : $x \sim y$ si et seulement si $f(x) = f(y)$.

Les propositions 1 et 2 seront fréquemment utilisées.

Proposition 1. Soient $f : X \rightarrow Y$ et $g : Y \rightarrow Z$, deux applications. Alors si l'application composée $g \circ f$ est surjective, l'application g l'est aussi.

Proposition 2. Soit S un ensemble fini, et soit f une fonction de S dans $\mathbb{N}$, alors la fonction f atteint son maximum et son minimum.

2. Rappels sur l'anneau $\mathbb{Z}$ des entiers relatifs et son corps des fractions $\mathbb{Q}$

L'anneau $\mathbb{Z}$ est muni de la division euclidienne :

Proposition 1. Pour tous entiers $n, k \in \mathbb{Z}$ il existe deux entiers uniques q et r tels que :

- $n = qk + r$,
- $0 \leq r < k$.

L'entier r est le reste de la division de n par q .

Si n est non nul, quand le reste est nul on dit que k divise n , on le note $k \mid n$.

Une conséquence fondamentale, qui sera redémontrée dans le chapitre 1 est que tous les sous-groupes de $\mathbb{Z}$ sont de la forme :

$$k\mathbb{Z} = \{ku \mid u \in \mathbb{Z}\}.$$

Dans l'anneau $\mathbb{Z}$, un élément a une décomposition en facteurs premiers, qui est unique à l'ordre près. Ceci permet de définir le plus grand commun diviseur (pgcd), et le plus petit commun multiple (ppcm) de deux entiers, et plus généralement d'une famille d'entiers.

L'identité de Bézout permet de définir le pgcd d'une autre façon :

Proposition 2. Soient $a, b \in \mathbb{Z}$, et soit d leur plus grand commun diviseur. Alors il existe $u, v \in \mathbb{Z}$ tels que $d = ua + bv$. Plus généralement soient $a_1, \dots, a_n \in \mathbb{Z}$, et soit d leur plus grand commun diviseur. Il existe $u_1, \dots, u_n$, tels que

$$a_1 u_1 + \dots + a_n u_n = d.$$

L'algorithme d'Euclide permet de calculer le pgcd de deux entiers a et b , on en déduit aussi un algorithme pour calculer les coefficients de la formule de Bézout. Supposons que $0 < a < b$ et que b_1 est le reste de la division de b par a on constate que :

- soit $b_1 = 0$ et le pgcd cherché est a ,
- soit le pgcd de a et b est aussi celui de a et b_1 .

Dans le second cas on peut itérer le processus et faire la division de a par b_1 et obtenir un reste $a_1 < a$ car $b_1 < a$. Encore une fois on a deux cas possibles :

- soit $a_1 = 0$ et le pgcd est b_1 ,
- soit le pgcd de a et b_1 est aussi celui a_1 et b_1 .

On définit ainsi itérativement a_k et b_k tels que le pgcd de a_k et b_{k+1} , ainsi que celui de a_k et b_k , est celui de a et b .

Dans ce processus on arrive, soit à un moment où a_n divise b_n et a_n est le pgcd, soit b_n divise a_{n-1} et b_n est le pgcd. Si on arrive à $a_n = 1$ ou $b_n = 1$ les entiers a et b sont premiers entre eux.

Rappelons maintenant qu'une *loi interne* sur un ensemble A est une application de $A \times A$ dans A . Passons à la construction du corps $\mathbb{Q}$ des fractions de $\mathbb{Z}$.

Construction du corps $\mathbb{Q}$. On considère l'ensemble $\mathbb{Z} \times \mathbb{Z}^*$ et on le munit de la relation d'équivalence suivante : $(a, b) \cong (u, v)$ si et seulement si $av = bu$.

On définit deux lois internes, addition et multiplication, par :

- $(a, b) + (u, v) = (av + bu, bv)$,
- $(a, b) \times (u, v) = (au, bv)$.

On vérifie aisément que la classe d'équivalence de la somme et du produit de deux éléments ne dépend que des classes d'équivalence des deux éléments. Ceci permet de définir l'addition et la multiplication sur l'ensemble quotient. Les deux lois sont évidemment commutatives.

La multiplication est distributive par rapport à l'addition. C'est-à-dire que :

$$(u, v) \times ((a, b) + (a', b')) = (u, v) \times (a, b) + (u, v) \times (a', b').$$

Par ailleurs l'ensemble des classes d'éléments de la forme $(k, 1)$ est isomorphe à $\mathbb{Z}$ et la classe de l'élément $(1, k)$, $k \neq 0$, est un inverse pour la multiplication de la classe de $(k, 1)$.

Le corps $\mathbb{Q}$ est l'ensemble de ces classes d'équivalence de $\mathbb{Z} \times \mathbb{Z}^*$ muni de ces deux lois internes.

3. Topologie de $\mathbb{R}^n$

Rappelons d'abord les propriétés fondamentales de $\mathbb{R}$.

Théorème 1. *Le corps $\mathbb{Q}$ est dense dans le corps $\mathbb{R}$ et $\mathbb{R}$ est complet : toutes les suites de Cauchy convergent dans $\mathbb{R}$.*

Théorème 2. *Les sous-espaces connexes de $\mathbb{R}$ sont les intervalles, c'est-à-dire les sous-espaces convexes.*

Théorème 3. *Les sous-espaces compacts de $\mathbb{R}$ sont les ensembles fermés et bornés.*

Ce résultat se généralise à $\mathbb{R}^n$:

Théorème 4. *Les sous-espaces compacts de $\mathbb{R}^n$ sont les ensembles fermés et bornés.*

On dit qu'un espace topologique X est connexe par arcs si pour toute paire de points $x, y \in X$, il existe une application continue, f de $[0, 1]$ dans X telle que $f(0) = x$ et $f(1) = y$. Un espace topologique connexe par arcs est toujours connexe. Le résultat inverse est, en général, faux, mais on a le résultat suivant :

Théorème 5. *Les ouverts connexes de $\mathbb{R}^n$ sont connexes par arcs.*

Pour terminer ce type de rappel mentionnons aussi l'équivalence des normes sur les espaces vectoriels réels de dimension finie [LA].

4. Le corps des nombres complexes $\mathbb{C}$

Le corps des nombres complexes $\mathbb{C}$ est construit comme suit.

Il est défini en tant qu'ensemble comme étant $\mathbb{R} \times \mathbb{R}$, que l'on munit des deux lois internes, addition et multiplication, suivantes :

- ▶ $(a, b) + (u, v) = (a + u, b + v),$
- ▶ $(a, b) \times (u, v) = (au - bv, av + bu).$

L'élément (a, b) est noté comme d'habitude $a + bi$, avec $i = (0, 1)$, et bien entendu $i^2 = -1 = (-1, 0)$.

Ces deux lois sont commutatives, associatives. La multiplication est distributive par rapport à l'addition. L'élément $0 = (0, 0)$, est élément neutre pour l'addition, alors que l'élément $1 = (1, 0)$ l'est pour la multiplication.

Le conjugué d'un élément $z = a + bi$, noté $\bar{z}$, est $a - bi$. Le module $|z|$ de l'élément $z \in \mathbb{C}$ est $\sqrt{a^2 + b^2}$ et $|z|^2 = z\bar{z}$. L'inverse d'un élément non nul z est $\frac{\bar{z}}{|z|^2}$.

Étant donné un nombre complexe z l'exponentielle, notée e^z , de z est définie comme étant la somme de la série absolument convergente ;

$$\sum_{n \geq 0} \frac{z^n}{n!}.$$

On a pour tous $z, z' \in \mathbb{C}$ la formule

$$e^z e^{z'} = e^{z+z'}.$$

Enfin on a :

Proposition 1. *Soit z un nombre complexe, les deux conditions suivantes sont équivalentes :*

- $e^z = 1$,
- il existe $k \in \mathbb{Z}$ tel que $z = 2k\pi i$.

Notons que l'application exponentielle est un homomorphisme de $\mathbb{C}$ muni de l'addition dans $\mathbb{C}^*$ muni de la multiplication. La proposition précédente en décrit le noyau.

Corollaire. *Soit z un nombre complexe. Les deux conditions suivantes sont équivalentes :*

- $z^n = 1$,
- $z = e^{\frac{2k\pi i}{n}}$ pour un certain entier k .

5. Espaces vectoriels et algèbre linéaire

Le but de cette section est de rappeler les résultats fondamentaux en algèbre linéaire. Ceux-ci ont été vus le plus souvent en DEUG (désormais LICENCE 1^{re} ET 2^e ANNÉE) ou CLASSES PRÉPARATOIRES dans le cadre des espaces vectoriels réels ou complexes. Il ne se passe absolument aucun changement pour les résultats rappelés ci-dessous si on remplace $\mathbb{R}$ ou $\mathbb{C}$ par un corps k quelconque. C'est-à-dire que les démonstrations sont mot pour mot les mêmes en substituant à $\mathbb{R}$ ou $\mathbb{C}$ le corps k .

Dans la mesure où le lecteur peut ne pas avoir manié d'autres corps que $\mathbb{R}$ ou $\mathbb{C}$, ni même avoir rencontré la définition générale d'un corps, nous lui proposons de substituer dans les énoncés qui suivent k à $\mathbb{R}$ ou $\mathbb{C}$. Puis, dès qu'il sera plus familier avec la notion de corps et qu'il aura à l'utiliser, ce qui se produira à partir du chapitre 3, de faire un retour en arrière, de prendre un manuel de première et deuxième années et de vérifier les énoncés pas à pas en procédant à la substitution inverse. Bien souvent dans les manuels de première et seconde année les énoncés commencent par : « Soient $k = \mathbb{R}$ ou $k = \mathbb{C}$, ... ». Il suffit d'oublier cette phrase.

Commençons par la définition. Soit donc k un corps et soit E un groupe abélien.

On dit que E est un espace vectoriel sur k si on a une application (aussi appelée loi externe ou opération) :

$$k \times E \longrightarrow E, (\lambda, v) \longmapsto \lambda v,$$

telle que :

- $1v = v,$
- $\lambda(v + w) = \lambda v + \lambda w,$
- $(\mu + \lambda)v = \mu v + \lambda v,$
- $\mu(\lambda v) = (\mu\lambda)v,$

ceci pour tous $v, w \in E$ et $\lambda, \mu \in k$.

Les conditions ci-dessus impliquent que pour tout $v \in E$, on a :

$$0 \times v = 0.$$

On a la notion de sous-espaces vectoriels. C'est un sous-groupe (donc non vide) $H \subset E$ tel que pour tous $v \in H$ et $\lambda \in k$ on ait $\lambda v \in H$.

L'intersection d'une famille de sous-espaces vectoriels est un sous-espace vectoriel. Étant donné un sous-ensemble A d'un espace vectoriel E l'intersection de tous les sous-espaces vectoriels de E contenant A est appelé le sous-espace vectoriel engendré par A . C'est le plus petit espace vectoriel qui contient A .

Soit E un espace vectoriel sur k , soit $(v_i), i \in I$ une famille d'éléments de E indexée par un ensemble I non nécessairement fini. De même manière soit (λ_i) une famille d'éléments de k indexée par I .

On dit que les éléments d'une telle famille sont presque tous nuls s'ils sont tous nuls sauf un nombre fini d'entre eux.

Si la famille (λ_i) est constituée d'éléments presque tous nuls, on peut définir la somme

$$\sum_{i \in I} \lambda_i v_i.$$

En effet, seuls un nombre fini d'éléments sont non nuls. Une telle somme est appelée une combinaison linéaire à coefficients dans k des vecteurs v_i .

- Une famille (v_i) d'éléments de E est génératrice sur k , ou constitue un système de générateurs sur k , si le plus petit sous-espace vectoriel de E contenant tous les v_i est E lui-même.

Si E a au moins un élément non nul ceci équivaut à dire que pour tout $v \in E$ il existe une famille d'éléments, presque tous nuls, de k indexée par I , soit (λ_i) telle que $v = \sum_i \lambda_i v_i$. Si E est réduit au vecteur nul la partie vide est génératrice.

- Une famille (v_i) d'éléments de E est libre sur k si la relation $\sum_i \lambda_i v_i = 0$, (λ_i) désignant une famille d'éléments de k , presque tous nuls, indexées par I , n'a lieu que si tous les éléments λ_i sont nuls.
- Une famille qui est génératrice et libre est appelée une base de E sur le corps k .

Il en résulte que, si E a une base sur k , tout élément de $v \in E$ s'écrit d'une manière et d'une seule comme combinaison linéaire des vecteurs de la base.

Théorème 1. *Un espace vectoriel E sur un corps k a au moins une base. Toutes les bases ont même nombre d'éléments, appelé la dimension de l'espace.*

On remarquera que l'espace vectoriel $\{0\}$ admet pour base l'ensemble vide et est de dimension zéro.

Rappelons le théorème de la base incomplète :

Théorème 2. *Soit E un espace vectoriel sur un corps k , et soit (v_i) , $i \in I$ une famille libre. Alors il existe au moins une autre famille libre (w_j) telle que la réunion des deux familles constitue une base de E .*

La démonstration fait un usage crucial de la possibilité de trouver un inverse pour les éléments non nuls de k .

On appelle hyperplan d'un espace vectoriel de dimension n un sous-espace de dimension $n - 1$, plan un sous-espace de dimension 2 et droite un sous-espace de dimension 1. On précise parfois droite vectorielle, plan vectoriel, ..., pour distinguer de la situation affine.

Soient deux espaces vectoriels E et F sur un corps k , une application f de E dans F est dite linéaire (k -linéaire si on tient à rappeler le corps) si :

- f est un homomorphisme de groupes abéliens, c'est-à-dire que $f(v + w) = f(v) + f(w)$ pour tous $v, w \in E$,
- $f(\lambda v) = \lambda f(v)$ pour tous $\lambda \in k$ et $v \in E$.

L'ensemble des applications k -linéaires de E dans F est lui-même un espace vectoriel sur k , il sera noté

$$\mathcal{L}_k(E, F).$$

Quand le choix du corps k ne sera pas ambigu l'indice k sera omis, ce sera souvent le cas quand $k = \mathbb{C}$. Si $E = F$ $\mathcal{L}_k(E, E)$ sera noté $\mathcal{L}_k(E)$.

La composition des applications détermine une structure d'anneau sur cet espace vectoriel. On signale deux notations concurrentes, de plus en plus en plus fréquentes, $\text{Hom}_k(E, F)$ pour $\mathcal{L}_k(E, F)$ et $\text{End}_k(E)$ pour $\mathcal{L}_k(E)$.

Le *dual* d'un espace vectoriel E sur un corps k est $\mathcal{L}(E, k)$ et est noté E^* . Rappelons qu'étant donnée une application linéaire $f : E \rightarrow F$ on peut définir son application duale $f^* : F^* \rightarrow E^*$ et que $(f \circ g)^* = g^* \circ f^*$ dès que la formule a un sens.

Étant donné un espace E on définit une application linéaire de E dans le bidual E^{**} par $x \mapsto \hat{x}$, où $\hat{x}$ est donnée par $\hat{x}(\lambda) = \lambda(x)$. Cette application est linéaire et est un isomorphisme si et seulement si E est de dimension finie.

Si α est une *forme linéaire* non nulle l'ensemble des $v \in E$ tels que $\alpha(v) = 0$ est un sous-espace de dimension $n - 1$ de E (si E est de dimension n). C'est un hyperplan de E , α est appelé une équation de l'hyperplan.

Nous introduisons maintenant brièvement la notion d'*espace vectoriel quotient*. Celle-ci devra être relue, par exemple, après l'étude des groupes quotients. Elle ne jouera de rôle que dans les chapitres 5 et 6.

Soit E un espace vectoriel et F un sous-espace. On définit sur E une relation d'équivalence par $v \mathcal{R} w$ si et seulement si $v - w \in F$. Une classe d'équivalence est un sous-ensemble de la forme $v + F = \{v + f \mid f \in F\}$.

On munit l'ensemble des classes d'équivalence d'une structure d'espace vectoriel sur k par les formules :

- $(v + F) + (w + F) = v + w + F$,
- $\lambda(v + F) = \lambda v + F$.

Cet espace est appelé l'espace vectoriel quotient de E par F et est noté E/F .

On vérifie que l'application $v \mapsto v + F$ de E vers E/F est une surjection linéaire.

Enfin, si on restreint cette application à un supplémentaire F' quelconque de F dans E , l'application de F' dans E/F est un isomorphisme. En particulier l'image d'une base de F' est une base de E/F et la dimension de E est la somme de la dimension de F' et de celle de E/F .

Le sous-ensemble de $\mathcal{L}(E)$ constitué par les applications bijectives est un groupe pour la composition. Entre autres, l'application réciproque d'une application linéaire bijective est elle-même linéaire. C'est le *groupe linéaire* de E et il est noté $\text{GL}(E)$.

6. Matrices, déterminants, systèmes linéaires

Soient E et F deux espaces vectoriels sur un corps k . Supposons données une base (e_i) , $i = 1, \dots, m$ de E et une base (f_j) , $j = 1, \dots, n$ de F . Soit f une application linéaire et soit

$$f(e_i) = \sum_{j=1}^n a_{j,i} f_j.$$

La matrice $A = (a_{\ell,h})$ de f dans les bases (e_i) et (f_j) est le tableau d'éléments de k à n lignes et m colonnes dont le terme situé sur la ℓ -ième ligne et la h -ième colonne est $a_{\ell,h}$.

L'ensemble des matrices à coefficients dans k à n lignes et m colonnes est un espace vectoriel sur k de dimension mn , isomorphe à $\mathcal{L}_k(E, F)$. Il est noté $M_{n,m}(k)$. Si $m = n$ on notera $M_n(k)$. À la composée de deux applications linéaires $f : E \rightarrow F$, $g : F \rightarrow G$ correspond le produit des matrices. Nous renvoyons à [LM] et [LA] pour les détails. En fait, $M_n(k)$ est un anneau. Les matrices inversibles constituent un groupe pour la multiplication noté $GL_n(k)$. Il est évidemment isomorphe à $GL(E)$.

Rappelons la formule du changement de base. Soient (e_i) et (e'_i) deux bases de E et soient (f_j) et (f'_j) deux bases de F . Soit P la matrice de passage de (e_i) à (e'_i) , c'est-à-dire la matrice dont les vecteurs colonnes sont les coefficients des e'_k dans la base (e_i) . Et soit Q la matrice de passage de (f_j) à (f'_j) . Soit A la matrice d'une application linéaire f par rapport aux bases (e_i) et (f_j) , et soit A' la matrice de f par rapport aux bases (e'_i) et (f'_j) . On a :

$$A' = Q^{-1}AP.$$

Pour terminer rappelons que le rang d'une matrice est égal à la dimension du sous-espace image de l'application linéaire associée. C'est aussi la dimension de la plus grande sous-matrice carrée extraite inversible. C'est-à-dire la dimension de la plus grande matrice extraite de déterminant non nul.

La formule pour le déterminant d'une matrice (n, n) , $A = (a_{i,j})$, est :

$$\det(A) = |a_{i,j}| = \sum_{s \in \mathfrak{S}_n} (-1)^{\text{sgn}(s)} \prod_i a_{i,s(i)}.$$

Ici $\mathfrak{S}_n$ est le groupe symétrique et sgn est la signature. Encore une fois on renvoie à [LA], [LM] pour les détails ainsi que pour les formules de Cramer donnant l'inverse d'une matrice.

Nous invitons le lecteur à relire toutes les propriétés du déterminant liées aux manipulations des lignes et des colonnes. Elles s'étendent immédiatement quand les coefficients sont dans un corps commutatif quelconque ou un anneau commutatif quelconque.

Nous concluons en donnant une partie du théorème de Rouché et Fontené :

Théorème. Soit un système de n équations linéaires non homogènes à p inconnues. Soit A la matrice de taille (p, n) du système, soit k le rang de A . Alors si $k = n$ le système admet une solution et l'ensemble des solutions un sous-espace affine de dimension $p - n$.

Chapitre 1

Groupes

Groupes quotients

Groupes abéliens

de type fini

Dans ce chapitre, on commence par rappeler les notions de base de la théorie des groupes : sous-groupes, homomorphismes, noyaux, images, intersections, produits. Le groupe des congruences modulo n est redéfini et étudié en détail. Puis on introduit la notion de système générateur d'un groupe, de groupe monogène et d'ordre d'un élément.

La deuxième section est consacrée à l'étude des relations d'équivalence associée à un sous-groupe, puis à la notion de sous-groupe distingué et à celle de groupe quotient. On classe alors les groupes d'ordre p où p est un nombre premier. Le produit semi-direct de deux groupes est étudié en détail.

Dans la troisième section on étudie la décomposition primaire des groupes abéliens finis, en liaison avec le « lemme Chinois ».

Les deux dernières sections sont elles consacrées à la structure des groupes abéliens de type fini. Il s'agit là de théorèmes plus difficiles mais il est souhaitable que les énoncés en soient connus en troisième année de LICENCE.

1. Groupes, générateurs, groupes monogènes

Cette section commence donc par un rappel des définitions de base :

Définition 1

Un *groupe* est un ensemble G muni d'une *loi interne*, i.e. d'une application $\mu : G \times G \rightarrow G$, qui satisfait aux conditions suivantes :

il existe un élément $e \in G$ tel que pour tout $g \in G$

$$\mu(g, e) = \mu(e, g) = g,$$

pour tout $g \in G$ il existe un élément $g' \in G$ tel que

$$\mu(g, g') = \mu(g', g) = e,$$

pour tous $g, h, k \in G$ on a

$$\mu(g, \mu(h, k)) = \mu(\mu(g, h), k).$$

Dans la suite un groupe sera noté (G, μ) . S'il n'y a pas lieu d'avoir une notation particulière pour la loi interne celle-ci sera omise dans la notation, le groupe sera seulement noté G et $\mu(g, h)$ sera simplement noté gh .

L'élément e est appelé l'élément neutre. L'élément g' est appelé l'inverse de g et est noté g^{-1} . La dernière condition s'appelle l'*associativité* de la loi interne.

La première condition implique l'unicité de e . En effet, soient deux éléments e et e' satisfaisant à la première condition. On a $ee' = e$ et $ee' = e'$ donc $e = e'$. La dernière condition implique elle l'unicité de g^{-1} . En effet soient g_0 et g_1 satisfaisant à la deuxième condition, on a $g_0 = g_0(gg_1) = (g_0g)g_1 = g_1$.

Définition 2

Un groupe (G, μ) est dit *abélien* ou *commutatif* si pour tous x et y éléments de G on a $\mu(x, y) = \mu(y, x)$.

Dans la suite, par convention quand on utilisera la terminologie « abélien » la loi interne sera notée $+$, et l'élément neutre 0 . Sinon, que le groupe soit commutatif ou non elle sera notée multiplicativement et l'élément neutre sera noté 1 ou e .

Dans le cas abélien, si n est un entier positif, la somme $\underbrace{x + \cdots + x}_{n \text{ fois}}$ sera notée nx .

Son élément opposé sera noté $-nx$. Dans le cas éventuellement non commutatif, les mêmes éléments seront notés x^n et x^{-n} . On a donc $x^n = \underbrace{x \times \cdots \times x}_{n \text{ fois}}$.

- Le premier exemple de groupe est celui des entiers naturels $\mathbb{Z}$ muni de l'addition,

- de même, l'ensemble $\mathbb{Q}$ des nombres rationnels, celui $\mathbb{R}$ des nombres réels, celui $\mathbb{C}$ des nombres complexes, tous munis de l'addition, sont des groupes,
- un ensemble à deux éléments notés 0 et 1 est muni d'une structure de groupe en déclarant que 0 est l'élément neutre et que 1 est son propre inverse. La table d'un groupe G est le tableau dont les colonnes et les lignes sont indexées par les éléments de G . La valeur $\mu(g, h)$ est portée à l'intersection de la ligne indexée par g et de la colonne indexée par h . La table du groupe précédent est :

μ	0	1
0	0	1
1	1	0

Il s'agit là évidemment du groupe $\mathbb{Z}/2\mathbb{Z}$.

Définition 3

Un *sous-groupe* d'un groupe (G, μ) est un sous-ensemble non vide H de G tel que l'application μ restreinte à $H \times H$ soit à valeurs dans H , et détermine une loi de groupe sur H .

La première condition nous dit que μ définit une loi interne pour H . On en conclut que μ est associative. En effet, les relations à satisfaire le sont dans G , donc par restriction dans H . La seconde condition est réalisée dès que tout élément de H a son inverse dans H . Ces conditions sont équivalentes aux suivantes :

- l'élément neutre appartient à H
- pour tous x et y dans H $\mu(x, y^{-1})$ appartient à H .

En effet, partons du second groupe de conditions. Si on remplace x par e dans la seconde, on voit que si $y \in H$ alors $y^{-1} \in H$. Si on y remplace y^{-1} par y , on obtient la première puis la seconde du premier groupe.

Inversement, si H est un sous-groupe, les conditions du second groupe sont satisfaites.

À partir de groupes ou de sous-groupes, un certain nombre de constructions permettent de construire de nouveaux groupes, en voici quelques unes :

Proposition 1. Soit H_i $i \in I$ une famille de sous-groupes de G , alors $\cap_i H_i$ est un sous-groupe de G .

On applique, par exemple, les critères suivants la définition 3. Le premier est immédiat ; pour le second on observe que si x et y sont dans H_i pour tout i alors $\mu(x, y^{-1}) \in H_i$ pour tout i , et donc $\mu(x, y^{-1}) \in \cap_i H_i$.

Produits, produits restreints, et sommes directes de groupes

Étant donnés une famille de groupes G_i , l'indice i décrivant un ensemble I considérons le produit

$$\prod_{i \in I} G_i$$

on en note les éléments par $(g_i)_{i \in I}$. On définit un produit par la formule :

$$(g_i)_{i \in I} \times (g'_i)_{i \in I} = (g_i g'_i)_{i \in I}$$

où g_i, g'_i sont éléments de G_i .

On considère aussi le produit restreint. On le définit comme étant le sous-ensemble du produit

$$\prod_{i \in I} G_i$$

constitué par les éléments $(g_1, \dots, g_i, \dots)$ tels que pour tout i , sauf un nombre fini, l'élément g_i est égal à l'élément neutre de G_i . C'est un sous-groupe du produit. Si l'ensemble d'indice I est fini, produit et produit restreint coïncident.

Si les groupes sont abéliens, le produit restreint est appelé la somme directe et est noté :

$$\bigoplus_{i \in I} G_i.$$

Les groupes sont importants tant pour eux-mêmes que pour les relations qu'ils ont entre eux. Celles-ci s'expriment par les applications d'un groupe dans un autre, compatibles aux lois à la source, et au but, plus précisément :

Définition 4

On appelle *homomorphisme* de (G, μ) dans (G', μ') une application ϕ de G dans G' telle que pour tous x et y dans G on ait

$$\phi(\mu(x, y)) = \mu'(\phi(x), \phi(y)).$$

On observe que l'image de l'élément neutre de G est l'élément neutre de G' . En effet $\phi(x) = \phi(\mu(x, e)) = \mu'(\phi(x), \phi(e))$ pour tout $x \in G$, en multipliant à gauche par l'inverse de $\phi(x)$, on obtient $\phi(e) = e'$ (avec les notations évidentes). Un homomorphisme bijectif est appelé un *isomorphisme*. On notera que l'application réciproque d'un homomorphisme bijectif est un homomorphisme. Si $G = G'$, un homomorphisme est appelé un *endomorphisme*, un isomorphisme est appelé un *automorphisme*. Si ϕ est un homomorphisme de G dans H et ψ un homomorphisme de H dans K , alors $\psi \circ \phi$ est un homomorphisme de G dans K .

Ces définitions donnent de nouveaux moyens de construire des groupes :

Définition 5

L'image d'un groupe G par un homomorphisme ϕ dans un groupe H est l'ensemble des éléments de la forme $\phi(x)$, lorsque x décrit G , on la note $\text{Im}(\phi)$.

Proposition 2. L'image d'un homomorphisme de G dans H est un sous-groupe de H .

Démonstration. On applique les critères suivant la définition 3. Soient $\phi(x)$ et $\phi(y)$ des éléments de $\text{Im}(\phi)$, alors $\phi(y)^{-1} = \phi(y^{-1})$ et $\phi(x)\phi(y^{-1}) = \phi(xy^{-1}) \in \text{Im}(\phi)$, le résultat suit. ■

Définition 6

On appelle *noyau* d'un homomorphisme

$$\phi : G \rightarrow G'$$

l'ensemble des éléments x de G tels que $\phi(x) = e'$, e' désignant l'élément neutre de G' . On le note $\text{Ker}(\phi)$.

Proposition 3. Le noyau d'un homomorphisme $\phi : G \rightarrow G'$ est un sous-groupe de G' .

Démonstration. On applique les critères suivant la définition 3. Soient x et y des éléments de $\text{Ker}(\phi)$, on a alors

$$\phi(x) = \phi(y)^{-1} = e'$$

et

$$\phi(xy^{-1}) = \phi(x)\phi(y)^{-1} = e' ,$$

le résultat suit. ■

On vérifiera qu'un homomorphisme est injectif si et seulement si son noyau est réduit à l'élément neutre.

Les groupes apparaissent en mathématiques dans une grande variété de situation. Les exemples qui suivent sont fondamentaux en algèbre, en géométrie, en combinatoire.

Classes de congruences modulo n

Définition 7

On dit que deux entiers k et l sont congruents modulo n si et seulement si la différence $k - l$ est divisible par n .

La *relation de congruence* est une relation d'équivalence sur les entiers. La classe d'équivalence d'un entier k est l'ensemble des entiers de la forme $k + an$, $a \in \mathbb{Z}$ quelconque, on la notera $\{k + n\mathbb{Z}\}$. Ces classes d'équivalence constituent une partition de $\mathbb{Z}$.

On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes d'équivalence dans $\mathbb{Z}$ pour la relation de congruence modulo n .

Définissons une loi interne, notée $+$, sur cet ensemble par :

$$\{k + n\mathbb{Z}\} + \{k' + n\mathbb{Z}\} = \{k + k' + n\mathbb{Z}\}.$$

Proposition 4. *Cette loi définit sur $\mathbb{Z}/n\mathbb{Z}$ une structure de groupe. De plus l'application $p : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ qui à un entier associe sa classe d'équivalence est un homomorphisme.*

Étant donné un entier a on notera $\bar{a}$ sa classe de congruence modulo n .

Groupes de permutations

Soit S un ensemble, l'ensemble $\mathcal{P}(S)$ des applications bijectives, ou permutations, de S dans lui-même est un groupe pour la composition des applications. La loi est interne car la composée de deux bijections est une bijection, elle est associative car la composition des applications l'est. L'élément neutre est l'application identité de X , notée Id_X définie par $\text{Id}_X(x) = x$ pour tout $x \in X$. L'inverse est l'application réciproque.

Voici un exemple particulier de sous-groupe de permutations. Soit E un espace vectoriel sur un corps K . L'ensemble des applications linéaires inversibles forme, pour la composition des applications, un groupe noté $\text{GL}(E)$. C'est un sous-groupe de $\mathcal{P}(E)$. Si E est de dimension finie n , et si on a choisi une base, le groupe $\text{GL}(E)$ est isomorphe au groupe $\text{GL}_n(K)$ des matrices inversibles (n, n) à coefficients dans K .

Éléments inversibles dans un anneau

Étant donné un anneau A , l'ensemble A^* des éléments inversibles pour la multiplication est un groupe pour cette loi. Par exemple $\mathbb{Z}^* = \{+1, -1\}$, $\mathbb{K}^* = \mathbb{K} - \{0\}$ si $\mathbb{K}$ est un corps. De même l'ensemble des nombres complexes de module 1 est un groupe pour la multiplication, cela résulte de la relation $|zz'| = |z||z'|$. On le note U .

Donnons maintenant deux exemples de sous-groupes, dont le premier en particulier est fondamental. Il en sera donné une généralisation dans la section 4.

Sous-groupes de $\mathbb{Z}$

Proposition 5. Soit H un sous-groupe de $\mathbb{Z}$. Alors, soit H est trivial c'est-à-dire réduit à l'élément neutre, soit H est de la forme $a\mathbb{Z} = \{au \mid u \in \mathbb{Z}\}$, pour un entier a non nul, bien défini au signe près.

Démonstration. On suppose que H n'est pas réduit à $\{0\}$. Le groupe H contient forcément des éléments positifs non nuls, en effet il contient des éléments non nuls, et étant un sous-groupe il contient leurs opposés. Soit a le plus petit élément positif non nul de H .

Montrons que tout élément de H est multiple de a . Il suffit de le montrer pour des éléments positifs. Soit $h \in H$, $h > 0$, faisons la division euclidienne de h par a : $h = aq + r$, $0 \leq r < a$. L'élément $r = h - aq$ appartient à H , en effet h et aq sont éléments de H . Comme a est le plus petit élément positif non nul de H on a $r = 0$, donc $H = a\mathbb{Z}$.

Si a est non nul, l'application

$$\begin{aligned}\mathbb{Z} &\longrightarrow a\mathbb{Z}, \\ k &\longmapsto ak\end{aligned}$$

est un isomorphisme, on en déduit donc que tout sous-groupe de $\mathbb{Z}$ est soit trivial, soit isomorphe à $\mathbb{Z}$. ■

Automorphismes, automorphismes intérieurs

Étant donné un groupe G , l'ensemble des automorphismes de G forme un groupe pour la loi de composition des applications. Il sera noté $\text{Aut}(G)$. Si on considère un élément $x \in G$ l'application

$$i_x : G \rightarrow G, \quad g \longmapsto xgx^{-1}$$

est un automorphisme de G . En effet pour tous $a, b \in G$ on a :

$$i_x(ab) = xabx^{-1} = xax^{-1}xbx^{-1} = i_x(a)i_x(b),$$

notons aussi que $(i_x)^{-1} = i_{x^{-1}}$. On dit que i_x est un automorphisme intérieur de G . Le sous-ensemble $\text{Int}(G) \subset \text{Aut}(G)$ constitué par les automorphismes intérieurs est un sous-groupe. En effet on observe que l'on a $i_x \circ i_y = i_{xy^{-1}}$.

L'application de G dans $\text{Aut}(G)$ qui à x associe i_x est un homomorphisme. Cela résulte de la formule $i_x \circ i_y = i_{xy}$. L'image est le sous-groupe $\text{Int}(G)$.

Un élément x est dans le noyau de cet homomorphisme si pour tout $y \in G$, $xyx^{-1} = x$, soit $xy = yx$. Ce qui veut dire que x commute à tout élément de G . Ce sous-groupe est appelé le *centre* de G , et est noté $Z(G)$, par définition il est commutatif.

L'exemple qui suit doit, de préférence, être lu après le chapitre 3. Il s'agit de la description des automorphismes du groupe des classes de congruence modulo n . Il est déterminé par :

Il y a un isomorphisme entre le groupe $\text{Aut}(\mathbb{Z}/n\mathbb{Z})$ et le groupe $(\mathbb{Z}/n\mathbb{Z})^*$

Notons $\text{End}(\mathbb{Z}/n\mathbb{Z})$ l'ensemble des endomorphismes de $\mathbb{Z}/n\mathbb{Z}$. C'est un groupe, car la somme $(f + g)$ de deux endomorphismes f et g donnée par $(f + g)(u) = f(u) + g(u)$ est un endomorphisme. Anticipons les résultats sur la structure d'anneau de $\mathbb{Z}/n\mathbb{Z}$ du chapitre 3 auquel le lecteur est invité à se reporter. On observe qu'un homomorphisme $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ est déterminé par l'image de la classe de congruence de 1. En effet si on a $f(\bar{1}) = \bar{a}$ on a $f(\bar{x}) = \bar{a} \bar{x}$. Autrement dit l'application :

$$\begin{array}{ccc} \text{End}(\mathbb{Z}/n\mathbb{Z}) & \longrightarrow & \mathbb{Z}/n\mathbb{Z}, \\ f & \longmapsto & f(\bar{1}) \end{array}$$

est une bijection. Plus précisément c'est un isomorphisme de groupes.

Si f est un automorphisme montrons que $\bar{a}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$. Soit g l'automorphisme réciproque de f . On a $g \circ f(\bar{1}) = \bar{1}$, soit $g(\bar{a}) = g(\bar{1}) \bar{a} = \bar{1}$. Ce qui veut dire que $\bar{a}$ est inversible. Ceci démontre en fait que l'application

$$\begin{array}{ccc} \text{Aut}(\mathbb{Z}/n\mathbb{Z}) & \longrightarrow & (\mathbb{Z}/n\mathbb{Z})^*, \\ f & \longmapsto & f(\bar{1}) \end{array}$$

est un isomorphisme de groupes. La loi à gauche étant la composition des applications, celle à droite la multiplication.

Systèmes de générateurs

Nous définissons maintenant la notion de système de générateurs pour un groupe.

Définition 8

Étant donné un groupe G et un sous-ensemble P de G , on appelle *sous-groupe engendré* par P l'intersection de tous les sous-groupes de G contenant P .

L'ensemble ainsi défini est un sous-groupe (proposition 1). Par définition, tout sous-groupe de G contenant P contient ce sous-groupe.

Soit $P = \{x_1, x_2, \dots, x_n\}$ une famille finie d'éléments de G . On notera

$$\langle x_1, x_2, \dots, x_n \rangle$$

le sous-groupe engendré par P . Les éléments x_i sont appelés les *générateurs*. L'ensemble $P = \{x_1, x_2, \dots, x_n\}$ est un *système de générateurs*. Il n'y a évidemment pas unicité des systèmes de générateurs.

Définition 9

On dira que G est engendré par n éléments s'il existe n éléments $x_1, \dots, x_n$ dans G tels que le sous-groupe engendré par ces éléments soit G lui-même.

Considérons par exemple le groupe abélien $\mathbb{Z}^n$. Soit un entier i tel que $1 \leq i \leq n$ et soit $e_i \in \mathbb{Z}^n$ l'élément $(0, 0, \dots, 1, 0, \dots, 0)$, l'élément 1 étant situé en i -ième position. Les éléments $e_1, e_2, \dots, e_n$ forment un système générateur de $\mathbb{Z}^n$ car $(\lambda_1, \dots, \lambda_n) = \lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n$.

En fait on peut caractériser les groupes abéliens engendrés par n éléments à partir de $\mathbb{Z}^n$.

Proposition 6. *Soit G un groupe abélien. Les deux conditions suivantes sont équivalentes :*

- G est engendré par n éléments,
- il existe un homomorphisme surjectif de $\mathbb{Z}^n$ sur G .

Démonstration. Soit $\{x_1, \dots, x_n\}$ un système de n générateurs. Soit ϕ l'application :

$$\begin{aligned} \mathbb{Z}^n &\longrightarrow G, \\ (\lambda_1, \dots, \lambda_n) &\longmapsto \lambda_1 x_1 + \dots + \lambda_n x_n. \end{aligned}$$

Comme G est abélien c'est un homomorphisme de groupes. En effet, on a $\sum_i \lambda_i x_i + \sum_i \mu_i x_i = \sum_i (\lambda_i + \mu_i) x_i$. L'image de l'homomorphisme est un sous-groupe, c'est le groupe G tout entier puisqu'il en contient un système de générateurs, donc ϕ est surjectif.

Inversement, soit ϕ un homomorphisme surjectif de $\mathbb{Z}^n$ sur G , et soit $\{e_1, \dots, e_n\}$ le système de générateurs de $\mathbb{Z}^n$ introduit plus haut. Notons d'abord que la surjectivité de ϕ implique que G est abélien. En effet, soient $x, y \in G$, il existe $x', y' \in \mathbb{Z}^n$ tels que $\phi(x') = x$ et $\phi(y') = y$. Donc

$$xy = \phi(x')\phi(y') = \phi(x'y') = \phi(y'x') = \phi(y')\phi(x') = yx.$$

Nous n'avions donc pas besoin de l'hypothèse G abélien pour cette partie de l'énoncé. Prenons maintenant la notation additive pour G . Posons $x_i = \phi(e_i)$, on montre que $\{x_1, \dots, x_n\}$ est un système de générateurs pour G , soit que tout élément de G peut s'écrire sous la forme $\sum_i \lambda_i x_i$, pour $\lambda_1, \dots, \lambda_n \in \mathbb{Z}^n$. En effet l'ensemble des éléments de la forme précédente est un sous-groupe de G ainsi qu'on l'a vu plus haut. Comme c'est l'image de ϕ et que ϕ est surjective, il est égal à G . Le résultat suit. ■

Le cas des groupes engendrés par un élément mérite une attention particulière. On remarque que ces groupes sont tous commutatifs. En effet si x désigne le générateur le sous-ensemble $\{x^n \mid n \in \mathbb{Z}\}$ du groupe est un sous-groupe qui s'identifie au groupe par hypothèse. Il est clair que x^m commute à x^n pour tous m et n .

Définition 10

Un groupe est dit monogène s'il admet un système générateur réduit à un élément. Un *groupe monogène fini* est appelé un *groupe cyclique*.

Les groupes $\mathbb{Z}^n$, avec $n > 1$, ne sont pas monogènes.

Introduisons maintenant :

Définition 11

Soit G un groupe et soit x un élément de G . On appelle *ordre* de x , s'il existe, le plus petit entier positif non nul tel que $x^n = e$. Si un tel entier n'existe pas on dit que l'élément est d'ordre infini. Un élément d'ordre fini est aussi dit de *torsion*.

Par exemple dans $\mathbb{Z}/n\mathbb{Z}$ la classe de 1 est d'ordre n . Dans le groupe multiplicatif U des nombres complexes de module 1 l'élément i est d'ordre 4, l'élément j est d'ordre 3, l'élément $-j$ est d'ordre 6, et l'élément ij est d'ordre 12.

Proposition 7. Soient G un groupe et x un élément de G d'ordre n . Alors tout entier k non nul tel que $x^k = e$ est divisible par n .

Démonstration. Soit k tel que $x^k = e$. Faisons la division euclidienne de k par n : $k = an + r$, $0 \leq r < n$. On a donc $e = x^k = x^{an+r} = (x^n)^a x^r = x^r$, donc $x^r = e$. Comme $0 \leq r < n$ on a par définition de n $r = 0$. ■

Définition 12

Le cardinal ou nombre d'éléments d'un groupe G est aussi appelé son *ordre*. On le note $\#G$.

Nous allons décrire les groupes monogènes.

Observons d'abord qu'un groupe monogène infini est isomorphe à $\mathbb{Z}$. Si x désigne un générateur l'isomorphisme est donné par l'application qui à n associe x^n . Cette application est surjective car x est un générateur, injective car x est d'ordre infini, c'est un isomorphisme.

Proposition 8. Un groupe cyclique G d'ordre n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Démonstration. Soit x un générateur de G , n son ordre. Ainsi qu'on l'a vu plus haut, G est abélien et faisons choix de la notation additive pour la loi de G . Considérons l'homomorphisme ϕ de $\mathbb{Z}$ dans G qui, à l'entier λ associe l'élément λx de G .

Puisque n est l'ordre de x , le noyau de ϕ est $n\mathbb{Z}$. Deux éléments ont même image si et seulement si ils sont congrus modulo n . Une classe de congruence modulo n est donc envoyée sur un élément de G . Soit $\bar{\phi}$ l'application ainsi déterminée. Deux classes distinctes sont envoyées sur des éléments distincts. En effet, si on a $\phi(\bar{k}) = \phi(\bar{l})$, on a $\phi(k - l) = 0$. Donc $k - l$ est dans le noyau de ϕ , donc k et l sont congrus modulo n . En conclusion, l'ensemble des classes de congruence modulo n s'envoie bijectivement par $\bar{\phi}$ sur le groupe G .

L'application ainsi définie est un homomorphisme de groupes. En effet l'équation

$$\bar{\phi}(\bar{a} + \bar{b}) = \bar{\phi}(\bar{a}) + \bar{\phi}(\bar{b}).$$

résulte de $(a + b)x = ax + bx$. ■

Corollaire. *L'ordre d'un élément est égal au cardinal du sous-groupe qu'il engendre.*

Démonstration. Supposons que l'ordre de l'élément est fini, soit n . Alors la démonstration précédente montre que le cardinal de $\langle x \rangle$ est n .

Si l'ordre de x est infini le groupe engendré par x , soit $\langle x \rangle$, est isomorphe à $\mathbb{Z}$ d'où le résultat. Pour montrer cela, il suffit de reprendre la démonstration précédente, en observant que le noyau de ϕ est trivial, i.e. réduit à $\{0\}$. ■

Voici un exemple de calcul de l'ordre d'un élément, le corollaire sera utilisé dans la dernière section du chapitre.

Proposition 9. *Soit $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$. L'ordre de $\bar{a}$ est égal à :*

$$\frac{n}{\text{pgcd}(a, n)}.$$

Démonstration. L'ordre de $\bar{a}$ divise $\frac{n}{\text{pgcd}(a, n)}$, en effet

$$n \mid \frac{n}{\text{pgcd}(a, n)} a$$

l'entier $a \frac{n}{\text{pgcd}(a, n)}$ est donc nul modulo n . Inversement, si k est l'ordre de $\bar{a}$ l'entier n divise ka . Écrivons a sous la forme $a' \text{pgcd}(a, n)$, en particulier a' est premier à n . L'entier $\frac{n}{\text{pgcd}(a, n)}$ divise ka' , donc divise k , d'où le résultat. ■

Corollaire. *Soit $x \in \mathbb{Z}/n\mathbb{Z}$, et soit d un diviseur de n . Il existe $y \in \mathbb{Z}/n\mathbb{Z}$ tel que $dy = x$ si et seulement si $\frac{n}{d}x = 0$.*

Démonstration. Soit $x = \bar{a}$, l'ordre de x est $\frac{n}{\text{pgcd}(a, n)}$. Supposons que $\frac{n}{d}x = 0$, ce qui signifie que

$$\frac{n}{\text{pgcd}(a, n)} \mid \frac{n}{d},$$

soit

$$d \mid \text{pgcd}(a, n).$$

On a donc :

$$x = d \left(\frac{\text{pgcd}(a, n)}{d} \times \left(\frac{a}{\text{pgcd}(a, n)} \right) \right).$$

L'implication inverse est immédiate. ■

Proposition 10. Soit H un sous-groupe de $\mathbb{Z}/n\mathbb{Z}$. Alors, soit H est trivial, c'est-à-dire réduit à l'élément neutre, soit H est isomorphe à $\mathbb{Z}/m\mathbb{Z}$ pour un entier $m > 1$ bien déterminé et divisant n .

Démonstration. Soit p la surjection canonique de $\mathbb{Z}$ sur $\mathbb{Z}/n\mathbb{Z}$, alors $p^{-1}(H)$ est un sous-groupe de $\mathbb{Z}$. Il est de la forme $a\mathbb{Z}$ avec a non nul si H est non trivial. La proposition précédente montre que le sous-groupe engendré par $\bar{a}$, qui est égal par construction à H puisque p est surjectif, est

$$\mathbb{Z} / \frac{n}{\text{pgcd}(a, n)} \mathbb{Z}.$$

■

Voici une reformulation :

Corollaire. Soient, n et d des entiers. Le sous-groupe $d(\mathbb{Z}/n\mathbb{Z})$ de $\mathbb{Z}/n\mathbb{Z}$, est isomorphe à $\mathbb{Z} / \frac{n}{\text{pgcd}(d, n)} \mathbb{Z}$.

2. Relations d'équivalence dans un groupe, groupes quotients

L'une des constructions les plus importantes pour étudier la structure des groupes est celle de *groupe quotient*. La première étape est d'étudier certaines relations d'équivalences, sur un groupe, associées à un sous-groupe. Puis on se pose la question de savoir si l'ensemble des classes d'équivalence peut être muni d'une loi compatible avec celle du groupe. Cette section est très importante pour la troisième année de LICENCE.

Soient G un groupe et H un sous-groupe.

Définition 1

On appelle relation à gauche (resp. à droite) associée au sous-groupe H la relation définie par

$$x \sim y \iff x^{-1}y \in H,$$

resp.

$$x \sim' y \iff yx^{-1} \in H.$$

Proposition 1. *Ces relations sont des relations d'équivalence.*

Démonstration. Cela résulte du fait que H est un sous-groupe. Prenons par exemple le cas de la relation à gauche. La réflexivité s'écrit $x^{-1}x \in H$, soit $e \in H$. La symétrie s'écrit $x^{-1}y \in H \iff y^{-1}x \in H$, mais $x^{-1}y$ est l'inverse de $y^{-1}x$, si l'un des éléments appartient au sous-groupe, l'autre aussi. Enfin la transitivité suit de ce que si on a $x^{-1}y \in H$ et $y^{-1}z \in H$ alors $x^{-1}z \in H$. ■

Proposition 2. *La classe d'équivalence d'un élément $x \in H$ pour la relation à gauche (resp. la relation à droite) est l'ensemble $xH = \{xh \mid h \in H\}$ (resp. $Gx = \{gx \mid g \in G\}$). Ces classes sont appelées respectivement classes à gauche et classes à droite selon H . Elles ont toutes même cardinal, à savoir celui de H .*

Ceci résulte du fait que l'application $h \mapsto xh$ de H dans xH est bijective, elle admet $h \mapsto x^{-1}h$ pour application réciproque.

L'ensemble des classes à gauche selon H est noté G/H , respectivement à droite est noté $H \backslash G$.

Voici l'application la plus célèbre de cette construction.

Théorème 1 (Lagrange). *L'ordre d'un sous-groupe divise l'ordre du groupe.*

Démonstration. Les classes à gauche (ou à droite) constituent une partition de G . Supposons que G ait un nombre fini d'éléments. Il y a alors un nombre fini de classes d'équivalence, soit $\{x_1, \dots, x_n\}$ un système de représentants de ces classes, c'est-à-dire un sous-ensemble de G contenant un élément de chaque classe et un seul. L'ordre de G est alors égal à $n(\#H)$. ■

Corollaire 1. *Soit G un groupe fini. L'ordre d'un élément divise l'ordre du groupe.*

Démonstration. Ceci résulte du théorème de Lagrange et du corollaire 1 de la première section : l'ordre du sous-groupe engendré par un élément x divise l'ordre du groupe, mais c'est aussi l'ordre de l'élément. ■

Ce corollaire permet de classer les groupes dont l'ordre est un nombre premier :

Corollaire 2. *Tout groupe d'ordre p , où p est un nombre premier est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.*

Démonstration. Soit G un groupe d'ordre p . Tout élément est d'ordre p ou 1. Seul l'élément neutre est d'ordre 1, il y a donc forcément un élément d'ordre p . On en choisit un. Le sous-groupe qu'il engendre est isomorphe à $\mathbb{Z}/p\mathbb{Z}$ et à G d'où le résultat. ■

Soit maintenant ϕ un homomorphisme d'un groupe G dans un groupe H . Soit $x \in \text{Ker}(\phi)$, soit $y \in G$. L'élément xyx^{-1} est dans $\text{Ker}(\phi)$, en effet :

$$\phi(yxy^{-1}) = \phi(y)\phi(x)\phi(y^{-1}) = \phi(y)\phi(y^{-1}) = \phi(yy^{-1}) = e.$$

Ceci justifie d'introduire la :

Définition 2

Un sous-groupe K d'un groupe G est dit distingué dans G dès que pour tout $a \in G$ l'ensemble $aKa^{-1} = \{aka^{-1} \mid k \in K\}$ est contenu dans K .

Proposition 3. *Le noyau d'un homomorphisme est un sous-groupe distingué.*

Démonstration. C'est ce qu'on a montré avant la définition. ■

En fait si K est distingué dans G , on a $aKa^{-1} = K$ pour tout $a \in G$, en effet de l'inclusion $a^{-1}Ka \subset K$, on conclut en multipliant à gauche par a et à droite par a^{-1} que $K \subset aKa^{-1}$, pour tout a élément de G .

Voici deux exemples classiques de sous-groupes distingués. D'abord, dans un groupe abélien, tous les sous-groupes sont distingués.

Le sous-groupe $\text{Int}(G)$ de $\text{Aut}(G)$ est distingué.

Démonstration. Soient $\psi \in \text{Aut}(G)$, $a \in G$, et $i_a \in \text{Int}(G)$. Alors

$$\psi \circ i_a \circ \psi^{-1}(x) = \psi(a\psi^{-1}(x)a^{-1}) = \psi(a)x\psi(a)^{-1} = i_{\psi(a)}(x).$$

Donc on a $\psi \circ i_a \circ \psi^{-1} = i_{\psi(a)}$, ce qui démontre le résultat. ■

Nous pouvons maintenant introduire la notion de groupe quotient :

Théorème 2. *Soit G un groupe et soit H un sous-groupe distingué. Il existe sur l'ensemble des classes à gauche G/H une structure de groupe et une seule telle que l'application canonique $p : G \rightarrow G/H$ soit un homomorphisme.*

Démonstration. On observe d'abord que si H est distingué dans G les classes à droite et les classes à gauche selon H coïncident. On a $aH = Ha$ pour tout $a \in G$ car $aH = a(a^{-1}Ha) = Ha$. Le théorème peut donc s'énoncer indifféremment avec les classes à gauche ou les classes à droite.

Considérons maintenant le « produit » de deux classes à gauche : $aHbH = \{axy | x, y \in H\}$. On a $aHbH = ab(b^{-1}Hb)H$; comme $b^{-1}Hb = H$ puisque H est distingué, on obtient en fin de compte $aHbH = abH$. Définissons une loi interne μ sur l'ensemble des classes à gauche comme suit : soient H_0 et H_1 deux classes, l'ensemble

$$H_0H_1 = \{kl | k \in H_0, l \in H_1\}$$

est une classe à gauche ainsi qu'on vient de le voir, on détermine ainsi une loi interne sur l'ensemble des classes.

Cette loi admet H pour élément neutre. Elle est associative :

$$\begin{aligned}(H_0H_1)H_2 &= \{km | k \in H_0H_1, m \in H_2\} = \dots \\ \dots &= \{klm | k \in H_0, l \in H_1, m \in H_2\} = H_0(H_1H_2).\end{aligned}$$

Enfin $a^{-1}H$ est l'inverse de aH .

Par ailleurs la formule

$$aHbH = abH$$

montre que l'application canonique $p : G \rightarrow G/H$, $a \mapsto aH$ est un homomorphisme de groupes. ■

Par exemple le groupe $\mathbb{Z}/n\mathbb{Z}$ est le quotient du groupe $\mathbb{Z}$ par le sous-groupe $n\mathbb{Z}$.

Proposition 4. Soit ϕ un homomorphisme d'un groupe G dans un groupe G' . Soit H un sous-groupe distingué de G , supposons que pour tout $g \in H$ $\phi(g) = e$. Alors il existe un unique homomorphisme ψ de G/H dans G' tel que $\phi = \psi \circ p$, où p est l'homomorphisme canonique.

Démonstration. On définit ψ par la formule $\psi(gH) = \phi(g)$. Pour vérifier que cette formule a un sens il faut montrer que, si $gH = g'H$, on a $\phi(g) = \phi(g')$. Ceci résulte de ce que $g^{-1}g' \in H$ implique $\phi(g) = \phi(g')$. Il reste à vérifier que ψ est un homomorphisme. On a

$$\psi(gHg'H) = \psi(gg'H) = \phi(gg') = \phi(g)\phi(g') = \psi(gH)\psi(g'H),$$

car H est distingué, le résultat suit.

Pour l'unicité il suffit d'observer que, si on a $\phi = \psi \circ p$, ceci implique que $\phi(g) = \psi \circ p(g) = \psi(gH)$. Donc la formule proposée est la seule possible. ■

Nous allons donner maintenant un certain nombre de conséquences classiques de cette construction : la *décomposition canonique des homomorphismes* et les *théorèmes d'isomorphismes*.

Corollaire. Soit $\phi : G \rightarrow H$ un homomorphisme. Notons p l'application canonique de G dans $G/\text{Ker}(\phi)$, i l'inclusion canonique de $\text{Im}(\phi)$ dans H . Il existe un unique isomorphisme ϕ' de $G/\text{Ker}(\phi)$ dans $\text{Im}(\phi)$ tel que $\phi = i \circ \phi' \circ p$. Autrement dit, on a un diagramme « commutatif » comme suit :

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ \downarrow p & & \uparrow i \\ G/\text{Ker}(\phi) & \xrightarrow{\phi'} & \text{Im}(\phi) \end{array}$$

Démonstration. Prenons dans la démonstration précédente $H = \text{Ker}(\phi)$. Prenons pour ϕ' l'homomorphisme de $G/\text{Ker}(\phi)$ dans $\text{Im}(\phi)$ défini par

$$\phi'(x \text{Ker}(\phi)) = \phi(x),$$

c'est un isomorphisme car il est injectif et surjectif par construction. Puis composons à gauche par l'inclusion i de l'image de ϕ dans H . ■

Théorème 3. Soient G un groupe, H et K deux sous-groupes distingués de G tels que $K \subset H$. Alors le groupe quotient H/K s'identifie à un sous-groupe distingué H' du groupe quotient G/K . De plus le groupe quotient G/H est isomorphe au groupe quotient $(G/K)/H'$.

Par abus de notation H' sera souvent noté H/K .

Démonstration. On considère l'homomorphisme canonique $p : G \rightarrow G/H$. Son noyau contient K qui est distingué, on en déduit un homomorphisme de G/K vers G/H . Cet homomorphisme associe à la classe aK la classe aH , il est surjectif. Son noyau H' consiste en l'ensemble des classes aK telles que aKH est égal H , c'est-à-dire à l'ensemble des classes aK telles que $a \in H$. Cet ensemble est isomorphe au groupe quotient H/K . Ceci démontre le résultat. ■

Théorème 4. Soient G un groupe, H un sous-groupe distingué de G , et K un sous-groupe de G . Alors l'ensemble $HK = \{hk \mid k \in K, h \in H\}$ est un sous-groupe de G . Par ailleurs le groupe quotient HK/H est isomorphe au groupe quotient $K/K \cap H$.

Démonstration. Pour ce qui est de la première assertion, on observe que l'ensemble KH est égal à l'ensemble HK . En effet, soient kh , $k \in K$ et $h \in H$, alors $kh = (khk^{-1})k$ et, comme H est distingué dans G , on a $khk^{-1} = h' \in H$. Donc on a

$kh = h'k$, et donc $KH \subset HK$. L'inclusion inverse se montre de manière analogue et le résultat suit.

Il en résulte que KH est un sous-groupe.

En effet un élément de la forme $khk'h'$ peut s'écrire sous la forme $khtk'$, $t \in H$, à cause de la première inclusion que l'on applique au produit $k'h'$. Puis ce produit s'écrit $kk'\bar{h}$, $\bar{h} \in H$, à cause de la seconde inclusion que l'on applique au produit htk' . Ceci montre que la première condition de la définition 3 de la section 1 est satisfaite. Pour ce qui est de la seconde condition, soit $(kh)^{-1} = h^{-1}k^{-1}$, le second terme s'écrit $k'h'$, $k' \in K$, $h' \in H$ à cause de la seconde inclusion. Ce que l'on cherchait.

Considérons maintenant l'homomorphisme de K dans HK/H qui est la composée de l'inclusion de H dans HK et de la projection canonique de HK sur HK/H . Cet homomorphisme est surjectif et a pour noyau $H \cap K$. Admettant ces deux propriétés pour un instant, le théorème résulte de la factorisation canonique.

Démontrons d'abord la surjectivité. L'homomorphisme étudié associe à la classe $k(K \cap H)$ la classe kH . Un élément de HK/H est une classe de la forme aH , $a \in HK = KH$. Soit $a = kh$, donc $aH = kH$. Il en résulte que la classe aH est l'image de $kK \cap H$. Ce qui démontre la surjectivité de l'homomorphisme considéré.

Soit maintenant un élément dans le noyau, c'est-à-dire un élément k tel que kH soit égal à H . Ceci est équivalent à $k \in H \cap K$. Le résultat suit. ■

Le produit semi-direct

Nous allons étudier maintenant une construction qui illustre les développements précédents : le *produit semi-direct*. Elle apparaît naturellement dans un grand nombre de situations.

On se donne un groupe G , un groupe F ainsi qu'un homomorphisme τ de G dans le groupe $\text{Aut}(F)$ des automorphismes de F . Un élément $g \in G$ étant donné, $\tau(g)$ est donc un automorphisme de F . Soient (f, g) , (f', g') des éléments du produit $F \times G$. On définit une loi interne sur $F \times G$ par la formule suivante

$$\mu((f, g), (f', g')) = (f\tau(g)(f'), gg').$$

Théorème 5. *L'ensemble $F \times G$ muni de cette loi interne est un groupe que l'on notera $F \times_{\tau} G$. On l'appelle le produit semi-direct de G par F relativement à τ . Par ailleurs :*

- ▶ l'ensemble des éléments de la forme $(1, g)$, $g \in G$, est un sous-groupe isomorphe à G ,
- ▶ l'ensemble des éléments de la forme $(f, 1)$, $f \in F$, est un sous-groupe distingué isomorphe à F , que l'on notera par abus également F ,
- ▶ l'intersection des deux sous-groupes précédents est réduite à l'élément neutre,
- ▶ le quotient de $F \times_{\tau} G$ par $\underset{F}{\mathcal{G}}$ est isomorphe à G .

On remarquera que si τ est l'identité on retrouve le produit ordinaire des groupes.

Démonstration. Nous devons vérifier que la loi définie a un élément neutre, qu'elle est associative et que chaque élément a un inverse. Pour ce qui est de l'élément neutre on vérifie que $(1, 1)$ en est un. Pour ce qui est de l'associativité on calcule, omettant μ dans les notations :

$$\begin{aligned} ((f_0, g_0)(f_1, g_1))(f_2, g_2) &= (f_0 \tau(g_0)(f_1), g_0 g_1)(f_2, g_2) = \dots \\ \dots &= (f_0 \tau(g_0)(f_1) \tau(g_0 g_1)(f_2), g_0 g_1 g_2), \end{aligned}$$

et

$$\begin{aligned} (f_0, g_0)((f_1, g_1)(f_2, g_2)) &= (f_0, g_0)(f_1 \tau(g_1)(f_2), g_1 g_2) = \dots \\ \dots &= (f_0 \tau(g_0)(f_1 \tau(g_1)(f_2)), g_0 g_1 g_2), \end{aligned}$$

mais on a

$$\tau(g_0)(f_1) \tau(g_0 g_1)(f_2) = \tau(g_0)(f_1) \tau(g_0)(\tau(g_1)(f_2)) = \tau(g_0)(f_1 \tau(g_1)(f_2)),$$

puisque τ est un homomorphisme de G dans $\text{Aut}(F)$. Le résultat suit.

Pour ce qui est de l'inverse de (f, g) on résout l'équation :

$$(f, g)(f', g') = (f \tau(g)(f'), gg') = (1, 1),$$

ce qui donne $gg' = 1$ soit $g' = g^{-1}$, et $f \tau(g)(f') = 1$ soit $\tau(g)(f') = f^{-1}$. Comme l'application réciproque de $\tau(g)$ est $\tau(g^{-1})$ ceci donne $f' = \tau(g^{-1})(f^{-1})$. On vérifie que $(f', g')(f, g) = (1, 1)$.

La démonstration de la première condition se fait en observant que l'ensemble considéré est un sous-groupe de $F \times_{\tau} G$ et l'application qui à $g \in G$ associe $(1, g)$ est un isomorphisme.

La démonstration de la seconde condition est analogue. Il faut montrer en plus que F est distingué. Le produit $(f, g)(f', 1)(\tau(g^{-1})(f^{-1}), g^{-1})$ est égal, tous calculs faits, à $(f \tau(g)(f') f^{-1}, e)$, ce qui donne le résultat.

La troisième condition est vérifiée par construction.

Pour ce qui est de la quatrième, on commence par remarquer que l'application ϕ qui à (f, g) associe g est un homomorphisme surjectif du produit semi-direct vers G . En effet

$$\phi((f, g)(f', g')) = \phi(f \tau(g)(f'), gg') = gg' = \phi(g)\phi(g').$$

Le noyau en est F , le résultat suit, par application de la factorisation canonique. ■

Les propriétés précédentes caractérisent le produit semi-direct :

Théorème 6. Soient G et F deux groupes. Soit L un groupe tel que :

- il existe un isomorphisme i de G sur un sous-groupe de L ,
- il existe un isomorphisme j de F sur un sous-groupe distingué de L ,
- l'intersection de $\text{Im}(i)$ et $\text{Im}(j)$ est réduite à l'élément neutre,
- la composée de l'inclusion de G dans L et de la projection canonique de L sur L/F est un isomorphisme.

Alors il existe un homomorphisme $\tau : G \rightarrow \mathbf{Aut}(L)$ tel que L soit isomorphe au produit semi-direct de G par F relativement à τ .

Démonstration. Construisons τ , soit $g \in G$, alors pour tout $f \in F$, on a $i(g)j(f)i(g)^{-1} \in j(F)$ car $j(F)$ est distingué. L'application

$$\gamma_g : j(F) \longrightarrow j(F), \quad j(f) \longmapsto i(g)j(f)i(g)^{-1}$$

est donc un homomorphisme de $j(F)$, dans la formule précédente j^{-1} désigne l'automorphisme réciproque de $j : F \rightarrow j(F)$.

C'est un automorphisme car son application réciproque est $\gamma_{g^{-1}}$.

Considérons maintenant l'automorphisme $\tau_g = j^{-1} \circ \gamma_g \circ j$ de F . L'application $\tau : g \mapsto \tau_g$ est un homomorphisme de G dans $\mathbf{Aut}(F)$, en effet on a $\tau_{gg'} = \tau_g \circ \tau_{g'}$ car :

$$\begin{aligned} \tau_{gg'}(f) &= j^{-1}(i(gg')j(f)i(gg')^{-1}) \\ &= j^{-1}(i(g)i(g')j(f)i(g')^{-1}i(g^{-1})) \\ &= j^{-1}(i(g)j(j^{-1}(i(g')j(f)i(g')^{-1})))i(g^{-1}) \\ &= j^{-1}(i(g)j(\tau_{g'}(f))i(g^{-1})) = \tau_g(\tau_{g'}(f)). \end{aligned}$$

On remarquera que :

$$\tau_g(f) = j^{-1}(i(g)j(f)i(g)^{-1}).$$

Construisons maintenant un homomorphisme du produit semi-direct vers L on le définit par la formule suivante :

$$\phi(f, g) = i(g)j(\tau_{g^{-1}}(f)) = j(f)i(g).$$

On doit vérifier que :

$$j(f)i(g)j(f')i(g') = j(f\tau_g(f'))i(gg'),$$

pour tous g, g', f, f' . Or le premier terme se réécrit

$$j(f)(i(g)j(f')i(g^{-1}))i(gg'),$$

soit

$$j(f\tau_g(f'))i(gg').$$

D'où le résultat. ■

On va donner un exemple classique de produit semi-direct, il sera suivi d'un autre exemple fondamental, celui du groupe quaternionique. Ceci permettra de décrire tous les groupes à 8 éléments (exercice).

Les groupes diédraux

Soit n un entier impair. Considérons $\mathbb{Z}/n\mathbb{Z}$, on le note multiplicativement, i.e. si a est un générateur les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont de la forme a^i , pour $0 < i < n$ et $a^0 = a^n = e$. Soit τ l'élément de $\text{Aut}(\mathbb{Z}/n\mathbb{Z})$ qui envoie la classe c sur son inverse c^{-1} .

Ceci détermine un homomorphisme de $\mathbb{Z}/2\mathbb{Z} \cong \{+1, -1\}$ vers $\text{Aut}(\mathbb{Z}/n\mathbb{Z})$: à $+1$ on associe l'élément neutre, à -1 on associe τ .

Le groupe diédral D_{2n} est le produit semi-direct de $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z}$ par cet homomorphisme. Notons b le générateur de $\mathbb{Z}/2\mathbb{Z}$, on a alors, par construction, $bab^{-1} = a^{n-1}$, $b^2 = e$, $a^n = e$. Ce groupe a $2n$ éléments, en tant qu'ensemble il peut s'écrire

$$D_n = \{e, a, a^2, \dots, a^{n-1}; b, ba, ba^2, \dots, ba^{n-1}\}.$$

Le groupe D_{2n} est isomorphe au groupe des isométries du plan euclidien $\mathbb{R}^2$ laissant invariant un polygone régulier à n cotés centré en l'origine.

En effet, considérons un tel polygone P_n qui a, par exemple, un sommet en $(1, 0)$. Ce polygone est invariant par le groupe des rotations de centre l'origine et dont l'angle est un multiple entier de $\frac{2\pi}{n}$, ce groupe est isomorphe $\mathbb{Z}/n\mathbb{Z}$, soit ρ un générateur. Le polygone est également invariant par rapport à la symétrie orthogonale σ autour de l'axe des x . Cette symétrie engendre un groupe isomorphe à $\mathbb{Z}/2\mathbb{Z}$. Les éléments ρ et σ engendrent le groupe considéré.

En effet, si on a une isométrie directe qui laisse invariant P_n , c'est une rotation dont l'angle est un multiple de $\frac{2\pi}{n}$. Si une isométrie indirecte γ laisse P_n invariant l'isométrie directe $\sigma\gamma$ laisse aussi P_n invariant, c'est donc une puissance de ρ , donc $\gamma = \sigma\rho^k$ est bien dans le sous-groupe engendré par σ et ρ . Le groupe a donc $2n$ éléments et en tant qu'ensemble est égal à

$$\{e, \rho, \rho^2, \dots, \rho^{n-1}; \sigma, \sigma\rho, \sigma\rho^2, \dots, \sigma\rho^{n-1}\}.$$

La rotation conjuguée $\sigma\rho\sigma$ est égale à ρ^{n-1} .

On considère alors l'application de D_{2n} dans ce groupe qui envoie a^i sur ρ^i et ba^j sur $\sigma\rho^j$, pour tout i et tout j . C'est une bijection, par construction, de plus, c'est un homomorphisme. Vérifions par exemple que ba^jba^l a bien pour image $\sigma\rho^j\sigma\rho^l$. On a $ba^jba^l = (bab^{-1})^ja^l = a^{(n-1)j+l}$ et $\sigma\rho^j\sigma\rho^l = (\sigma\rho\sigma^{-1})^j\rho^l = \rho^{(n-1)j+l}$. D'où le résultat.

En particulier le groupe D_8 est un groupe d'ordre 8 non commutatif. C'est le groupe des isométries planes laissant invariant un carré. On notera que le sous-groupe à deux éléments engendré par b n'est pas distingué, ceci évidemment, est aussi vrai si $n \neq 2$.

On observera que D_8 a un élément d'ordre 1, 5 éléments d'ordre 2 et 2 éléments d'ordre 4.

Voici la table du groupe D_8 :

μ	e	a	a^2	a^3	b	ba	ba^2	ba^3
e	e	a	a^2	a^3	b	ba	ba^2	ba^3
a	a	a^2	a^3	e	ba^3	b	ba	ba^2
a^2	a^2	a^3	e	a	ba^2	ba^3	b	ba
a^3	a^3	e	a	a^2	ba	ba^2	ba^3	b
b	b	ba	ba^2	ba^3	e	a	a^2	a^3
ba	ba	ba^2	ba^3	b	a^3	e	a	a^2
ba^2	ba^2	ba^3	b	ba	a^2	a^3	e	a
ba^3	ba^3	b	ba	ba^2	a	a^2	a^3	e

Donnons un autre exemple de produit semi-direct. Considérons $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, et soit

$$\tau \in \text{Aut}(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})$$

l'automorphisme qui envoie la classe (a, b) sur la classe (b, a) . Ceci détermine un homomorphisme de $\mathbb{Z}/2\mathbb{Z} \cong \{+1, -1\}$ vers $\text{Aut}(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})$: à $+1$ on associe l'élément neutre, à -1 on associe τ . On considère le produit semi-direct de $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ par cet homomorphisme. Notons c et c' les deux générateurs évidents de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ et ϵ le générateur de $\mathbb{Z}/2\mathbb{Z}$. On a, par construction, $\epsilon c = c' \epsilon$ et $\epsilon c' = c \epsilon$. Ce groupe a 8 éléments, en tant qu'ensemble il est égal à :

$$\{e, c, c', cc'; \epsilon, \epsilon c, \epsilon c', \epsilon cc'\}.$$

Exercice. Vérifier que ce groupe, que l'on appellera H , est isomorphe à D_8 .

Étant donné que l'on a la table de D_8 et la liste des éléments du groupe en question il suffit de construire une application f de D_8 vers H qui soit une bijection et un homomorphisme. On commence par observer que nécessairement on a $f(e) = e$ et que l'image d'un élément d'ordre 2 (resp. 4) doit être d'ordre 2 (resp. 4). Il y a dans les deux groupes 2 éléments d'ordre 4, pour H il s'agit de ϵc et $\epsilon c'$. Posons donc $f(a) = \epsilon c$, $f(a^3) = \epsilon c'$ ce qui donne $f(a^2) = cc'$. Puis posons $f(b) = \epsilon$, ce qui implique $f(ba) = c$, $b(ba^2) = \epsilon cc'$, $f(ba^3) = c'$.

On vérifie à la main que l'application ainsi définie est un homomorphisme.

Le groupe quaternionien

Pour conclure cette section on va donner un autre exemple de groupe à 8 éléments non abélien. En fait, on vérifie (exercice) qu'un groupe non abélien à 8 éléments est isomorphe soit à D_8 soit au groupe que l'on va décrire.

On considère l'ensemble à 8 éléments constitué par les symboles

$$1, -1, i, -i, j, -j, k, -k.$$

On note Q cet ensemble. On définit sur Q une loi interne par les conditions suivantes :

- 1 est élément neutre,
- $(-1)i = -i$, $(-1)j = -j$, $(-1)k = -k$,
- $i^2 = j^2 = k^2 = -1$,
- $ij = k$, $jk = i$, $ki = j$,
- $ji = -k$, $kj = -i$, $ik = -j$.

Voici la table du groupe :

μ	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	$-k$	$-j$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

On vérifie sans peine que la loi ainsi définie est associative et que chaque élément a un inverse. L'ensemble Q est donc un groupe avec cette loi interne. Il y a un élément d'ordre 1, un élément d'ordre 2 et 6 éléments d'ordre 4. Ce groupe ne peut donc être isomorphe au groupe D_8 . En effet, s'ils l'étaient, ils auraient nécessairement autant d'éléments d'ordre 1, 2 et 4, ce qui n'est pas le cas.

3. Décomposition primaire des groupes abéliens finis

Dans toute cette section, G désigne un groupe abélien fini. On va distinguer les éléments de G suivant le « type » de leur torsion, et en déduire des conséquences sur la structure du groupe. Le théorème 1, le lemme 1, la proposition 2 sont particulièrement importants.

Définition 1

Soit p un nombre premier, soit G_p le sous-ensemble de G constitué par les éléments dont l'ordre est une puissance de p . Un tel élément est dit de p -torsion.

Le sous-ensemble G_p est en fait un sous-groupe de G comme on le vérifie aisément : si un élément x est d'ordre p^k et si un élément y est d'ordre p^l l'ordre de $x - y$ est un diviseur de $p^{\max(k,l)}$ donc une puissance de p .

Proposition 1. *Pour que G_p soit non réduit à l'élément neutre il faut que p divise l'ordre de G .*

Démonstration. En effet on sait (corollaire 1, section 2) que l'ordre d'un élément divise l'ordre du groupe. Donc s'il y a dans le groupe un élément distinct de l'élément neutre dont l'ordre est une puissance de p , p premier, cela entraîne que p divise l'ordre du groupe. ■

Théorème 1. *Pour tous les nombres premiers, sauf un nombre fini, le sous-groupe G_p est réduit à $\{0\}$. Enfin G est isomorphe au produit des groupes G_p , produit pris sur les nombres premiers p divisant l'ordre de G :*

$$\prod_{p \mid \#G} G_p.$$

On observera que l'on peut prendre le produit sur tous les nombres premiers ou seulement sur ceux pour lesquels G_p est non-trivial, le résultat est le même. La démonstration de ce théorème va prendre toute cette section. On va d'abord énoncer un certain nombre de lemmes. Le premier d'entre eux est connu sous le nom de lemme de Cauchy. Nous ne l'énonçons ici que dans le cas où G est abélien. On verra au chapitre suivant que l'on peut lever cette restriction.

Lemme 1. *Tout groupe abélien fini G dont l'ordre est divisible par un nombre premier p contient un élément d'ordre p .*

Démonstration du lemme. On raisonne par récurrence sur l'ordre de G .

Soit donc n l'ordre de G , on suppose que p divise n . Si $n = p$ le résultat est une conséquence de l'étude des groupes cycliques, faite dans la section 1. Le groupe étant alors isomorphe à $\mathbb{Z}/p\mathbb{Z}$. Supposons le résultat démontré pour tout groupe abélien fini dont l'ordre est strictement plus petit que n . Soit $g \in G$ un élément distinct de l'élément neutre.

Si le sous-groupe engendré par g , $\langle g \rangle$ est égal à G cela veut dire que G est cyclique et que g est d'ordre n . Alors l'élément $\frac{n}{p}g$ est d'ordre p .

Si $\langle g \rangle$ n'est pas égal à G et si p divise le cardinal de $\langle g \rangle$ on peut appliquer l'hypothèse de récurrence.

Sinon on peut considérer le groupe quotient $G/\langle g \rangle$ et lui appliquer l'hypothèse de récurrence. En effet, son ordre est divisible par p car c'est le quotient de n par l'ordre de g , qui est premier à p . Soit x un élément d'ordre p du groupe quotient. Soit $y \in G$ un élément tel que $\pi(y) = x$, où π désigne la projection canonique de G sur $G/\langle g \rangle$. L'ordre de y , que nous appellerons k , est un multiple de l'ordre de x , c'est-à-dire de p . L'élément $\frac{k}{p}y$ est alors d'ordre p , d'où le résultat. ■

Démonstration du théorème 1. Soient $p_1, p_2, \dots, p_n$ les nombres premiers divisant l'ordre de G . On considère l'homomorphisme suivant :

$$\prod_{i=1, \dots, n} G_{p_i} \longrightarrow G$$

$$(x_1, \dots, x_n) \longmapsto x_1 + x_2 + \dots + x_n$$

où $x_k \in G_{p_k}$.

Faisons d'abord l'observation suivante. Supposons que l'on ait $x, y \in G$, avec G abélien et que x et y soient d'ordre premiers entre eux. Supposons alors que $x + y = 0$, alors x et y sont nuls.

C'est une application de l'identité de Bézout. Soit u l'ordre de x et v l'ordre de y . Il existe a et b tels que $au + bv = 1$. On obtient donc

$$au(x + y) = 0$$

et

$$au(x + y) = aux + auy = auy = auy + bvy = (au + bv)y = y,$$

donc $y = 0$. De même on obtient $x = 0$.

Revenons à la démonstration du théorème. Supposons que l'on ait une équation de la forme $\sum_i x_i = 0$. Soient $p_1, p_2, \dots, p_n$ les nombres premiers divisant l'ordre de G .

On a alors $x_1 + \sum_{i \geq 2} x_{p_i} = 0$, l'ordre de x_1 est une puissance de p_1 , et celui de $\sum_{i \geq 2} x_i$, qui divise une puissance de $p_2 p_3 \cdots p_n$, sont premiers entre eux. Ces deux

éléments sont donc nuls par application du lemme. En particulier x_1 est nul, on considère après la relation $\sum_{i \geq 2} x_i = 0$ et on conclut par récurrence. Ceci montre l'injectivité de l'homomorphisme considéré.

Il reste à montrer la surjectivité. Pour ce faire considérons

$$n = \prod_{i=1, \dots, n} p_i^{\alpha_i}$$

la décomposition de $\#G$ en produit de puissances de nombres premiers, et posons

$$u_i = \prod_{j \neq i} p_j^{\alpha_j}.$$

Les entiers $u_1, u_2, \dots, u_n$ sont premiers entre eux, i.e. il n'y a pas d'entiers distincts de 1 ou -1 les divisant tous. On peut donc écrire l'identité de Bézout : il existe des entiers $a_1, \dots, a_n$ tels que $a_1 u_1 + \dots + a_n u_n = 1$. Soit alors $x \in G$, on a $x = a_1 u_1 x + \dots + a_n u_n x$. L'ordre de l'élément $u_i x$ (et donc $a_i u_i x$) divise $p_i^{\alpha_i}$. Cet élément est donc dans G_{p_i} . D'où le résultat. ■

La proposition suivante doit être vue comme une généralisation du lemme de Cauchy :

Proposition 2. *Le cardinal de G_{p_i} est égal à $p_i^{\alpha_i}$.*

Démonstration. Le lemme de Cauchy implique que le cardinal de G_{p_i} est une puissance de p_i , soit $p_i^{\beta_i}$. En effet, tous les éléments de G_{p_i} sont, par hypothèse, une puissance de p_i . Si le cardinal de G_{p_i} était divisible par un nombre premier p distinct de p_i , il contiendrait des éléments d'ordre p , une contradiction. On a donc $\Pi_i p_i^{\beta_i} = \Pi_i p_i^{\alpha_i}$, donc $\beta_i = \alpha_i$ pour tout i . ■

Voici un résultat classique :

Proposition 3 (lemme chinois). *Soient $(u_1, u_2, \dots, u_n)$ une famille d'entiers deux à deux premiers entre eux, et $(a_1, a_2, \dots, a_n)$ une famille d'entiers quelconques. Il existe un entier n tel que pour tout i on a $n = a_i \bmod u_i$.*

Démonstration. La projection canonique

$$\mathbb{Z}/(u_1 u_2 \cdots u_n) \mathbb{Z} \longrightarrow \prod_{i=1, \dots, n} \mathbb{Z}/u_i \mathbb{Z},$$

qui à une classe modulo $u_1 u_2 \cdots u_n$ associe les classes modulo $u_1, u_2, \dots, u_n$, est injective. Il suffit de montrer que si un entier est divisible par chacun des u_i il est divisible par leur produit, ce qui est vrai car ils sont premiers entre eux deux à deux. Comme les deux groupes ont même cardinal l'homomorphisme est bijectif. ■

Le lemme chinois peut évidemment être démontré directement ! On le fait en exercice.

Voici une application du théorème 1 dont on aura besoin dans la section 5 :

Corollaire. Soit G un groupe abélien fini. Soit m le ppcm des ordres de ses éléments, alors il existe un élément d'ordre m dans G .

Démonstration. On peut se limiter au cas où l'ordre de G est une puissance d'un nombre premier p . En effet si $G \cong \prod_{i=1, \dots, n} G_{p_i}$, et si on a un élément x_i qui convient pour G_{p_i} , l'élément $x_1 + \dots + x_n$ conviendra pour G . Si $\#G$ est une puissance d'un nombre premier le corollaire résulte de la définition. ■

Voici un exemple de décomposition primaire. Soit le groupe

$$G = \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/38\mathbb{Z} \times \mathbb{Z}/21\mathbb{Z},$$

on a :

- $\mathbb{Z}/24\mathbb{Z} \cong \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z},$
- $\mathbb{Z}/38\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/19\mathbb{Z},$
- $\mathbb{Z}/21\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}.$

Donc

$$G \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/19\mathbb{Z}.$$

4. Structure des groupes abéliens de type fini, définitions, énoncés, structure des groupes libres

Définition 1

On dit qu'un groupe abélien G est un *groupe de type fini* s'il a un nombre fini de générateurs.

Les groupes abéliens de type fini, c'est-à-dire les quotients de $\mathbb{Z}^n$ pour un certain entier n , se rencontrent dans une grande variété de contextes mathématiques. Nous allons étudier leur structure. Ils se « répartissent » en deux « types » :

Définition 2

On dit qu'un groupe abélien est un *groupe de torsion* si tout élément de G est d'ordre fini, (on rappelle qu'un élément est dit de torsion si son ordre est fini).

Soit p un nombre premier.

Définition 3

Un groupe abélien est dit de p -torsion si l'ordre de tout élément est une puissance de p . On dit aussi qu'un élément dont l'ordre est une puissance de p est de p -torsion.

Proposition 1. *Un groupe abélien G de torsion et de type fini est fini.*

Démonstration. En effet c'est un quotient de $\mathbb{Z}^n$, n désignant le cardinal d'un système de générateurs (proposition 6, section 1).

Soit alors k le produit des ordres des éléments du système de générateurs. Puisque le groupe est abélien l'ordre de tout élément du groupe est un diviseur de k . L'homomorphisme surjectif de $\mathbb{Z}^n$ sur G est donc nul sur $(k\mathbb{Z})^n$. Le quotient de $\mathbb{Z}^n$ par $(k\mathbb{Z})^n$ est isomorphe à $(\mathbb{Z}/k\mathbb{Z})^n$.

Le groupe G est donc un quotient de $(\mathbb{Z}/k\mathbb{Z})^n$ (voir section 2 proposition 3) et est donc fini. ■

Définition 4

On appelle *sous-groupe de torsion* d'un groupe abélien G le sous-groupe constitué par les éléments de torsion.

Un groupe G dont le sous-groupe de torsion est réduit à $\{0\}$ sera dit sans torsion.

Soit G un groupe abélien de type fini,

Définition 5

Le groupe abélien G sera dit libre s'il est isomorphe à $\mathbb{Z}^n$ pour un certain entier n .

Un groupe abélien libre est sans-torsion. Les résultats suivants décrivent la structure des groupes abéliens de type fini. Le premier résultat sur la structure des groupes abéliens de type fini est que les deux types décrits ci-dessus permettent de les décrire tous :

Théorème 1. *Tout groupe abélien de type fini G est somme directe d'un groupe abélien libre L et d'un groupe de torsion T . Si on a des isomorphismes $G \cong L \oplus T$ et $G \cong L' \oplus T'$, L et L' libres, T et T' de torsion alors L est isomorphe à L' et T est isomorphe à T' .*

Il nous faut maintenant étudier les deux types en question. On commence par le cas des groupes libres. Pour le cas des groupes de torsion on renvoie à la section suivante.

Théorème 2. Soit L un groupe abélien sans torsion de type fini, alors L est isomorphe à $\mathbb{Z}^n$ pour un entier n bien déterminé.

Ce théorème n'est pas vrai si on remplace la mention sans torsion de type fini par sans torsion.

Les théorèmes 1 et 2 seront démontrés plus loin. On commencera par énoncer et démontrer la proposition suivante :

Proposition 2. Soit G un sous-groupe de $\mathbb{Z}^n$, alors G est isomorphe à $\mathbb{Z}^r$ pour un entier $r \leq n$. L'entier r est unique.

Démonstration. La démonstration se fait par récurrence sur n . Le théorème a déjà été démontré pour $n = 1$ dans la première section. Supposons le donc démontré pour $n - 1$.

Soit π la projection sur le dernier facteur de $\mathbb{Z}^n$ qui à $(x_1, \dots, x_n)$ associe x_n . L'image $\pi(G)$ est un sous-groupe de $\mathbb{Z}$ et est donc égal à $a\mathbb{Z}$ pour un certain entier a éventuellement nul. Si $a = 0$ on est ramené au cas $n - 1$, car

$$G \subset \text{Ker}(\pi) \cong \mathbb{Z}^{n-1},$$

et l'hypothèse de récurrence nous donne le résultat.

Si $a \neq 0$ soit ω un élément tel que $\pi(\omega) = a$. Le sous-groupe

$$\text{Ker}(\pi) \cap G \subset \text{Ker}(\pi) \cong \mathbb{Z}^{n-1}$$

est isomorphe à $\mathbb{Z}^{r'}$ pour un entier $r' \leq n - 1$. Soit ψ un tel isomorphisme. Étant donné un élément $x \in G$ l'élément

$$x - \frac{\pi(x)}{a}\omega$$

est dans $\text{Ker}(\pi)$. Considérons l'homomorphisme

$$\begin{aligned} G &\longrightarrow (\text{Ker}(\pi) \cap G) \times \mathbb{Z}, \\ x &\longmapsto \left(x - \frac{\pi(x)}{a}\omega, \frac{\pi(x)}{a} \right). \end{aligned}$$

C'est un isomorphisme, l'inverse est donné par l'application

$$(y, t) \longmapsto y + t\omega.$$

Donc G est isomorphe à $\text{Ker}(\pi) \cap G \oplus \mathbb{Z}$ qui est lui-même isomorphe via $\psi \oplus \text{Id}$ à $\mathbb{Z}^{r'+1}$, $r' + 1 \leq n$.

Montrons maintenant que l'entier r est bien déterminé. Cela résulte du :

Lemme 1. Si on a un isomorphisme ϕ entre $\mathbb{Z}^r$ et $\mathbb{Z}^t$ alors $r = t$.

Démonstration du lemme. Considérons $\mathbb{Z}^r$ comme sous-groupe de $\mathbb{Q}^r$ et $\mathbb{Z}^t$ comme sous-groupe de $\mathbb{Q}^t$. L'isomorphisme ϕ « s'étend » en un isomorphisme Φ d'espaces vectoriels sur $\mathbb{Q}$ de $\mathbb{Q}^r$ dans $\mathbb{Q}^t$. Ce qui implique que $r = t$ par unicité de la dimension.

Il nous reste à construire Φ . Soit $(x_1, \dots, x_r) \in \mathbb{Q}^r$, on peut trouver un entier μ tel que $\mu(x_1, \dots, x_r) \in \mathbb{Z}^r$, il suffit pour cela de prendre un multiple du ppcm m des dénominateurs des x_i mis sous forme de fractions irréductibles. Définissons Φ par :

$$\Phi(x_1, \dots, x_r) = \frac{\phi(\mu x_1, \dots, \mu x_r)}{\mu}.$$

La valeur ne dépend pas du choix de μ car le terme de droite est toujours égal à

$$\frac{\phi(mx_1, \dots, mx_r)}{m}.$$

On vérifie que Φ est une application linéaire d'espaces vectoriels sur $\mathbb{Q}$. Elle est injective car

$$\frac{\phi(mx_1, \dots, mx_r)}{m} = 0$$

implique

$$\phi(mx_1, \dots, mx_r) = 0,$$

soit $(mx_1, \dots, mx_r) = 0$ par injectivité de ϕ , donc $(x_1, \dots, x_r) = 0$. Elle est surjective : si on a $(y_1, \dots, y_t) \in \mathbb{Q}^t$ soit m' tel que $m'(y_1, \dots, y_t) \in \mathbb{Z}^t$. Il existe $(x_1, \dots, x_r) \in \mathbb{Z}^r$ tel que $\phi(x_1, \dots, x_r) = m'(y_1, \dots, y_t)$. Enfin on vérifie que

$$\Phi\left(\frac{(x_1, \dots, x_r)}{m'}\right) = (y_1, \dots, y_t). \quad \blacksquare$$

Démonstration du théorème 2. Soit L un groupe abélien de type fini ayant un système de m générateurs, L est donc un quotient de $\mathbb{Z}^m$ via un homomorphisme p .

Considérons un système $(x_1, \dots, x_n)$ d'éléments de L tels que :

- si on a une relation $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$ tous les λ_i sont nuls,
- il n'existe pas de familles de ce type comportant plus d'éléments.

La première condition équivaut à dire que $\langle x_1, \dots, x_n \rangle$ est isomorphe à $\mathbb{Z}^n$.

On dira qu'une telle famille est un système libre maximal.

Des familles satisfaisant à la première condition et ayant un cardinal non nul existent dès que le groupe n'est pas un groupe de torsion. En effet, il suffit de prendre la famille réduite à un élément qui n'est pas de torsion. L'existence d'un système libre maximal est garantie par la proposition précédente.

En effet, si G est un quotient de $\mathbb{Z}^m$ et si $\{x_1, \dots, x_n\}$ est une famille de G satisfaisant à la première condition, on va montrer que $n \leq m$. Soit π l'application canonique de $\mathbb{Z}^m$ dans G . L'inégalité résulte de l'observation suivante : soient $x'_i \in \mathbb{Z}^m$ tels que $\pi(x'_i) = x_i$. Le système formé par les x'_i satisfait à la première condition, il engendre un groupe isomorphe à $\mathbb{Z}^n$, donc $n \leq m$ d'après la proposition 2.

Soit $\{x_1, \dots, x_n\}$ une famille satisfaisant aux deux conditions. Considérons un système de générateurs $\{y_1, \dots, y_m\}$ de L , puis le système $\{x_1, \dots, x_n, y_j\}$. Pour tout $1 \leq j \leq m$ Il existe des entiers $\lambda_{i,j}$, $1 \leq i \leq n$, et μ_j non tous nuls tels que

$$\lambda_{1,j}x_1 + \dots + \lambda_{n,j}x_n + \mu_j y_j = 0$$

d'après la seconde condition. La première condition implique que $\mu_j \neq 0$. Notons F le sous-groupe de L engendré par $\{x_1, \dots, x_n\}$, il est isomorphe à $\mathbb{Z}^n$.

Soit $t = \prod_{j=1, \dots, m} \mu_j$, les relations précédentes montrent que pour tout j on a $ty_j \in F$, donc $tL \subset F$, puisque les y_j engendrent L . Il en résulte que le sous-groupe tL de F est isomorphe à $\mathbb{Z}^k$, $k \leq n$.

Par ailleurs, comme $t \neq 0$ donc L est isomorphe à tL via l'application qui à $x \in L$ associe tx . Cet homomorphisme est surjectif par définition, il est injectif parce que le groupe est sans torsion. On en conclut que L est isomorphe à $\mathbb{Z}^k$. Comme $F \cong \mathbb{Z}^n$ est un sous-groupe de $L \cong \mathbb{Z}^k$ on a forcément $n \leq k$, donc $k = n$.

Enfin, on notera que deux systèmes libres maximaux ont même cardinal. Soient en effet $\{y_1, \dots, y_n\}$ et $\{y'_1, \dots, y'_{n'}\}$ deux tels systèmes, F et F' les sous-groupes qu'ils engendrent. L'argument précédent montre que $tF' \subset F$ donc que $n \leq n'$. On montre de manière similaire l'inégalité $n' \leq n$, et donc $n = n'$. $\blacksquare$

Démonstration du théorème 1. Le quotient de G par son sous-groupe de torsion T est un groupe libre L car il est sans torsion et de type fini. Le groupe G est isomorphe à $L \oplus T$. Ceci résulte du lemme suivant :

Lemme 1. Soit π un homomorphisme surjectif d'un groupe abélien sur $\mathbb{Z}^n$. Il existe un homomorphisme $s : \mathbb{Z}^n \rightarrow G$ tel que l'homomorphisme composé $\pi \circ s$ soit l'identité de $\mathbb{Z}^n$.

Démonstration du lemme. Soient $e_1, \dots, e_n$ les générateurs canoniques de $\mathbb{Z}^n$. Soient $e'_1, \dots, e'_n$ des éléments de G tels que pour tout $i = 1, \dots, n$, $s(e'_i) = e_i$.

L'application qui à l'élément $\lambda_1 e_1 + \dots + \lambda_n e_n$ associe l'élément $\lambda_1 e'_1 + \dots + \lambda_n e'_n$ est un homomorphisme de $\mathbb{Z}^n$ dans G qui satisfait à la condition requise. $\square$

On revient à la démonstration du théorème. Le groupe L est isomorphe à $\mathbb{Z}^n$ pour un certain entier n . Soit s un homomorphisme de L dans G tel que $\pi \circ s$ soit l'identité de L , où $\pi : G \rightarrow L$ est la projection canonique. Pour tout $x \in G$ l'élément $x - s(\pi(x))$ est dans le sous-groupe de torsion. Il suffit de vérifier que

$\pi(x - s \circ \pi(x)) = 0$, mais on a

$$\pi(x - s \circ \pi(x)) = \pi(x) - (\pi \circ s)(\pi(x)) = \pi(x) - \pi(x) = 0.$$

Considérons alors l'homomorphisme de G dans $L \oplus T$ qui à

$$x \longmapsto (\pi(x), x - s(\pi(x))).$$

C'est un isomorphisme car il admet pour inverse l'homomorphisme qui à (x, y) associe $s(x) + y$.

Démontrons maintenant que si on a deux décompositions on a forcément $L \cong L'$ et $T \cong T'$. Les groupes T et T' sont nécessairement tous les deux isomorphes au sous-groupe de torsion de G , donc ils sont isomorphes. Les groupes L et L' sont eux tous les deux isomorphes au quotient de G par son sous-groupe de torsion. Ils sont donc tous les deux isomorphes. ■

Voici un exemple de sous-groupe de $\mathbb{Z}^4$. On considère le sous-ensemble T des (x, y, z, t) tels que $x + 2y + 3z = 0$ et $2y + 5t = 0$. Les éléments $e = (3, 0, -1, 0)$ et $e' = (-10, 5, 0, -2)$ sont dans T . C'est un système générateur libre de T . Soit en effet $v = (x, y, z, t) \in T$, de la seconde équation on déduit que $5 \mid y$ et $2 \mid t$. En particulier $w = v - \frac{y}{5}e'$ est encore dans T .

Soit w est de la forme $(x', 0, z', 0)$, x' est divisible par 3 et est égal à $\frac{x'}{3}e$. Le sous-groupe est isomorphe à $\mathbb{Z}^2$.

5. Structure des groupes abéliens de type fini de torsion

Dans cette section nous allons étudier le cas des groupes de torsion. La formulation proposée est due à A. Liulevičius. Voici l'énoncé principal :

Théorème 1. *Soit G un groupe abélien fini, c'est-à-dire un groupe égal à son sous-groupe de torsion. Il existe une unique suite d'entiers strictement supérieurs à 1, $(m_1, \dots, m_n)$, tels que :*

- pour tout i , $1 \leq i \leq n-1$, $m_{i+1} \mid m_i$;
- le groupe G est isomorphe à $\mathbb{Z}/m_1\mathbb{Z} \times \dots \times \mathbb{Z}/m_n\mathbb{Z}$.

Les entiers m_i sont appelés les facteurs invariants du groupe.

Avant d'entamer la démonstration de ce théorème énonçons et démontrons un lemme.

Lemme 1. *Soit G un groupe abélien fini dont l'ordre de tous les éléments divise un entier n donné. Soit H un sous-groupe de G , et soit $\phi : H \rightarrow \mathbb{Z}/n\mathbb{Z}$ un homomorphisme. Il existe un homomorphisme $\psi : G \rightarrow \mathbb{Z}/n\mathbb{Z}$ tel que ψ restreint à H soit égal à ϕ .*

Démonstration du lemme. Supposons d'abord que G soit engendré par H et un élément $x \in G$. Ceci signifie que tout élément de G s'écrit sous la forme $h + \lambda x$, pour $h \in H$ et $\lambda \in \mathbb{Z}$. Cette écriture n'est pas nécessairement unique. Si on a $h + \lambda x = h' + \lambda' x$, on a $h - h' = (\lambda' - \lambda)x \in H \cap \langle x \rangle$. Le sous-groupe $\langle x \rangle$ est isomorphe à $\mathbb{Z}/k\mathbb{Z}$, k désignant l'ordre de x , par hypothèse $k \mid n$. Le sous-groupe $H \cap \langle x \rangle$ de $\langle x \rangle$ est cyclique et admet pour générateur un multiple de x de la forme lx avec $l \mid k$ (corollaire 2, section 2).

Considérons maintenant la valeur $y = \phi(lx) \in \mathbb{Z}/n\mathbb{Z}$. Il existe un élément $w \in \mathbb{Z}/n\mathbb{Z}$ tel que $lw = y$. En effet, pour que ceci ait lieu il faut et suffit que $\frac{n}{l}y = 0$ (corollaire à la proposition 9 de la section 1). Or on a

$$\frac{n}{l}y = \frac{n}{l}\phi(lx) = \phi\left(\frac{n}{l}lx\right) = \phi(nx) = \phi(0) = 0.$$

Ayant choisi un tel élément w , définissons ψ par la formule

$$\psi(h + \lambda x) = \phi(h) + \lambda w.$$

Vérifions que ceci est bien défini. Si on a $h + \lambda x = h' + \lambda' x$, on doit montrer que $\phi(h) + \lambda w = \phi(h') + \lambda' w$, soit $\phi(h - h') = (\lambda' - \lambda)w$.

Mais

$$(\lambda' - \lambda)x = h - h' \in H \cap \langle x \rangle,$$

donc $(\lambda' - \lambda)x = \alpha lx$, en particulier

$$\lambda' - \lambda \equiv \alpha l \pmod{k}.$$

Donc :

$$\phi(h - h') = \phi(\alpha lx) = \alpha \phi(lx) = \alpha y = \alpha lw,$$

et

$$\lambda w - \lambda' w = -\alpha lw,$$

car l'ordre de w est un diviseur de k . L'application est donc bien définie.

La démonstration du lemme s'achève par une récurrence. Le groupe G est engendré par H et un nombre fini d'éléments x_i , $i = 1, \dots, k$. On fait une récurrence sur i variant de 1 à k . On vient de faire le pas initial de la récurrence. On suppose avoir prolongé l'homomorphisme au sous-groupe engendré par H et $x_1, \dots, x_i$. De nouveau, on prolonge à celui engendré par H et $x_1, \dots, x_{i+1}$ comme il a été indiqué. ■

Démonstration du théorème. Considérons un élément x d'ordre maximal m_1 de G , un tel élément existe car G est fini. En fait, d'après le corollaire 2 à la proposition 2 de la section 3, l'ordre de tout autre élément de G divise m_1 . Le sous-groupe engendré par x est isomorphe à $\mathbb{Z}/m_1\mathbb{Z}$.

On peut alors appliquer le lemme 1 au sous-groupe $\mathbb{Z}/m_1\mathbb{Z}$ de G , avec ϕ l'identité Id de $\mathbb{Z}/m_1\mathbb{Z}$. Il existe donc $\pi : G \rightarrow \mathbb{Z}/m_1\mathbb{Z}$ qui restreint à $\mathbb{Z}/m_1\mathbb{Z}$ est l'identité.

L'homomorphisme π composé à l'inclusion i de $\mathbb{Z}/m_1\mathbb{Z}$ dans G détermine un homomorphisme p de G dans lui-même, on a $p \circ p = p$. Définissons un homomorphisme de G dans $\mathbb{Z}/m_1\mathbb{Z} \oplus \text{Ker}(\pi)$ par

$$x \mapsto (\pi(x), x - p(x)).$$

C'est un isomorphisme, car il admet pour isomorphisme réciproque :

$$(x, y) \mapsto i(x) + y.$$

On peut alors faire une récurrence, car l'ordre des éléments de $\text{Ker}(\pi)$ divise m_1 et soit donc m_2 l'ordre maximal d'un élément de $\text{Ker}(\pi)$...

On va maintenant démontrer l'unicité des m_i . Supposons que l'on ait un isomorphisme de G avec $\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_n\mathbb{Z}$ et avec $\mathbb{Z}/m'_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m'_n\mathbb{Z}$ avec les conditions requises sur les m_i et sur les m'_j . Soit p un nombre premier divisant m_n . L'ensemble des éléments d'ordre p de

$$\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_n\mathbb{Z}$$

et donc de G est isomorphe au produit de n copies de $\mathbb{Z}/p\mathbb{Z}$, et a donc p^n éléments. Supposons que $n' \leq n$ alors

$$\mathbb{Z}/m'_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m'_{n'}\mathbb{Z}$$

a au plus $p^{n'}$ éléments d'ordre p . On en déduit que $n = n'$ et que tout nombre premier divisant m_n divise m'_n .

Faisons alors une hypothèse de récurrence sur l'ordre de G . Et soit p divisant m_n et m'_n . Le sous-groupe pG de G est isomorphe à

$$\mathbb{Z}/\frac{m_1}{p}\mathbb{Z} \times \cdots \times \mathbb{Z}/\frac{m_n}{p}\mathbb{Z}$$

et à

$$\mathbb{Z}/\frac{m'_1}{p}\mathbb{Z} \times \cdots \times \mathbb{Z}/\frac{m'_n}{p}\mathbb{Z}.$$

On en déduit par récurrence que, pour tout i , on a $\frac{m_i}{p} = \frac{m'_i}{p}$, d'où le résultat. ■

Voici un exemple : Soit le groupe

$$G = \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/38\mathbb{Z} \times \mathbb{Z}/21\mathbb{Z},$$

rappelons que l'on a :

$$G \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/19\mathbb{Z}.$$

ce qui se réécrit :

$$G \cong (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \times \mathbb{Z}/19\mathbb{Z}),$$

soit

$$G \cong \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/3192\mathbb{Z}.$$

Exercices

Certains exercices font appel à des définitions énoncées au chapitre 2. Entre autres celle du groupe symétrique, le lecteur, au cas il ne saurait pas cette définition, s'y reportera.

I. Généralités

1. a) Soit G un groupe dont tous les éléments, distincts de l'élément neutre, sont d'ordre 2. Montrer qu'il est commutatif.
b) Montrer que si G est fini son cardinal est une puissance de 2, on pourra raisonner par récurrence sur l'ordre de G .
c) Montrer que G est isomorphe au groupe $(\mathbb{Z}/2\mathbb{Z})^n$.
2. Écrire la table de tous les groupes à 4 et 6 éléments.
3. Soit p un nombre premier, montrer que tout groupe d'ordre $2p$ contient un élément d'ordre p , on utilisera l'exercice 1.
4. Pour quelles valeurs de n l'entier $n^3 - 2n + 7$ est-il divisible par 3, 5 et 7 ?
5. Trouver les solutions entières de :

$$\begin{cases} 5x + 3y \equiv 1 \\ x + 2y \equiv 4 \end{cases}$$

pour la relation de congruence modulo 3, respectivement pour la relation de congruence modulo 7.

6. Supposons donnés un groupe abélien G , un sous-groupe H , et un homomorphisme φ de H dans $\mathbb{Q}$. On suppose de plus que G est engendré par H et par un élément x . Montrer qu'il existe un homomorphisme $\bar{\varphi}$ de G dans $\mathbb{Q}$, qui restreint à H , est égal à φ . On fera attention à ce que l'intersection du sous-groupe engendré par x avec H n'est pas nécessairement réduite à $\{0\}$.
7. Montrer que les matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

forment un système générateur de $\mathrm{GL}_2(\mathbb{Z})$.

II. Sous-groupes distingués, groupes quotients

1. Soient G un groupe et H un sous-groupe tel que l'ensemble G/H ait deux éléments. Montrer que H est distingué dans G .

2. a) Soit G un groupe, on appelle commutateur un élément de G de la forme $xyx^{-1}y^{-1}$, $x, y \in G$. On appelle *groupe dérivé* de G et on note $D(G)$ le sous-groupe engendré par les commutateurs. Montrer que $D(G)$ est distingué dans G . Montrer que le groupe quotient $G/D(G)$ est commutatif.
- b) Montrer que tout homomorphisme de G dans un groupe abélien A factorise au travers de $G/D(G)$. On appelle $G/D(G)$ l'abélianisé de G .
- c) Calculer $D(G)$ pour $\mathfrak{S}_3, \mathfrak{S}_4$.
3. a) Montrer que le groupe $\mathbb{Q}/\mathbb{Z}$ a un sous-groupe à n éléments et un seul.
- b) Montrer que le quotient de $\mathbb{Q}/\mathbb{Z}$ par un sous-groupe fini est isomorphe à $\mathbb{Q}/\mathbb{Z}$.
4. Déterminer tous les homomorphismes de $\mathbb{Q}/\mathbb{Z}$ dans $\mathbb{Z}$.
5. On considère le groupe quotient $\mathbb{Q}/\mathbb{Z}$. Supposons donnés un groupe abélien G , un sous-groupe H , et un homomorphisme φ de H dans $\mathbb{Q}/\mathbb{Z}$. On suppose de plus que G est engendré par H et un élément supplémentaire x . Montrer qu'il existe un homomorphisme $\bar{\varphi}$ de G dans $\mathbb{Q}/\mathbb{Z}$ qui restreint à H est égal à φ .
6. a) Soit p un nombre premier. On note $\mathbb{Z}/p^\infty$ le sous-ensemble de $\mathbb{Q}/\mathbb{Z}$ constitué par les classes des fractions $\frac{k}{p^n}$, $k \in \mathbb{Z}$. Montrer que c'est un sous-groupe.
- b) Déterminer l'ensemble des homomorphismes de $\mathbb{Z}/p^\infty$ dans lui même : on montrera que c'est un groupe G pour l'addition ; puis qu'il existe un homomorphisme surjectif $\rho_n : G \rightarrow \mathbb{Z}/p^n\mathbb{Z}$ tel que pour tout n on ait $p_n \circ \rho_n = \rho_{n-1}$, où p_n est la surjection canonique $p_n : \mathbb{Z}/p^n \rightarrow \mathbb{Z}/p^{n-1}$. On montrera enfin qu'un homomorphisme g est déterminé par la suite des $\rho_n(g)$.
7. Soit G un groupe. Montrer que le centre $Z(G)$ de G est un sous-groupe distingué.
- ' Montrer que pour tout automorphisme ψ du groupe G on a $\psi(Z(G)) \subset Z(G)$.
8. Soit G un groupe, montrer que, si le groupe quotient $G/Z(G)$ est cyclique, le / groupe G est abélien.
9. a) (Voir [Del]) Montrer que si $G/Z(G)$ est fini il en est de même de $D(G)$ (exercice 3). On commencera par montrer que l'ensemble des commutateurs de paires d'éléments de G est fini et stable par les automorphismes intérieurs. Majorer le cardinal ℓ de cet ensemble en fonction du cardinal n du groupe quotient. On notera ℓ ce cardinal.
- b) Montrer que si γ est un commutateur sont des commutateurs un produit de $p+q$ commutateurs comportant γ au moins p fois peut se réécrire sous la forme $\gamma^p u$ où u est un produit de q commutateurs.

c) Montrer que l'on a la relation :

$$[x, y^2][xyx^{-1}, y]^{n-1} = [x, y]^{n+1},$$

pour tout x et y , on rappelle que $[x, y] = xyx^{-1}y^{-1}$. Montrer finalement qu'on peut ramener tout produit de commutateurs à au plus $n\ell$ termes.

10. Soient G un groupe et H et K deux sous-groupes. Montrer que l'ensemble HK est un sous-groupe si et seulement si il est égal au sous-ensemble KH .
11. Soient G un groupe et H et K des sous-groupes. Montrer que si les cardinaux de G/H et de G/K sont finis et premiers entre eux alors $G = HK$.
12. Montrer que $\mathfrak{S}_3$ est un produit semi-direct. Montrer que le groupe orthogonal $O(2)$ est le produit semi-direct de $\mathbb{Z}/2\mathbb{Z}$ et de $SO(2)$.
13. Calculer le centre des groupes D_8 et Q . Calculer les groupes quotients. De même rechercher les groupes dérivés $D(G)$ pour D_8 et Q et déterminer les groupes abélianisés.
14. a) Montrer que tous les sous-groupes du groupe quaternionien Q sont distingués. Montrer que le groupe Q n'est pas le produit semi-direct d'un groupe d'ordre 2 par un groupe d'ordre 4.
b) Identifier Q à un sous-groupe de $GL_2(\mathbb{C})$.
15. Décrire tous les groupes d'ordre 8. On commencera par décrire tous les groupes abéliens, puis on passera au cas non abélien.

III. Groupes abéliens, groupes abéliens de type fini

1. a) Montrer que $\mathbb{Q}$ n'a pas de système générateur libre.
b) Montrer que tout sous-groupe de $\mathbb{Q}$ ayant un nombre fini de générateurs est isomorphe à $\mathbb{Z}$.
c) Le résultat précédent reste-t-il vrai si on remplace $\mathbb{Q}$ par $\mathbb{R}$?
2. Écrire la décomposition primaire des groupes suivants :
 - $\mathbb{Z}/26\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/54\mathbb{Z}$,
 - $\mathbb{Z}/36\mathbb{Z} \times \mathbb{Z}/120\mathbb{Z} \times \mathbb{Z}/55\mathbb{Z}$,
 - $\mathbb{Z}/30\mathbb{Z} \times \mathbb{Z}/48\mathbb{Z} \times \mathbb{Z}/54\mathbb{Z} \times \mathbb{Z}/26\mathbb{Z}$.
3. Trouver les facteurs invariants des groupes suivants :
 - $\mathbb{Z}/36\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/90\mathbb{Z}$,
 - $\mathbb{Z}/27\mathbb{Z} \times \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/55\mathbb{Z}$,
 - $\mathbb{Z}/63\mathbb{Z} \times \mathbb{Z}/48\mathbb{Z} \times \mathbb{Z}/54\mathbb{Z}$.

4. On dit qu'un élément x dans un groupe abélien G est infiniment divisible si pour tout entier n non nul il existe un élément de G tel que $ny = x$. Montrer que le groupe quotient $\mathbb{Z}^{\times\mathbb{N}}/\mathbb{Z}^{\mathbb{N}}$ contient des éléments non nuls infiniment divisibles. On rappelle que $\mathbb{Z}^{\times\mathbb{N}}$ est l'ensemble des suites d'entiers, et que $\mathbb{Z}^{\mathbb{N}}$ est l'ensemble des suites d'entiers nuls, sauf pour un nombre fini d'entre eux.
5. Soit $v = (a_1, \dots, a_n) \in \mathbb{Z}^n$. Montrer que les conditions suivantes sont équivalentes :
- il existe un homomorphisme $\mu : \mathbb{Z}^n \rightarrow \mathbb{Z}$ tel que $\mu(v) = 1$,
 - les entiers $a_1, \dots, a_n$ sont premiers entre eux dans leur ensemble,
 - il existe une matrice (n, n) à coefficients entiers de déterminant égal à 1 dont la première colonne est v .
6. Montrer que les automorphismes de $(\mathbb{Z}/2\mathbb{Z})^2$ (resp. $(\mathbb{Z}/4\mathbb{Z})^2$) sont en bijection avec les matrices $(2, 2)$ à coefficients dans $\mathbb{Z}/2\mathbb{Z}$ de déterminant non-nul (resp. les matrices $(2, 2)$ à coefficients dans $\mathbb{Z}/4\mathbb{Z}$ de déterminant inversible).
7. a) Décrire le groupe des automorphismes de $(\mathbb{Z}/2\mathbb{Z})^2$. Montrer qu'il est isomorphe au groupe symétrique $\mathfrak{S}_3$.
- b) Décrire le groupe des automorphismes de $(\mathbb{Z}/2\mathbb{Z}) \times \mathbb{Z}/4\mathbb{Z}$: construire un homomorphisme dans le groupe des automorphismes de $(\mathbb{Z}/2\mathbb{Z})^2$. Déterminer l'image et le noyau de cet homomorphisme. Calculer le cardinal du groupe et l'identifier.
- c) Décrire le groupe des automorphismes de $(\mathbb{Z}/4\mathbb{Z})^2$: en construire un quotient isomorphe au groupe des automorphismes de $(\mathbb{Z}/2\mathbb{Z})^2$. Expliciter le noyau de l'application quotient.

■ Quelques réponses ou indications ■

I. Généralités

1. a) Utiliser la relation $(xy)^2 = 1 : xyxy = 1$ implique $xy = y^{-1}x^{-1}$, soit $xy = yx$, car $x = x^{-1}$ et $y = y^{-1}$.
- b) c) On considère un élément d'ordre 2, le sous-groupe qu'il engendre, le quotient de G par ce sous-groupe et on fait une récurrence.
2. Pour ce qui est des groupes d'ordre 4, selon que le groupe a ou n'a pas un élément d'ordre 4 on a à considérer les groupes $\mathbb{Z}/4\mathbb{Z}$ ou $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.
- Pour ce qui est des groupes d'ordre 6, on montre que si le groupe est abélien il est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ et que s'il est non-abélien il l'est au groupe symétrique $\mathfrak{S}_3$. En effet, en utilisant l'exercice 1 on montre qu'il y a au moins un élément d'ordre 3. Puis on montre qu'il y a au moins un élément d'ordre 2 par un argument de comptage car les

éléments c d'ordre 3 vont par paire (c et c^2). Le groupe $\mathfrak{S}_3$ est isomorphe à D_3 , voici sa table en notant c un élément d'ordre 3 et τ un élément d'ordre 2 :

μ	e	c	c^2	τ	τc	τc^2
e	e	c	c^2	τ	τc	τc^2
c	c	c^2	e	τc^2	τ	τc
c^2	c^2	e	c	τc	τc^2	τ
τ	τ	τc	τc^2	e	c	c^2
τc	τc	τc^2	τ	c^2	e	c
τc^2	τc^2	τ	τc	c	c^2	e

3. Observer que les éléments distincts de l'élément neutre sont d'ordre 2 ou p , puis montrer à l'aide de l'exercice 1 qu'il y en a au moins un d'ordre p .
6. Soit ψ l'extension cherchée. Si l'élément x est de torsion on pose $\psi(x) = 0$. Sinon on considère le plus petit entier positif non nul tel que $nx \in H$ et on pose

$$\psi(x) = \frac{\phi(nx)}{n}.$$

Dans les deux cas on doit vérifier que ceci détermine bien un homomorphisme.

7. Soit H le sous-groupe engendré par les matrices considérées. Montrer que les matrices

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

$k \in \mathbb{Z}$, sont dans H . Puis, en utilisant la division euclidienne, montrer qu'en multipliant à gauche et à droite par des matrices de H une matrice A quelconque, on peut annuler un coefficient.

II. Sous-groupes distingués, groupes quotients

2. a) b) Observer que le groupe dérivé est contenu dans le noyau de tout homomorphisme de G vers un groupe abélien et appliquer la factorisation canonique.
- c) On trouve les groupes alternés $\mathcal{A}_3$ et $\mathcal{A}_4$.
3. a) Il s'agit du sous-groupe constitué par les classes d'éléments de la forme $\frac{a}{n}$, avec a entier.
- b) Si n est l'ordre du groupe par lequel on quotiente considérer la multiplication par n de $\mathbb{Q}/\mathbb{Z}$ dans lui même.
4. Dans $\mathbb{Q}/\mathbb{Z}$ tous les éléments sont de torsion, il ne peut donc y avoir d'homomorphisme non trivial dans un groupe sans éléments non triviaux de torsion.
5. On reprend la démonstration du lemme 2 de la dernière section du chapitre. Le point fondamental est de remarquer que pour tout $x \in \mathbb{Q}/\mathbb{Z}$ et tout n non nul, il existe $y \in \mathbb{Q}/\mathbb{Z}$ tel que $ny = x$.

6. b) Étant donné un homomorphisme on observe que sa restriction au sous-groupe $\mathbb{Z}/p^n\mathbb{Z}$ prend valeurs dans $\mathbb{Z}/p^n\mathbb{Z}$.
8. Choisir un relèvement du générateur g de $G/Z(G)$, observer que ce relèvement et $Z(G)$ engendrent G . Donc tout élément s'écrit $g^k z$, $z \in Z(G)$. On a alors $g^k z g^h z' = g^k z z' g^h = z z' g^h g^k = z' g^h z g^k$.
9. a) On choisit un système (fini!) de représentants du groupe quotient dans G . C'est-à-dire qu'on choisit pour chaque élément du quotient un élément de G dont il est l'image. Puis on montre qu'il suffit de prendre les commutateurs et produits de commutateurs de ces éléments pour avoir tous les commutateurs. Puis on utilise que $zxyx^{-1}y^{-1}z^{-1} = zxxz^{-1}zyz^{-1}zx^{-1}z^{-1}zy^{-1}z^{-1}$.
- b) On écrit l'élément considéré $x_1 \dots u_k \gamma x$ faisant apparaître la première occurrence de γ . L'élément x est donc produit de $p+q-1-k$ termes et fait apparaître γ $p-1$ fois. On peut appliquer une hypothèse de récurrence. Pour la débiter on remarque que $u\gamma = \gamma\gamma^{-1}u\gamma$.
- c) Pour démontrer la formule on utilise l'hypothèse : si u est un commutateur u^n est dans le centre.
10. Introduire le sous-groupe $H \cap K$ et comparer différents indices en utilisant la condition de primalité.
11. Appliquer la caractérisation des produits semi-directs au sous-groupes du groupe affine constitués d'une part par les translations et d'autre part par les transformations linéaires.
12. Appliquer la caractérisation des produits semi-directs aux sous-groupes de $O(2)$ constitués d'une part par $SO(2)$ et d'autre part par

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

13. On trouve dans les deux cas $\mathbb{Z}/2\mathbb{Z}$ comme centre et $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ comme quotient. Le groupe dérivé est le centre dans les deux cas.
14. a) Utiliser la table ou l'exercice 1.
b) Se reporter au chapitre 6 et à la section sur les quaternions.
15. Pour les groupes abéliens utiliser le théorème général. Pour les groupes non-abéliens, on montre que G est isomorphe à D_8 ou à Q . Pour distinguer les cas on pourra discuter en fonction du nombre d'éléments d'ordre 4. On montrera également qu'il y a toujours un sous-groupe d'ordre 4, que l'on appellera H . On distinguera suivant les cas selon que la projection $\pi : G \rightarrow G/H$ admet un homomorphisme réciproque s (i.e. $s \circ \pi = \text{Id}_{G/H}$) à gauche ou non.

III. Groupes abéliens, groupes abéliens de type fini

1. Montrer en particulier qu'il n'y a pas de systèmes libres à 2 éléments, puisque pour quels que soient les rationnels non nuls a et b on peut toujours trouver un entier k tel que $ka + kb = 0$. Pour la seconde question on fera une récurrence sur le nombre de générateurs. Pour la dernière la réponse est non, on considérera le sous-groupe engendré par 1 et $\sqrt{2}$.
4. Considérer la suite $(1, 2!, 3!, \dots, n!, \dots)$.

5. Écrire la forme générale d'un homomorphisme de $\mathbb{Z}^n$ dans $\mathbb{Z}$ et utiliser Bézout et le développement du déterminant.
6. Exprimer les images des générateurs par un automorphisme comme combinaison linéaires des générateurs. Puis procéder comme d'ordinaire en algèbre linéaire en vérifiant soigneusement les étapes.
7. a) Se reporter au chapitre 2, étudier l'action du groupe des automorphismes sur $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ et montrer qu'il laisse invariant le sous-ensemble des éléments non nuls, et donc en induit une permutation, or cet ensemble a 3 éléments.
 b) On considère le sous-groupe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. À un automorphisme de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ on associe sa restriction à ce sous-groupe. C'est un automorphisme de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. On montrera que l'image est d'ordre 2 en considérant l'action sur les générateurs de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$: on observera qu'on obtient des matrices triangulaires supérieures. Le noyau est isomorphe au groupe des automorphismes de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ laissant fixe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$: les automorphismes de $\mathbb{Z}/4\mathbb{Z}$, soit $\mathbb{Z}/2\mathbb{Z}$.
 c) On considère le sous-groupe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$. À un automorphisme de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ on associe sa restriction à ce sous-groupe. On montrera dans ce cas que l'homomorphisme est surjectif en faisant une construction explicite : un automorphisme de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ est donné par une matrice $(2, 2)$ à coefficients dans $\mathbb{Z}/2\mathbb{Z}$ de déterminant non-nul, un automorphisme de $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ est donné par une matrice $(2, 2)$ à coefficients dans $\mathbb{Z}/4\mathbb{Z}$ de déterminant inversible. Donc étant donnée une matrice $(2, 2)$ à coefficients dans $\mathbb{Z}/2\mathbb{Z}$ de déterminant non-nul on trouvera une matrice $(2, 2)$ à coefficients dans $\mathbb{Z}/4\mathbb{Z}$ de déterminant inversible dont la réduction modulo 2 est la matrice initiale. Pour ce qui est du noyau montrer qu'il est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ en considérant les matrices $(2, 2)$ à coefficients dans $\mathbb{Z}/4\mathbb{Z}$ laissant fixe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \subset \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$.

Chapitre 2

Actions de groupes Groupes symétriques

Ce chapitre est consacré à l'étude des actions de groupes sur des ensembles. Les groupes apparaissent le plus souvent dans ce contexte et non pour eux-mêmes. C'est le cas en géométrie : par exemple les groupes laissant invariants des polyèdres ou tout autre figure géométrique.

L'action d'un groupe sur lui-même par conjugaison est un outil puissant pour comprendre sa structure. On le verra à l'occasion de l'étude des théorèmes de Sylow. De ce point de vue, c'est également une méthode utile pour classer les « petits » groupes finis à isomorphisme près. On donnera des exemples dans le cours, et d'autres dans les exercices.

Enfin, le groupe symétrique mérite une étude à part, à cause de la grande variété de situations où il apparaît, et à cause de ses liens avec la combinatoire. On en décrira les classes de conjugaison, ainsi que des systèmes de générateurs.

1. Actions de groupes, exemples

Commençons par introduire les définitions de base, soient G un groupe et S un ensemble.

Définition 1

On appelle *action à gauche* de G sur S une application $\phi : G \times S \rightarrow S$ telle que :

- l'application $s \mapsto \phi(1, s)$ est l'application identité de S ,
- pour tous $g, g' \in G$, $s \in S$ on a :

$$\phi(gg', s) = \phi(g, \phi(g', s)).$$

Pour alléger les notations on écrira la plupart du temps $g.s$ ou gs à la place de $\phi(g, s)$.

On peut aussi définir aussi des actions à droite, ce ne sera pas utile ici et comme dans la suite on considérera uniquement le cas des actions à gauche on dira simplement « action ».

On va donner maintenant une définition équivalente à la définition 1. Considérons donc une action d'un groupe G sur un ensemble S . Soit $g \in G$, définissons une application

$$\phi_g : S \rightarrow S \quad \text{par} \quad \phi_g(s) = \phi(g, s).$$

Observons d'abord que l'application ϕ_g est une bijection. On sait que ϕ_1 est l'identité. La seconde condition de la définition 1 où l'on substitue à l'élément g' l'élément g^{-1} montre que $\phi_g \circ \phi_{g^{-1}}$ est l'identité de E . De même on montre que $\phi_g^{-1} \circ \phi_g = \phi_1$. Donc les applications ϕ_g et $\phi_{g^{-1}}$ sont réciproques l'une de l'autre.

Proposition 1. *L'application $g \mapsto \phi_g$ est un homomorphisme de G dans $\mathcal{P}(S)$. On l'appellera l'homomorphisme associé.*

Démonstration. On observe que la deuxième condition de la définition 1 se réécrit $\phi_{gg'} = \phi_g \circ \phi_{g'}$. Ce qu'on a en déjà utilisé dans le cas où $g' = g^{-1}$. Détaillons un peu, soit $s \in S$ on a :

$$\phi_{gg'}(s) = \phi(gg', s) = \phi(g, \phi(g', s)) = \phi_g(\phi_{g'}(s)).$$

ce qui est bien le résultat cherché. ■

Considérons maintenant la situation réciproque : soient de nouveau G un groupe et S un ensemble. Supposons donné un homomorphisme ϕ de G dans $\mathcal{P}(S)$. La formule $\bar{\phi}(g, s) = \phi(g)(s)$ détermine une action à gauche de G sur S . En effet on a :

$$\bar{\phi}(1, s) = \phi(1)(s) = \text{Id}_S(s) = s$$

et :

$$\bar{\phi}(gg', s) = \phi(gg')(s) = \phi(g)(\phi(g')(s)) = \bar{\phi}(g, \bar{\phi}(g', s)),$$

pour tous $g, g' \in G$ et $s \in S$.

Nous allons maintenant décrire des exemples fondamentaux.

Soit S un ensemble ; le groupe $\mathcal{P}(S)$ agit sur S par l'application $\mathcal{P}(S) \times S \rightarrow S$ donnée par

$$(\phi, s) \mapsto \phi(s) \quad \phi \in \mathcal{P}(S), \quad s \in S.$$

L'homomorphisme associé est l'identité de $\mathcal{P}(S)$.

Soit G un sous-groupe de $\mathcal{P}(S)$. L'application $\mathcal{P}(S) \times S \rightarrow S$ restreinte à $G \times S$ détermine une action de G sur S . L'homomorphisme associé de G dans $\mathcal{P}(S)$ est l'inclusion de G dans $\mathcal{P}(S)$. On dit que l'on restreint l'action à G . Par exemple si $\mathbb{K}$ est un corps et si $S = E$ est un $\mathbb{K}$ -espace vectoriel, on peut se restreindre au sous-groupe $\text{GL}(E)$ de $\mathcal{P}(S)$ constitué par les bijections $\mathbb{K}$ -linéaires. De même, si $\mathbb{K} = \mathbb{R}$ et si E est muni d'un produit scalaire, on peut se restreindre au sous-groupe $\text{O}(E)$ des transformations orthonormales.

Un autre exemple capital est celui de l'action des groupes symétriques sur les polynômes en plusieurs variables.

Soit $\mathbb{K}[X_1, \dots, X_n]$ l'algèbre des polynômes en n variables sur $\mathbb{K}$. Soient $\mathfrak{S}_n$ le groupe symétrique, $\sigma \in \mathfrak{S}_n$, et $P \in \mathbb{K}[X_1, \dots, X_n]$. Posons

$$\sigma P(X_1, \dots, X_n) = P(X_{\sigma(1)}, \dots, X_{\sigma(n)}).$$

Cette formule détermine une action à gauche de $\mathfrak{S}_n$ sur $\mathbb{K}[X_1, \dots, X_n]$. Les éléments fixes de cette action seront décrits au chapitre 3.

Action par translation

Soit G un groupe, on le fait agir sur lui-même par translation à gauche, en définissant $\phi : G \times G \rightarrow G$ par $\phi(g, h) = gh$. L'application ϕ_g est donnée par $\phi_g(h) = gh$. Cette action s'appelle l'action de translation à gauche de G sur lui-même.

L'homomorphisme ainsi déterminé de G dans $\mathcal{P}(G)$ est injectif. En effet si on a $\phi_g = \text{Id}_G$, c'est-à-dire $\phi_g(h) = h$ pour tout $h \in G$ cela entraîne que $g = e$.

Supposons G fini et soit $\#G = n$. Le groupe $\mathcal{P}(G)$ est isomorphe au groupe symétrique $\mathfrak{S}_n$ (voir section 4). On obtient :

Théorème 1 (Cayley). *Tout groupe fini est isomorphe à un sous-groupe d'un groupe symétrique.*

Démonstration. En effet on a démontré qu'il existait un homomorphisme injectif de G dans $\mathcal{P}(G) \cong \mathfrak{S}_n$; le groupe G est isomorphe à son image. ■

On peut généraliser la construction précédente de la manière suivante. Soient G un groupe et H un sous-groupe. L'application de

$$\begin{aligned} G \times G/H &\longrightarrow G/H, \\ (g, g'H) &\longmapsto gg'H \end{aligned}$$

détermine une action de G sur G/H . Soit ϕ_H l'homomorphisme associé de G dans $\mathcal{P}(G/H)$.

Il est intéressant dans ce cas d'étudier le noyau de l'homomorphisme associé :

Proposition 2. *Le noyau de l'homomorphisme associé ϕ_H est le sous-groupe $\bigcap_{g \in G} gHg^{-1}$.*

Démonstration. Soit k un élément du noyau. Pour tout élément $g \in G$ on a $kgH = gH$, soit $k \in gHg^{-1}$. Réciproquement soit x appartenant à $\bigcap_{g \in G} gHg^{-1}$. Pour tout élément g de G on a

$$x \in gHg^{-1},$$

donc $x = ghg^{-1}$ pour un certain $h \in H$. Donc on a $xgH = ghg^{-1}gH = ghH = gH$. On en conclut que $xgH = gH$ pour tout $g \in G$, donc que $x \in \text{Ker}(\phi_H)$. ■

Le groupe $\bigcap_{g \in G} gHg^{-1}$ est distingué par construction. C'est en fait le plus grand sous-groupe distingué de G contenu dans H : tout sous-groupe distingué de G contenu dans H est contenu dedans.

Action par conjugaison d'un groupe sur lui-même

Soit G un groupe, l'application suivante $G \times G \rightarrow G$, $(g, h) \mapsto ghg^{-1}$ définit une action de G sur lui-même, appelée action par conjugaison. L'homomorphisme de G dans $\mathcal{P}(G)$ associé est la composition de l'homomorphisme de G dans $\text{Aut}(G)$ qui à g associe l'automorphisme intérieur i_g , suivit de l'inclusion de $\text{Aut}(G)$ dans $\mathcal{P}(G)$.

Définition 2

Soit G un groupe et soient x et y deux éléments de G . On dit que x et y sont conjugués dans G s'il existe $g \in G$ tel que $gxg^{-1} = y$.

Proposition 3. *La relation de conjugaison est une relation d'équivalence dans G .*

Démonstration. La réflexivité est claire, on prend $g = 1$. Pour la symétrie, si on a $gxg^{-1} = y$ on a $g^{-1}yg = x$. Enfin pour la transitivité si on a $gxg^{-1} = y$ et $g'yg'^{-1} = z$ on a $gg'x(gg')^{-1} = z$. ■

Terminons cette section par des définitions.

Définition 3

L'ensemble des éléments conjugués à un élément x de G est appelé sa *classe de conjugaison*.

Par exemple dans un groupe abélien toute classe de conjugaison est réduite à un élément. Dans le cas du groupe D_8 , reprenant les notations du chapitre 1, on trouve comme classes de conjugaison $\{e\}$, $\{a^2\}$, $\{a, a^3\}$, $\{b, ba^2\}$, et $\{ba, ba^3\}$.

Définition 4

Soient A et B deux parties de G on dit qu'elles sont conjuguées s'il existe $g \in G$ tel que $gAg^{-1} = B$.

Soit E un ensemble sur lequel un groupe G agit, et soit x un élément de E , et A une partie de E :

Définition 5

On dit que x est un point fixe de l'action si : $\forall g \in G \quad gx = x$.

Le *fixateur* de A dans G est défini et noté par :

$$\text{Fix}_G(A) = \{g \in G \mid \forall a \in A, ga = a\},$$

Le stabilisateur de A dans G est défini et noté par :

$$\text{Stab}_G(A) = \{g \in G \mid gA = A\}.$$

Le *stabilisateur* d'une partie A contient, mais est (en général) différent du fixateur. On dit qu'un élément du stabilisateur de A laisse A fixe globalement, alors qu'un élément du fixateur laisse A fixe point par point.

Proposition 4. Les sous-ensembles de la définition précédente sont des sous-groupes.

Démonstration. La démonstration est facile, faisons la pour le cas du stabilisateur. Si on a $g, g' \in \text{Stab}_G(A)$ on veut montrer que $gg' \in \text{Stab}_G(A)$. On a $(gg')A = g(g'A) = gA = A$. De même si $g \in \text{Stab}_G(A)$ on a $g^{-1}A = g^{-1}(gA) = (g^{-1}g)A = A$. ■

Voici deux exemples de stabilisateur.

Si on considère l'ensemble $\{1, \dots, n\}$ l'ensemble de ses bijections est le groupe symétrique $\mathfrak{S}_n$ (section 4). Le stabilisateur du point n est le groupe symétrique $\mathfrak{S}_{n-1}$.

Soit maintenant $\text{GL}_n(\mathbb{R})$ le groupe des matrices (n, n) inversibles à coefficients dans $\mathbb{R}$. Il agit sur $\mathbb{R}^n$. Le stabilisateur du vecteur $(1, 0, \dots, 0)$ est constitué par les matrices de la forme

$$\begin{pmatrix} 1 & a_{1,2} & \dots & a_{1,n} \\ 0 & & & \\ \vdots & & (B) & \\ 0 & & & \end{pmatrix},$$

où B est une matrice $(n-1, n-1)$ inversible. On notera que l'ensemble quotient du groupe linéaire par ce sous-groupe est en bijection avec $\mathbb{R}^n - \{0\}$.

2. Orbites, ensembles transitifs, décomposition en orbites

Nous allons maintenant étudier de plus près l'action d'un groupe sur un ensemble. En particulier on décomposera l'ensemble en sous-ensembles « stables » par l'action de G . Soit donc G un groupe agissant sur un ensemble E , soit e un élément de E .

Définition 1

On appelle *orbite* $\mathcal{O}$ de e sous l'action du groupe G l'ensemble $\{ge \mid g \in G\}$.

Par construction, l'orbite $\mathcal{O}$ de e est stable sous l'action de G . C'est-à-dire que, pour tout élément x dans $\mathcal{O}$ et tout $g \in G$, l'élément gx est encore dans $\mathcal{O}$. En effet, si $p = g'e$ on a $gp = (gg')e$. Ceci veut dire que l'orbite elle-même est un ensemble muni d'une action de G .

Par exemple, dans le cas de l'action de translation de G sur lui-même ou celui de l'action par translation sur un ensemble quotient G/H , l'orbite de tout point est l'ensemble tout entier. En effet on a la relation suivante pour tous $g, g' \in G$: $g'g^{-1}(gH) = g'H$.

Dans le cas de l'action par conjugaison, l'orbite d'un élément est sa classe de conjugaison.

Si on considère un espace vectoriel E sur un corps $\mathbb{K}$, l'orbite sous l'action de $\text{GL}(E)$ d'un vecteur v est réduite à $\{0\}$ si $v = 0$, et $E - \{0\}$ si $v \neq 0$.

Revenons au cas général d'un ensemble E muni d'une action d'un groupe G . On définit une relation sur E en disant que x est en relation avec y si et seulement si y est dans l'orbite de x . Cette relation est une relation d'équivalence.

Elle est réflexive car $ex = x$; symétrique car $y = gx$ implique $x = g^{-1}y$; enfin elle est transitive car $y = gx$ et $z = g'y$ impliquent $z = g'gx$. C'est donc une relation d'équivalence dont les classes d'équivalence sont les orbites. On en déduit que :

Proposition 1. *Les orbites de E sous l'action de G constituent une partition de G .*

Définition 2

Si l'ensemble E sous l'action de G a une seule orbite on dit qu'il est *transitif*.

Par exemple l'action par translation d'un groupe G sur lui-même ou sur un quotient G/H est transitive.

En conclusion un ensemble E sur lequel agit un groupe G est réunion de sous-ensembles transitifs sous l'action de G , à savoir les orbites.

Structure des ensembles transitifs, applications

On va maintenant décrire les ensembles transitifs. Soient G un groupe, E, E' deux ensembles munis d'une action de G . On va définir une notion d'isomorphisme entre ces ensembles.

Définition 3

On dit que E et E' sont isomorphes s'il existe une application $f : E \rightarrow E'$ telle que :

- f est une bijection,
- pour tous $g \in G, s \in E$ on a $f(gs) = gf(s)$.

La seconde condition nous dit que f est compatible aux actions dans les deux membres. Étant donnée la décomposition en orbites décrite plus haut ceci réduit l'étude de tous les ensembles munis d'une action d'un groupe G à l'étude des ensembles de la forme G/H , comme le montre le théorème ci-dessous.

Théorème 1. Soit E un ensemble muni d'une action transitive du groupe G . Alors E est isomorphe, comme ensemble muni d'une action de G , à un ensemble de la forme G/H muni de l'action par translation, où H est le fixateur d'un point quelconque de E .

Démonstration. Soit s un point de E , et soit H le fixateur de s . On définit une application π de G dans E par $\pi(g) = gs$. Par définition de H , cette application factorise au travers de l'espace quotient G/H . En effet deux éléments g et g' ont même image si et seulement si $g^{-1}g'$ est élément du fixateur de s , soit H . Soit $\bar{\pi} : G/H \rightarrow E$ l'application induite. Cette application est surjective par hypothèse de transitivité de l'action de G sur E . Elle est injective par construction, en effet si on a $gs = g's$ cela veut dire que $g^{-1}g's = s$, donc que $g^{-1}g'$ est élément du stabilisateur de s dans G , donc $gH = g'H$. Enfin l'application ainsi définie vérifie bien $\bar{\pi}(gg'H) = gg's = g(\pi(g'H))$. ■

Considérons maintenant le cas particulier où E et G sont finis. On a alors des propriétés intéressantes sur les cardinaux.

Corollaire. Si l'ensemble E est transitif sous l'action de G on a :

$$\# E = \frac{\# G}{\# \text{Fix}_G(s)},$$

où s est un point quelconque de E . En particulier l'entier $\# E$ divise $\# G$.

Ceci montre en particulier que le cardinal du fixateur d'un point d'un ensemble transitif E ne dépend pas de ce point.

On a plus :

Proposition 2. Soient G, E un ensemble muni d'une action transitive de G . Soient $s, s' \in E$, alors $\text{Fix}_G(s)$ et $\text{Fix}_G(s')$ sont conjugués.

Démonstration. Les détails de la démonstration sont donnés à faire en exercice. On se contente d'indiquer que si $g \in \text{Fix}_G(s)$ et si $s' = hs$ pour un certain élément $h \in G$ alors $hgh^{-1} \in \text{Fix}_G(s')$. ■

Ce qui implique qu'ils ont même cardinal, car deux parties conjuguées ont même nombre d'éléments.

Revenons au cas d'un ensemble E muni d'une action de G non nécessairement transitive. Soit $\{\mathcal{O}_i \mid i \in I\}$ l'ensemble des orbites. Supposons que $\mathcal{O}_i = Gs_i, s_i \in E$. Comme les orbites constituent une partition de E , on obtient :

$$\# E = \sum_{i \in I} \frac{\# G}{\# \text{Fix}_G(s_i)}.$$

En application de ce qui précède, on va donner une démonstration du petit théorème de Fermat. Étant donné un nombre premier p , on veut démontrer que pour tout entier $a \in \mathbb{N}$, $a^p - a$ est divisible par p . Soit donc A un ensemble fini ayant a éléments. Considérons l'ensemble A^p et munissons-le d'une action de $\mathbb{Z}/p\mathbb{Z}$ de la manière suivante. Soit $(a_1, \dots, a_p)$ un élément de A^p . Soit maintenant $\bar{k}$ un élément de $\mathbb{Z}/p\mathbb{Z}$ posons :

$$\bar{k}(a_1, \dots, a_p) = (a_{1+\bar{k}}, \dots, a_{p+\bar{k}}).$$

Dans la formule précédente on identifie une classe de congruence modulo p , $\bar{u}$, avec son représentant compris entre 1 et p . On vérifie que ceci détermine bien une action de $\mathbb{Z}/p\mathbb{Z}$ sur A^p . Les orbites ont soit p éléments, soit 1 élément comme p est premier et que le cardinal d'une orbite divise p . L'ensemble A^p est donc réunion d'orbites à un élément (les points fixes) et d'orbites à p éléments. Il en résulte que le cardinal de A^p , soit a^p , est congru modulo p au cardinal de l'ensemble des points fixes. Or l'ensemble des points fixes est l'ensemble des éléments diagonaux, c'est-à-dire de la forme $(x, \dots, x)$. Il y en a clairement a . Soit le résultat cherché.

Formule des classes

Considérons maintenant le cas où l'action est la conjugaison d'un groupe G sur lui-même. Soit $g \in G$, dans ce cas, le sous-groupe $\text{Fix}_G(g)$ est l'ensemble des éléments h du groupe G tels que $hgh^{-1} = g$, c'est-à-dire tels que $hg = gh$, c'est-à-dire qui commutent avec g . On l'appelle le *centralisateur* de g et on le note $C_G(g)$.

Rappelons que l'ensemble des éléments de G , commutant avec tous les éléments de G est un sous-groupe distingué de G appelé le centre de G et est noté $Z(G)$.

La classe de conjugaison d'un élément du centre est réduite à lui-même. Soit maintenant $\{g_i \mid i \in I\}$, I étant un ensemble d'indices, une famille d'éléments de G telle que :

- g_i n'est pas élément du centre,
- si $i \neq j$ g_i et g_j ne sont pas dans la même classe de conjugaison,
- pour toute classe de conjugaison C non réduite à un élément (qui n'est pas celle d'un élément du centre) il existe $i \in I$ tel que $g_i \in C$.

Les éléments g_i constituent donc un système de représentants des classes de conjugaison de cardinal strictement plus grand que 1. La formule se réécrit alors :

$$\#G = \#Z(G) + \sum_{i \in I} \frac{\#G}{\#C_G(g_i)}.$$

La première partie de la somme est le nombre d'orbites à un élément. Dans la seconde partie tous les entiers qui interviennent sont strictement plus grands que 1. Cette formule est connue sous le nom de *formule des classes*. Insistons bien sur le résultat suivant :

Proposition 3. *Soit G un groupe fini, et soit $g \in G$ quelconque. Le cardinal de G est égal au produit du nombre d'éléments dans la classe de conjugaison de g par le cardinal du centralisateur de g dans G .*

Donnons quelques exemples.

Si on considère un groupe abélien, alors chaque classe de conjugaison est réduite à un élément, donc la formule des classes est triviale.

Considérons le groupe Q , dont la table est donnée dans le chapitre 1. Le centre est constitué par 1 et -1 . Le centralisateur de l'élément i est le sous-groupe constitué par 1, -1 , i et $-i$. La classe de conjugaison de i est donc constituée par i et $-i$. Le résultat est analogue pour j et k . Donc la formule des classes s'écrit :

$$8 = 1 + 1 + 2 + 2 + 2.$$

Les deux premiers termes de la somme correspondent au centre.

Dans le cas de D_8 , dont la table est également donnée dans le chapitre 1, le centre est constitué par 1 et a^2 . La classe de conjugaison de a est constituée par a et a^3 , la classe de conjugaison de b est constituée de b et ba^2 , enfin celle de ba est constituée par ba et ba^3 . L'équation des classes est donc encore :

$$8 = 1 + 1 + 2 + 2 + 2.$$

Anticipant l'analyse du groupe symétrique, qui sera faite après, voici l'équation des classes pour S_3 . Le centre est réduit à l'élément neutre. Il y a 3 éléments d'ordre 2 tous conjugués et 2 éléments d'ordre 3 conjugués. L'équation des classes s'écrit donc :

$$6 = 1 + 3 + 2.$$

Voici une application classique de la formule des classes :

Théorème 2. *Soit G un groupe non réduit à un élément, dont l'ordre est une puissance d'un nombre premier p . Alors le centre de G contient au moins p éléments.*

Démonstration. En effet si on applique la formule des classes à G , on constate que la somme de droite est divisible par p , étant égale à l'ordre de G . Par ailleurs, tous les termes de la seconde partie de la somme le sont aussi, étant des quotients de l'ordre de G non égaux à 1. Donc l'ordre du centre est divisible par p , étant plus grand que 1 il est au moins égal à p . ■

Corollaire 1. *Soit p un nombre premier, tous les groupes d'ordre p^2 sont abéliens.*

Démonstration. En effet si le centre du groupe n'est pas égal à G il est d'ordre p . Le quotient de G par son centre est alors d'ordre p donc cyclique, l'exercice II. 7 du chapitre 1 montre alors que G est abélien. ■

Voici une autre application. Soit G un groupe dont l'ordre est une puissance p^n d'un nombre premier p , alors pour tout entier $h < n$ il existe au moins un sous-groupe H de G dont l'ordre est p^h . Les détails sont donnés à faire en exercice. L'idée est de faire une récurrence sur l'entier n en utilisant le théorème.

Action par conjugaison sur l'ensemble des sous-groupes

Soit maintenant G un groupe fini, et soit $\mathcal{S}$ l'ensemble de ses sous-groupes. Si H est un élément de $\mathcal{S}$ et si g est un élément de G , le sous-groupe conjugué gHg^{-1} est dans $\mathcal{S}$. Ceci détermine donc une action de G sur $\mathcal{S}$. Par ailleurs il est clair qu'un sous-groupe H et un conjugué gHg^{-1} ont même cardinal. Ceci implique :

Proposition 4. *Soit d un entier divisant le cardinal de G . Le sous-ensemble $\mathcal{P}_d$ de $\mathcal{S}$ constitué par les sous-groupes de cardinal d , est stable sous l'action de G par conjugaison. Cela définit donc une action de G sur $\mathcal{P}_d$.*

Observons qu'un sous-groupe est distingué dans G si et seulement si il est égal à tous ses conjugués, autrement dit si son orbite sous l'action de G par conjugaison est réduite à lui-même. Plus généralement l'orbite d'un sous-groupe étant un ensemble transitif sous l'action de G son cardinal, c'est-à-dire le nombre de conjugués du sous-groupe, est un diviseur de G .

Par exemple, si on considère le groupe symétrique $\mathfrak{S}_3$, on constate facilement que l'ensemble des sous-groupes d'ordre 2 (il y en a 3, correspondant aux 3 transpositions distinctes) est transitif sous l'action de $\mathfrak{S}_3$. Il s'agit là évidemment d'un cas particulier du théorème de Sylow.

3. Théorèmes de Sylow

Le théorème de Lagrange dit que l'ordre d'un sous-groupe d'un groupe fini est un diviseur de l'ordre du groupe. Inversement il est faux que, si on s'est donné un diviseur de l'ordre du groupe, l'on puisse trouver un sous-groupe dont l'ordre est ce diviseur, on verra à ce propos l'exercice 2 dans la section portant sur les groupes symétriques. Néanmoins le résultat est vrai si on se restreint à des diviseurs qui sont une puissance d'un nombre premier p .

Définition

Soit p^a la plus grande puissance de p divisant le cardinal de G , on appelle p -sous-groupe de Sylow de G un sous-groupe dont l'ordre est p^a .

Le premier résultat est un théorème d'existence :

Théorème 1 (premier théorème de Sylow). Soit G un groupe fini, et soit p un nombre premier divisant $\#G$. Le groupe G a au moins un p -sous-groupe de Sylow.

L'importance des sous-groupes de Sylow, tient entre autres, à ce qu'ils permettent de comprendre la structure des groupes. L'étude de leurs propriétés fournit de très beaux exemples d'actions de groupes.

Démonstration. Elle se fait par récurrence sur l'ordre du groupe. Le cas initial étant évident. Admettons le résultat pour tous les groupes d'ordre strictement inférieur à celui de G . Soit p^a la plus grande puissance de p divisant $\#G$, on considère deux cas : ou il existe un sous-groupe strict (distinct de G) de G dont l'ordre est divisible par p^a ou non.

Dans le premier cas par hypothèse de récurrence ce sous-groupe a un p -sous-groupe de Sylow, ce p -sous-groupe de Sylow en est un aussi pour G . Ce cas est donc traité.

Dans le second cas tous les sous-groupes stricts de G ont un ordre non divisible par p^a , donc pour tout sous-groupe strict H l'entier $\frac{\#G}{\#H}$ est divisible par p .

Si on considère la formule des classes, on en conclut que l'entier $\#Z(G)$ est divisible par p . Le groupe $Z(G)$ étant commutatif, on peut appliquer le lemme 1 de la section 3 du chapitre 1, dit de Cauchy. Il y a donc au moins un sous-groupe E d'ordre p dans $Z(G)$. Un tel sous-groupe est distingué dans G . En effet tout sous-groupe du centre est distingué dans G .

Considérons alors le quotient G/E , et $\pi : G \rightarrow G/E$ l'homomorphisme canonique. Par hypothèse de récurrence, G/E a un p -sous-groupe de Sylow S' , qui est de cardinal p^{a-1} . Le sous-groupe $\pi^{-1}(S')$ est de cardinal p^a et est donc un p -sous-groupe de Sylow de G . ■

Notons le corollaire, démontré au cours du développement, et qui n'est autre que le lemme de Cauchy dans le cas général :

Corollaire. *Un groupe fini G dont l'ordre est divisible par un nombre premier p contient au moins un élément d'ordre p .*

Dans le chapitre 1 on a déjà décrit les sous-groupes de Sylow dans le cas des groupes abéliens finis. Voici deux exemples dans le cas non-abélien. Le lecteur doit si nécessaire commencer par se reporter aux sections suivantes concernant le groupe symétrique.

Commençons par fixer un nombre premier p quelconque. La cardinal du groupe symétrique $\mathfrak{S}_p$ est $p!$. Donc le cardinal d'un p -sous-groupe de Sylow est p . Le cycle d'ordre p

$$(1, \dots, p)$$

engendre un sous-groupe d'ordre p . C'est un sous-groupe de Sylow.

Le second exemple concerne le groupe alterné $\mathcal{A}_4$. Il a 12 éléments. On vérifie facilement que l'ensemble constitué de l'élément neutre et des permutations suivantes :

$$(1, 2)(3, 4) \quad (1, 3)(2, 4) \quad (1, 4)(2, 3)$$

est un sous-groupe de $\mathcal{A}_4$. En fait, c'est l'ensemble constitué par les éléments d'ordre 2 et l'élément neutre. Il a 4 éléments c'est bien un 2-sous-groupe de Sylow. En fait, il est distingué dans $\mathcal{A}_4$.

Le théorème suivant montre comment les p -sous-groupes de Sylow sont reliés les uns aux autres ainsi qu'aux sous-groupes dont l'ordre est une puissance de p .

Théorème 2 (deuxième et troisième théorèmes de Sylow).

Soit G un groupe fini, et soit p un nombre premier divisant $\#G$.

- *Tout sous-groupe dont l'ordre est une puissance de p est contenu dans un p -sous-groupe de Sylow,*
- *deux p -sous-groupes de Sylow sont conjugués entre eux,*
- *le nombre des p -sous-groupes de Sylow est congru à 1 modulo p .*

Démonstration. On va commencer par démontrer la première assertion. On choisit donc un sous-groupe H dont l'ordre est une puissance de p . On veut montrer qu'il est contenu dans un p -sous-groupe de Sylow.

La démonstration repose sur une analyse de l'action de H sur l'ensemble des p -sous-groupes de Sylow de G . Notons $\mathcal{P}_p$ cet ensemble, il est non vide à cause du théorème précédent. Le groupe G agit sur l'ensemble $\mathcal{P}_p$ par conjugaison, en effet tout conjugué d'un p -sous-groupe de Sylow est un p -sous-groupe de Sylow.

Considérons maintenant un élément S de $\mathcal{P}_Y$. Commençons par observer que l'orbite de S dans $\mathcal{P}_Y$ sous l'action de G a un ordre non divisible par p . En effet le fixateur de S dans G est le sous-groupe des éléments $g \in G$ tels que $gSg^{-1} = S$. Ce sous-groupe contient manifestement le sous-groupe S , donc son ordre est divisible par p^a . Par conséquent le cardinal de l'orbite de S , que l'on notera $\mathcal{O}$, est un diviseur de $\frac{\#G}{p^a}$, et n'est donc pas divisible par p .

Soit maintenant H un sous-groupe de G . Le sous-groupe de H agit aussi sur $\mathcal{P}_Y$ et toute orbite sous l'action de G de $\mathcal{P}_Y$ est stable sous l'action du sous-groupe H . En particulier l'orbite $\mathcal{O}$ est stable sous l'action de H .

Mais l'ensemble $\mathcal{O}$ sous l'action de H n'est pas nécessairement transitif. Supposons maintenant que l'ordre de H soit une puissance de p . Les orbites de $\mathcal{O}$ sous l'action de H , sont toutes de cardinal divisant celui de H , donc une puissance de p . Elles sont, soit réduite à un élément, soit de cardinal divisible par p . Comme on l'a vu plus haut, le cardinal de $\mathcal{O}$ n'est pas divisible par p . Il y a au moins une orbite sous l'action de H dont le cardinal n'est pas divisible par p . Donc, il existe une orbite qui est réduite à un élément. Ce qui veut dire qu'il existe un conjugué S' de S tel que pour tout $h \in H$ on a $hS'h^{-1} = S'$.

On va maintenant montrer que $H \subset S'$, ceci achèvera la démonstration.

Considérons l'ensemble HS' . Le théorème 4 de la section 2 du chapitre 1 implique que c'est un sous-groupe. En fait, les hypothèses du théorème 4 ne sont pas tout à fait réalisées. On sait que pour tout $h \in H$ on a $hS'h^{-1} = S'$. Si on reprend la démonstration du théorème avec $H = H$ et $K = S'$, on constate que cette hypothèse est suffisante pour faire la démonstration.

L'ordre de HS' est une puissance de p . En effet, S' est distingué dans HS' et donc le groupe quotient HS'/H est isomorphe au groupe $S'/H \cap S'$ toujours par le théorème 4, section 2 du chapitre 1, et donc on a :

$$\frac{\#(HS')}{\#(H)} = \#(S'/H \cap S'),$$

soit

$$\#(HS') = \#(H) \times \#(S'/H \cap S').$$

Or, le membre de droite est une puissance de p . On en conclut que $HS' = S'$. En effet, si H n'est pas contenu dans S' , celui-ci est sous-groupe strict de HS' . L'ordre du sous-groupe HS' est alors une puissance de p strictement supérieure à p^a . C'est impossible par hypothèse, donc $H \subset S'$.

En remplaçant H par un p -sous-groupe de Sylow on obtient le second énoncé.

Pour le troisième, considérons l'action de S sur $\mathcal{P}_Y$. L'orbite de S lui-même est réduite à un point, celle de tout autre élément de $\mathfrak{S}$ n'est pas réduite à un point, en effet l'argument utilisé plus haut montre que si l'orbite de S' est réduite à

un point, on a $S = S'$. L'orbite de tout $S' \neq S$ a donc un cardinal divisible par p puisque c'est un diviseur non-trivial de p^a . L'ensemble considéré est donc réunion disjointe d'une orbite réduite à un point et d'une famille d'orbites dont le cardinal est divisible par p , le résultat suit. ■

On va donner une application de ce qui précède à la structure des groupes qui ont pq éléments, p et q étant des nombres premiers distincts. Soit donc G un tel groupe, G a au moins un p -sous-groupe de Sylow K , et au moins un q -sous-groupe de Sylow L . Le nombre des conjugués de K est congru à 1 modulo p (théorème 3) et de plus il divise pq , donc il divise q . On a donc :

Proposition 1. *Si q n'est pas congru à 1 modulo p il n'y a qu'un seul p -sous-groupe de Sylow. Donc il est distingué.*

Notons que les remarques précédentes impliquent aussi que si $q < p$ le p -sous-groupe de Sylow est distingué. Car $lp + 1$ ne peut diviser q que si $l = 0$.

Le corollaire suivant est laissé en exercice.

Corollaire. *Sous les hypothèses de la proposition précédente le groupe G est un produit semi-direct.*

Plus précisément un groupe d'ordre 15 est commutatif. En effet il y a un seul 3-Sylow et un seul 5-Sylow. Par ailleurs, il n'y a pas d'homomorphisme non trivial de $\mathbb{Z}/5\mathbb{Z}$ dans $\text{Aut}(\mathbb{Z}/3\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$ donc on ne peut avoir de produit semi-direct non trivial.

Nous laissons en exercice la proposition qui suit.

Proposition 2. *Un groupe d'ordre $2p$, p premier, est, soit commutatif, soit diédral.*

4. Groupes symétriques, classes de conjugaison

Notons $[n]$ l'ensemble $\{1, \dots, n\}$, et notons $\mathfrak{S}_n$ son groupe de permutations. Dans ce cas on parlera aussi de substitutions. On l'appelle le *groupe symétrique* de l'ensemble à n éléments. La première observation importante est que ce groupe ne dépend pas de l'ensemble à n éléments choisi au départ, ce qui justifie la terminologie.

Proposition 1. *Soit E un ensemble à n éléments. Le groupe des permutations de E est isomorphe à $\mathfrak{S}_n$.*

Démonstration. Soit ϕ une bijection entre E et $[n]$. L'application de $\mathcal{P}(E)$ dans $\mathfrak{S}_n$ donnée par $f \mapsto \phi \circ f \circ \phi^{-1}$ admet pour réciproque $g \mapsto \phi^{-1} \circ g \circ \phi$. C'est donc une bijection. C'est un homomorphisme car :

$$(\phi \circ f \circ \phi^{-1}) \circ (\phi \circ g \circ \phi^{-1}) = \phi \circ f \circ g \circ \phi^{-1}.$$

■

Le premier résultat important sur le groupe symétrique est le calcul de son cardinal :

Théorème 1. *Le groupe symétrique de l'ensemble à n éléments a $n!$ éléments.*

Démonstration. La démonstration se fait par récurrence sur n . Si $n = 1$ le groupe symétrique $\mathfrak{S}_1$ a un seul élément. Supposons donc le résultat connu pour $n - 1$. Considérons le sous-groupe de $\mathfrak{S}_n$ constitué par les éléments σ tels que $\sigma(n) = n$. Il est isomorphe à $\mathfrak{S}_{n-1}$, il a donc $(n - 1)!$ éléments par hypothèse de récurrence.

On construit alors une bijection entre $\mathfrak{S}_n$ et $[n] \times \mathfrak{S}_{n-1}$ comme suit. À la permutation σ , on associe

$$(\sigma(n), \tau_{n, \sigma(n)} \sigma).$$

Dans cette formule $\tau_{n, \sigma(n)}$ désigne l'unique permutation qui échange n et $\sigma(n)$ et qui est fixe sur les autres éléments. Le second terme peut être identifié à un élément de $\mathfrak{S}_{n-1}$ car $\tau_{n, \sigma(n)} \sigma(n) = n$ et l'ensemble des permutations de $\mathfrak{S}_n$ satisfaisant à cette propriété est un sous-groupe qui s'identifie à $\mathfrak{S}_{n-1}$. Admettant que l'application est une bijection la démonstration se termine en calculant le nombre d'éléments $[n] \times \mathfrak{S}_{n-1}$, soit $n \times (n - 1)! = n!$. Revenons sur la bijectivité, et construisons une application réciproque. Considérons un élément $i \in [n]$ et une permutation quelconque $\phi \in \mathfrak{S}_{n-1}$. Soit $\tau_{i, n}$ la permutation qui échange i et n et qui laisse les autres éléments fixes. On vérifie que :

$$[n] \times \mathfrak{S}_{n-1} \longrightarrow \mathfrak{S}_n, (i, \phi) \longmapsto \tau_{i, n} \phi$$

est l'application réciproque de celle définie plus haut.

■

Un élément σ de $\mathfrak{S}_n$ sera noté

$$\begin{pmatrix} 1, & \dots, & n \\ \sigma(1), & \dots, & \sigma(n) \end{pmatrix}.$$

On va maintenant étudier certains éléments particuliers du groupe $\mathfrak{S}_n$. Introduisons d'abord une définition :

■ Définition 1

On appelle *support* d'un élément $\sigma \in \mathfrak{S}_n$ l'ensemble des $i \in \{1, \dots, n\}$ qui ne sont pas des points fixes de σ , c'est-à-dire l'ensemble des i tels que $\sigma(i) \neq i$.

Proposition 2. *Le support de σ est stable sous l'action du sous-groupe engendré par σ .*

Démonstration. Il suffit de vérifier que, si i est dans le support de σ , il en est de même pour $\sigma(i)$. Or, dire que i est dans le support de σ équivaut à dire que $\sigma(i) \neq i$. Comme σ est injective ceci est équivalent à $\sigma^2(i) \neq \sigma(i)$. Ce qui implique $\sigma(i)$ est dans le support de σ . On en conclut donc que σ induit une bijection du support dans lui-même. Ceci implique que le support de σ est un ensemble muni d'une action du sous-groupe $\langle \sigma \rangle$ engendré par l'élément σ . ■

Décompositions en cycles

Définition 2

Soit k un entier tel que $1 < k \leq n$. On appelle *k-cycle* un élément σ de $\mathfrak{S}_n$ dont le support S a k éléments et tel que l'orbite de tout élément du support sous l'action du sous-groupe $\langle \sigma \rangle$ soit le support tout entier.

La dernière condition est équivalente à l'une des conditions suivantes :

- il existe un élément $s \in S$ tel que l'ensemble $\{s, \sigma(s), \dots, \sigma^{k-1}(s)\}$ est de cardinal k et est donc égal à S ,
- le sous-groupe $\langle \sigma \rangle$ engendré par σ est isomorphe à $\mathbb{Z}/k\mathbb{Z}$.

La première condition est juste une reformulation de la condition de la définition dans un cas particulier. Par ailleurs le cas particulier implique la cas général. On le montre en remarquant d'abord que $\sigma^k(s) = s$ et que donc plus généralement $\sigma^i(s) = \sigma^{i+k}(s)$. En effet, sinon, on aurait $\sigma^k(s) = \sigma^i(s)$ pour $0 < i < k$, soit $\sigma^{k-i}(s) = s$ pour $0 < k-i < k$, ce qui est impossible. Il en résulte que, l'orbite d'un point quelconque, $\sigma^i(s)$ est égale à :

$$\{\sigma^i(s), \sigma^{i+1}(s), \dots, \sigma^{i+k-1}(s)\}$$

soit

$$\{s, \sigma(s), \dots, \sigma^{k-1}(s)\},$$

car $\sigma^{k+j}(s) = \sigma^j(s)$.

Pour ce qui est de la seconde condition, il suffit de montrer que les éléments $1, \sigma, \dots, \sigma^{k-1}$ sont deux à deux distincts. Ceci entraîne que $\# \langle \sigma \rangle = k$, le résultat suit car $\sigma^k = 1$. Mais l'argument précédent démontre la condition requise.

Quand le support est réduit à deux éléments on dit que le 2-cycle est une *transposition*.

Les permutations suivantes sont des transpositions :

$$\begin{pmatrix} 1, 2, 3, 4, 5 \\ 1, 3, 2, 4, 5 \end{pmatrix}, \quad \begin{pmatrix} 1, 2, 3, 4 \\ 2, 1, 3, 4 \end{pmatrix}, \quad \begin{pmatrix} 1, 2, 3, 4, 5, 6, 7 \\ 7, 2, 3, 4, 5, 6, 1 \end{pmatrix}.$$

Voici des exemples de 3-cycles :

$$\begin{pmatrix} 1, & 2, & 3 \\ 2, & 3, & 1 \end{pmatrix}, \quad \begin{pmatrix} 1, & 2, & 3 \\ 3, & 1, & 2 \end{pmatrix}, \quad \begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 3, & 2, & 5, & 4, & 1 \end{pmatrix}.$$

Voici deux exemples de 5-cycles :

$$\begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 2, & 3, & 4, & 5, & 1 \end{pmatrix}, \quad \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6 \\ 2, & 4, & 3, & 6, & 1, & 5 \end{pmatrix}.$$

Introduisons maintenant une notation pour un k -cycle σ . Soit $\alpha_1, \dots, \alpha_k$ le support. On suppose que $\sigma(\alpha_i) = \alpha_{i+1}$ si $i = 1, \dots, k-1$ et que $\sigma(\alpha_k) = \alpha_1$, en particulier les α_i ne sont pas supposés rangés dans l'ordre croissant. Le cycle sera noté $(\alpha_1, \dots, \alpha_k)$.

Voici un pas essentiel dans la description des classes de conjugaison de $\mathfrak{S}_n$:

Théorème 2. *Le conjugué d'un k -cycle est un k -cycle. Deux k -cycles quelconques sont conjugués.*

Démonstration. Démontrons la deuxième propriété d'abord. Considérons deux k -cycles quelconques σ et τ . Notons $\{\alpha_1, \dots, \alpha_k\}$ le support du premier, $\{\beta_1, \dots, \beta_k\}$ celui du second. On suppose que $\sigma(\alpha_i) = \alpha_{i+1}$ si $i = 1, \dots, k-1$ et que $\sigma(\alpha_k) = \alpha_1$; on suppose que $\tau(\beta_i) = \beta_{i+1}$ si $i = 1, \dots, k-1$ et respectivement que $\tau(\beta_k) = \beta_1$. On définit maintenant une permutation ϕ par $\phi(\alpha_i) = \beta_i$ pour tout i , ϕ est défini de manière arbitraire sur le complémentaire de $\alpha_1, \dots, \alpha_k$, i.e. comme une bijection quelconque du complémentaire de $\alpha_1, \dots, \alpha_k$ sur le complémentaire de $\beta_1, \dots, \beta_k$.

Calculons la permutation $\phi^{-1} \circ \tau \circ \phi$. Pour un élément α_i , $1 \leq i < k$, on a $\phi^{-1} \circ \tau \circ \phi(\alpha_i) = \phi^{-1} \circ \tau(\beta_i) = \phi^{-1}(\beta_{i+1}) = \alpha_{i+1}$. De même, on a $\phi^{-1} \circ \tau \circ \phi(\alpha_k) = \alpha_1$.

Pour un élément u qui est dans le complémentaire de $\{\alpha_1, \dots, \alpha_k\}$ on a $\phi^{-1} \circ \tau \circ \phi(u) = \phi^{-1} \circ \phi(u) = u$. Donc $\phi^{-1} \circ \tau \circ \phi = \sigma$, ce qui donne le résultat.

Pour ce qui est de la première propriété soit $c = (\alpha_1, \dots, \alpha_k)$ un k -cycle et ϕ une permutation. Alors la permutation $\phi \circ c \circ \phi^{-1}$ est le k -cycle $(\phi(\alpha_1), \dots, \phi(\alpha_k))$. ■

Théorème 3. *Soit $\sigma \in \mathfrak{S}_n$. Il existe une famille, unique à l'ordre près, de k -cycles c_i , $i = 1, \dots, t$ telle que :*

- $\sigma = c_1 \circ c_2 \circ \dots \circ c_t$;
- pour tous i, j on a $c_i \circ c_j = c_j \circ c_i$;
- les supports des cycles c_i sont deux à deux disjoints.

Cette décomposition est appelée la décomposition canonique en cycles de σ .

Dans la pratique on dira décomposition en cycles, sauf s'il y a ambiguïté.

Démonstration. Soit σ la permutation considérée, et soit $H = \langle \sigma \rangle$ le sous-groupe qu'elle engendre. Considérons la décomposition en orbites de l'ensemble $\{1, \dots, n\}$ sous l'action de H , soient $\mathcal{O}_1, \dots, \mathcal{O}_t$ les orbites non réduites à un élément. Définissons le cycle c_i comme étant égal à σ sur l'orbite $\mathcal{O}_i$ et l'identité sur le complémentaire. C'est un cycle d'ordre $\# \mathcal{O}_i$.

On a besoin du lemme suivant dont la démonstration est laissée au lecteur.

Lemme 1. *Deux cycles dont le support est disjoint commutent.*

De cela il résulte que la permutation σ est le produit des c_i pris dans un ordre arbitraire.

Inversement, soit une permutation σ qui s'écrit comme produit de cycles commutant deux à deux. Les orbites des éléments de $\{1, \dots, n\}$, non réduites à un élément, sous l'action de $\langle \sigma \rangle$ sont les supports des cycles considérés. Les cycles c_i sont donc uniquement déterminés comme étant les permutations qui sont restrictions de σ aux orbites de $\langle \sigma \rangle$, non réduites à un élément et l'identité sur le complémentaire de l'orbite. ■

Corollaire. *Deux permutations $\sigma, \sigma' \in \mathfrak{S}_n$ sont conjuguées si et seulement si, dans leur décomposition canonique en cycles, apparaissent le même nombre de k -cycles pour tout entier k , $2 \leq k \leq n$.*

Démonstration. Supposons que les deux permutations σ et σ' soient conjuguées par ϕ , conjuguons alors la décomposition de l'un d'eux. Rappelons que le conjugué d'un k -cycle est un k -cycle. Rappelons aussi que, si on conjugue par un même élément deux éléments qui commutent, les éléments conjugués commutent encore. Ceci implique que si $\sigma = c_1 \circ c_2 \circ \dots \circ c_t$ est la décomposition en cycles de σ et si $\sigma' = \phi \circ \sigma \circ \phi^{-1}$, alors $\sigma' = (\phi \circ c_1 \circ \phi^{-1}) \circ (\phi \circ c_2 \circ \phi^{-1}) \circ \dots \circ (\phi \circ c_t \circ \phi^{-1})$ est la décomposition en cycles de σ' .

Considérons maintenant l'implication réciproque. Soient $\mathcal{O}_1, \dots, \mathcal{O}_t$ les orbites non réduites à un élément de $\langle \sigma \rangle$, soient $\mathcal{O}'_1, \dots, \mathcal{O}'_{t'}$ les orbites non réduites à un élément de $\langle \sigma' \rangle$. Observons d'abord qu'il y en a autant, c'est-à-dire que $t = t'$, ceci résulte du fait qu'il y a autant de cycles dans les deux décompositions. Comme pour tout k , il y a autant de k -cycles dans les deux décompositions quitte à réordonner, on peut supposer que $\# \mathcal{O}_i = \# \mathcal{O}'_i$ pour tout i . On définit alors ϕ de la manière suivante : pour tout i , c'est une bijection de $\mathcal{O}_i$ dans $\mathcal{O}'_i$, définie comme dans la démonstration du théorème 2, et c'est une bijection quelconque du complémentaire de $\cup_i \mathcal{O}_i$ vers le complémentaire de $\cup_i \mathcal{O}'_i$. La vérification est alors la même que dans la démonstration du théorème 2. ■

On va maintenant relier ceci à une donnée purement combinatoire :

Définition 3

On appelle *partition d'un entier n* toute suite décroissante d'entiers positifs non nuls $(\lambda_1, \dots, \lambda_h)$ telle que

$$\lambda_1 + \dots + \lambda_h = n.$$

On remarquera qu'une partition de n comporte au plus n entiers. Voici toutes les partitions de 5 :

$$(5) \quad (4, 1) \quad (3, 2) \quad (3, 1, 1) \quad (2, 2, 1) \quad (2, 1, 1, 1) \quad (1, 1, 1, 1, 1).$$

On fera attention à ne pas confondre la notion de partition d'un entier n avec celle de partition de l'ensemble $\{1, \dots, n\}$. C'est-à-dire avec avec la donnée d'une famille de sous-ensembles A_i , $i \in I$ de sous-ensembles de $\{1, \dots, n\}$ dont la réunion est égale à $\{1, \dots, n\}$ et qui sont disjoints deux à deux. En fait si on a une telle partition la famille des entiers $\# A_i$ est une partition de l'entier n . Mais on vérifiera aisément que deux partitions distinctes de $\{1, \dots, n\}$ peuvent créer la même partition de n . C'est le cas pour les partitions $\{1\}, \{2, 3\}$ et $\{1, 2\}, \{3\}$ de $[3]$.

Proposition 3. *Le nombre de classes de conjugaison du groupe $\mathfrak{S}_n$ est le nombre de partitions de l'entier n .*

Démonstration. Soit σ une permutation, et soit $\mathcal{O}_1, \dots, \mathcal{O}_t$ les orbites non réduites à un élément de $\{1, \dots, n\}$ sous l'action de $\langle \sigma \rangle$. Quitte à réordonner, on peut supposer que les cardinaux décroissent. Associons à σ la partition de n suivante $(\# \mathcal{O}_1, \dots, \# \mathcal{O}_t, 1, \dots, 1)$ où il y a $n - \sum_i \# \mathcal{O}_i$ fois 1 à la suite de $(\# \mathcal{O}_1, \dots, \# \mathcal{O}_t)$. On a bien là une partition de n , qui ne dépend que de la classe de conjugaison de σ , d'après la décomposition en cycles des permutations et la caractérisation des classes de conjugaison. On a donc une application de l'ensemble des classes de conjugaison vers les partitions.

Inversement si on a une partition de n , soit $(\lambda_1, \dots, \lambda_u)$ avec

$$\lambda_1, \dots, \lambda_v > 1 \text{ et } \lambda_{v+1} = \dots = \lambda_u = 1$$

on lui associe la classe de conjugaison de la permutation suivante donnée par le produit de v cycles de longueur respective $\lambda_1, \dots, \lambda_v$ à supports disjoints. Les supports peuvent être choisis disjoints car $\lambda_1 + \dots + \lambda_v \leq n$. $\square$

Signature d'une permutation

La fin de cette section est consacrée à l'étude de la *signature* d'une permutation.

Définition 4

Soit $\sigma \in \mathfrak{S}_n$, et soient $i, j \in \{1, \dots, n\}$ tels que $i < j$. On dit que σ présente une *inversion* en (i, j) si $\sigma(i) > \sigma(j)$. On appelle nombre d'inversions de σ , et on note $\nu(\sigma)$, le nombre de couples (i, j) , $i < j$, tels que σ présente une inversion en (i, j) .

Définition 5

Soit $\sigma \in \mathfrak{S}_n$, on appelle *signature* de σ , et on note $\varepsilon(\sigma)$, l'entier suivant égal à 1 ou -1 :

$$\frac{\prod_{i < j} \sigma(i) - \sigma(j)}{\prod_{i < j} i - j}.$$

Cet entier est aussi égal à $(-1)^{\nu(\sigma)}$.

Proposition 4. L'application signature $\varepsilon : \mathfrak{S}_n \rightarrow \{+1, -1\}$ est un homomorphisme de groupes, l'ensemble $\{+1, -1\}$ étant muni de la multiplication.

Démonstration. Calculons $\varepsilon(\sigma \circ \tau)$. On a :

$$\begin{aligned} \varepsilon(\sigma \circ \tau) &= \frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} i - j} = \dots \\ &= \frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} \tau(i) - \tau(j)} \times \frac{\prod_{i < j} \tau(i) - \tau(j)}{\prod_{i < j} i - j} = \frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} \tau(i) - \tau(j)} \times \varepsilon(\tau). \end{aligned}$$

Mais par ailleurs

$$\frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} \tau(i) - \tau(j)} = \varepsilon(\sigma).$$

En effet τ détermine une bijection T de l'ensemble $P = \{(i, j) \mid i > j\}$ dans lui-même par la formule $T((i, j)) = (\tau(i), \tau(j))$ si $\tau(i) > \tau(j)$ et $T((i, j)) = (\tau(j), \tau(i))$ si $\tau(i) < \tau(j)$, on note P^+ le premier sous-ensemble de P , P^- le second ainsi

déterminés. On a :

$$\frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} \tau(i) - \tau(j)} = \frac{\prod_{(i < j) \in P^+} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{(i < j) \in P^+} \tau(i) - \tau(j)} \times \frac{\prod_{(i < j) \in P^-} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{(i < j) \in P^-} \tau(i) - \tau(j)},$$

soit

$$\frac{\prod_{i < j} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{i < j} \tau(i) - \tau(j)} = \frac{\prod_{\tau(i) < \tau(j)} \sigma \circ \tau(i) - \sigma \circ \tau(j)}{\prod_{\tau(i) < \tau(j)} \tau(i) - \tau(j)} \times \frac{\prod_{\tau(j) < \tau(i)} \sigma \circ \tau(j) - \sigma \circ \tau(i)}{\prod_{\tau(j) < \tau(i)} \tau(j) - \tau(i)}.$$

On reconnaît $\varepsilon(\sigma)$. ■

La signature d'une transposition est -1 , celle d'un k -cycle est $(-1)^{k+1}$.

Définition 6

Le noyau de l'homomorphisme signature est appelé le *groupe alterné*. On le note $\mathcal{A}_n$.

Proposition 5. Le groupe alterné $\mathcal{A}_n$ est distingué dans le groupe symétrique $\mathfrak{S}_n$. Son ordre est $\frac{n!}{2}$.

Démonstration. Le groupe quotient $\mathfrak{S}_n / \mathcal{A}_n$ est isomorphe à $\mathbb{Z}/2\mathbb{Z}$. Donc :

$$\# \mathcal{A}_n = \frac{\# \mathfrak{S}_n}{2}. \quad \blacksquare$$

On va terminer cette section en donnant une liste des classes de conjugaison, de l'ordre d'un élément dans la classe, du nombre d'éléments dans la classe pour les groupes $\mathfrak{S}_4$, et $\mathfrak{S}_5$. Le cas de $\mathfrak{S}_3$ ayant été donné plus haut. Les classes de conjugaison sont indexées comme il a été indiqué plus haut sur les partitions. La partition est donc indiquée sur la première colonne, l'ordre sur la seconde, le nombre d'éléments dans la classe dans la troisième.

partition	ordre	cardinal de la classe
(1, 1, 1, 1)	1	1
(2, 1, 1)	2	6
(2, 2)	2	3
(3, 1)	3	8
(4)	4	6

Voici quelques indications sur cette table. À la partition $(1, 1, 1, 1)$ correspond évidemment l'élément neutre. À la partition $(2, 1, 1)$ correspondent les transpositions. Il y en a autant que de sous-ensembles à 2 éléments, soit 6. À la partition $(2, 2)$ correspondent les produits de transpositions à supports disjoints, il y en a 3, $(1, 2)(3, 4)$, $(1, 3)(2, 4)$, et $(1, 4)(2, 3)$. À la partition $(3, 1)$ correspondent les 3-cycles. Il y en a 8, en effet pour déterminer un 3-cycle on commence par déterminer le support, on a 4 choix possibles, puis sur ce support on a deux choix possibles correspondant aux deux 3-cycles de $\mathfrak{S}_3$. Enfin à (4) correspondent les 4-cycles. Il y en a 6. En fait (exercice !) il y a $(n - 1)!$ n -cycles dans $\mathfrak{S}_n$.

Voici la même table pour $\mathfrak{S}_5$.

partition	ordre	cardinal de la classe
$(1, 1, 1, 1, 1)$	1	1
$(2, 1, 1, 1)$	2	10
$(2, 2, 1)$	2	15
$(3, 1, 1)$	3	20
$(3, 2)$	6	20
$(4, 1)$	4	30
(5)	5	24

5. Groupes symétriques, générateurs, simplicité du groupe alterné

Dans cette section, on va donner des compléments sur les groupes symétriques. On va commencer par étudier les systèmes de générateurs du groupe symétrique et du groupe alterné. Puis, on étudiera les sous-groupes distingués du groupe symétrique et du groupe alterné.

Théorème 1. *Les cycles engendrent le groupe symétrique $\mathfrak{S}_n$.*

Démonstration. C'est une conséquence directe de la décomposition en cycles, c'est-à-dire du théorème 3 de la section 4. □

On peut préciser ce résultat. Notons $\tau_{i,j}$, $i \neq j$ la transposition (i, j) , et τ_i la transposition $(i, i + 1)$, $1 \leq i \leq n - 1$.

Théorème 2. Les transpositions τ_i , $1 \leq i \leq n-1$, forment un système générateur de $\mathfrak{S}_n$.

Démonstration. Notons H_n le sous-groupe de $\mathfrak{S}_n$ engendré par les transpositions τ_i , $1 \leq i \leq n-1$. On raisonne par récurrence sur n . Le cas $n=1$ est immédiat.

Passons du cas $n-1$ au cas n . Le groupe $\mathfrak{S}_{n-1}$ s'identifie au sous-groupe de $\mathfrak{S}_n$ constitué par les éléments laissant fixes n , on sait par récurrence qu'il est engendré par les transpositions τ_i , $1 \leq i \leq n-2$. Soit $\sigma \in \mathfrak{S}_n$, si on a $\sigma(n) = n$ alors $\sigma \in \mathfrak{S}_{n-1}$. Alors σ est dans le sous-groupe engendré par les τ_i , $1 \leq i \leq n-2$, donc dans le sous-groupe engendré par les τ_i , $1 \leq i \leq n-1$. Sinon, soit $\sigma(n) = k$, $k \neq n$, et soit $\mu \in \mathfrak{S}_{n-1}$ tel que $\mu(k) = n-1$. Alors $\tau_{n-1} \circ \mu \circ \sigma(n) = \tau_{n-1} \circ \mu(k) = \tau_{n-1}(n-1) = n$ et donc, on a $\tau_{n-1} \circ \mu \circ \sigma \in \mathfrak{S}_{n-1}$. Comme $\tau_{n-1} \circ \mu \in H_n$ et que $\mathfrak{S}_{n-1} \subset H_n$ on conclut que $\sigma \in H_n$. Ce qui démontre le résultat. $\square$

Corollaire. Les 3-cycles constituent un système de générateurs de $\mathcal{A}_n$.

Démonstration. Considérons une permutation paire μ , et écrivons une décomposition en produit de transpositions : $\mu = t_1 t_2 \cdots t_k$. Observons d'abord que l'entier k est pair en effet la signature de μ est $(-1)^k$. On a alors besoin du :

Lemme 1. Le produit de deux transpositions dont le support a en commun un élément est un 3-cycle.

Démonstration. Le produit de $\tau_{j,k}$ par $\tau_{i,j}$ est le 3-cycle (i, k, j) . $\square$

Revenons à la démonstration du corollaire et considérons un sous-produit $t_{2k-1} t_{2k}$ dans la décomposition de μ en produit de transpositions. Si les deux transpositions sont identiques ce sous-produit disparaît. Si les supports des deux transpositions ont un élément et un seul en commun le sous-produit est un 3-cycle. Enfin si les supports des deux transpositions sont d'intersection vide introduisons une transposition $\tau_{i,j}$ où i appartient au support de t_{2k-1} et j appartient au support de t_{2k} . Alors $t_{2k-1} t_{2k} = t_{2k-1} \tau_{i,j} \tau_{i,j} t_{2k}$, et $t_{2k-1} \tau_{i,j}$ et $\tau_{i,j} t_{2k}$ sont des 3-cycles. On en conclut que tout produit d'un nombre pair de transpositions est un produit de 3-cycles. Le résultat suit. $\square$

Pour finir cette section on va démontrer le théorème suivant dû à Galois.

Théorème 3. Soit n un entier supérieur ou égal à 5. Alors le groupe alterné $\mathcal{A}_n$ n'a pas de sous-groupes distingués distincts de $\{1\}$ et de lui-même.

La démonstration va être coupée en plusieurs étapes. On va d'abord étudier la question des sous-groupes distingués de $\mathfrak{S}_n$.

Proposition 1. *Le groupe alterné $\mathcal{A}_n$ est engendré par les carrés d'éléments de $\mathfrak{S}_n$.*

Démonstration. Cela résulte du corollaire du théorème 2. En effet tout 3-cycle c est un carré car $c = c^4 = (c^2)^2$. Par ailleurs tout carré est de signature 1 donc élément de $\mathcal{A}_n$. ■

Proposition 2. *Le groupe alterné $\mathcal{A}_n$ est stable par tout automorphisme de $\mathfrak{S}_n$.*

Démonstration. Tout automorphisme envoie un carré sur un carré, et donc tout produit de carrés sur un produit de carrés. Donc tout automorphisme de $\mathfrak{S}_n$ envoie $\mathcal{A}_n$ dans $\mathcal{A}_n$. ■

Théorème 4. *Si $n \geq 5$ le groupe alterné $\mathcal{A}_n$ est le seul sous-groupe distingué de $\mathfrak{S}_n$ distinct de $\{1\}$ et de lui-même.*

Démonstration. Soit H un sous-groupe distingué non réduit à $\{1\}$ et distinct de $\mathfrak{S}_n$. On va démontrer qu'il contient un 3-cycle. Soit $\mu \in H$ distinct de l'élément neutre et soit τ une transposition qui ne commute pas avec μ . Une telle transposition existe à cause du :

Lemme 2. *Le centre de $\mathfrak{S}_n$ est réduit à $\{1\}$ si $n \geq 3$.*

Démonstration du lemme. Comme les transpositions engendrent le groupe $\mathfrak{S}_n$, il suffit de montrer qu'un élément σ qui commute à toutes les transpositions est l'identité. Raisonnons par l'absurde. Soit i tel que $i \neq \sigma(i)$, et soit j distinct de i et de $\sigma(i)$, rappelons que $n \geq 3$. Alors la transposition $(\sigma(i), j)$ ne commute pas à σ comme on le voit en évaluant sur i . □

Continuons la démonstration du théorème 4.

Considérons le produit $\tau \circ \mu \circ \tau \circ \mu^{-1}$. Il s'écrit $(\tau \circ \mu \circ \tau) \circ \mu^{-1}$ et est donc dans H puisque H est distingué. Par ailleurs, en l'écrivant $\tau \circ (\mu \circ \tau \circ \mu^{-1})$, on voit que c'est un produit de transpositions. Cet élément est différent de l'élément neutre car τ ne commute pas avec μ .

C'est donc soit un 3-cycle, dans ce cas on a le résultat, car alors H étant distingué, contient tous les 3-cycles, puisque qu'il en contient un et que tous les 3-cycles sont conjugués. Il est égal à $\mathcal{A}_n$ par le corollaire du théorème 2.

Ou bien c'est un produit $\tau_{i,j} \tau_{k,l}$ de transpositions à supports disjoints. Dans ce cas, comme H est distingué et comme $n \geq 5$, H contient aussi le produit $\tau_{i,j} \tau_{k,m}$ où i, j, k, l, m sont deux à deux distincts. Il contient donc le produit $\tau_{i,j} \tau_{k,l} \tau_{i,j} \tau_{k,m}$ qui est égal à $\tau_{k,l} \tau_{k,m}$ et est un 3-cycle. La démonstration s'achève alors comme précédemment. ■

On va maintenant démontrer le théorème 3 en s'inspirant de ce qui précède.

Lemme 3. *Le centre de $\mathcal{A}_n$ est réduit à $\{1\}$ si $n \geq 4$.*

Démonstration du lemme. Comme les 3-cycles engendrent le groupe $\mathcal{A}_n$ il suffit de montrer qu'un élément σ qui commute à tous les 3-cycles est l'identité. Raisonnons par l'absurde. Soit i tel que $i \neq \sigma(i)$, et soit j distinct de i , $\sigma(i)$ et $\sigma^2(i)$. Un tel élément existe, rappelons que $n \geq 4$. Alors le 3-cycle $c = (i, \sigma(i), j)$ ne commute pas à σ comme on le voit en évaluant $c \circ \sigma$ et $\sigma \circ c$ sur i . ■

Démonstration du théorème 3. Soit H un sous-groupe distingué de $\mathcal{A}_n$ et soit μ un élément de H . Soit c un 3-cycle qui ne commute pas à μ . Considérons alors le produit $c \circ \mu \circ c^{-1} \circ \mu^{-1}$. Il s'écrit $(c \circ \mu \circ c^{-1}) \circ \mu^{-1}$, et est donc dans H puisque H est distingué dans $\mathcal{A}_n$. Par ailleurs, en l'écrivant $c \circ (\mu \circ c^{-1} \circ \mu^{-1})$, on voit que c'est un produit de 3-cycles. Cet élément est différent de l'élément neutre car c ne commute pas avec μ .

Il y a quatre possibilités concernant les supports de c et $\mu \circ c^{-1} \circ \mu^{-1}$:

- les supports sont égaux, dans ce cas $c \circ (\mu \circ c^{-1} \circ \mu^{-1})$ est différent de l'élément neutre par hypothèse, et est un 3-cycle ;
- soit les supports sont disjoints ;
- les supports ont un point en commun, auquel cas on vérifie que le produit des deux trois cycles est un 5-cycle ;
- soit les supports ont deux points en commun, auquel cas, considérant la signature, on constate que le produit des deux 3-cycles est égal au produit de deux transpositions à supports disjoints ou à un 3-cycle.

Dans chacun de ces cas, on va démontrer que le sous-groupe H contient un 3-cycle, il les contiendra donc tous et sera donc égal à $\mathcal{A}_n$, d'après le théorème 2 de la section 4 et le lemme suivant :

Lemme 4. *Si $n \geq 5$ tous les 3-cycles sont conjugués non seulement dans $\mathfrak{S}_n$ mais aussi dans $\mathcal{A}_n$.*

Démonstration. Soient σ et τ les 3-cycles considérés. On sait (section 4) qu'ils sont conjugués par un élément de $\mathfrak{S}_n$. Reprenons les notations de la démonstration du théorème 2 de la section 4. Si l'élément ϕ construit dans la démonstration de ce théorème n'est pas dans $\mathcal{A}_n$, on peut le modifier sur le complémentaire du support de σ en le décomposant par une transposition t . Cela est possible car $n \geq 5$. Alors $\phi \circ t$ conjugue σ et τ et est dans $\mathcal{A}_n$. □

Dans le premier cas et le second du quatrième la démonstration est terminée.

Proposition 3. *Si $n \geq 8$ les produits de 3-cycles à supports disjoints, les 5-cycles, les produits de transpositions à supports disjoints de $\mathfrak{S}_n$ sont conjugués par un élément de $\mathcal{A}_n$.*

Démonstration. La démonstration précédente marche, à ceci près qu'il nous faut avoir deux éléments en dehors du support des permutations considérées. Ce qui implique que l'on doit avoir $n \geq 8$. En fait pour les 5-cycles, $n \geq 7$ suffit et pour les produits de transpositions à supports disjoints $n \geq 6$ suffit. $\square$

Considérons maintenant le second cas. Le sous-groupe H contient le produit de deux 3-cycles c et c' qui commutent, soit cc' . Supposons $n \geq 8$. À cause du lemme 4 H contient cc'^2 , donc le produit $cc'cc'^2 = c^2$ qui est un 3-cycle. Il contient donc tous les 3-cycles et est égal à $\mathcal{A}_n$ (lemme 4).

Dans ce cas il reste à traiter les valeurs $n = 6$ et $n = 7$. Le lemme 4 reste vrai :

Proposition 4. *Si $n = 6$ ou $n = 7$ les produits de 3-cycles à supports disjoints de $\mathfrak{S}_n$ sont conjugués par un élément de $\mathcal{A}_n$.*

Démonstration. On indique les grandes lignes de la démonstration pour $n = 6$, elle est identique pour $n = 7$. Les détails sont laissés en exercice. On constate par un argument de dénombrement, qu'il y a 40 éléments qui sont des produits de 3-cycles à supports disjoints. Par ailleurs, le groupe $\mathcal{A}_6$ a 360 éléments. Enfin on vérifie que l'ensemble des éléments de $\mathcal{A}_6$ commutant au produit des 3-cycles à supports disjoints c et c' est le sous-groupe engendré par c et c' qui a 9 éléments. La classe de conjugaison a donc 40 éléments, c'est donc l'ensemble des produits de 3-cycles à supports disjoints. $\square$

Considérons maintenant le troisième cas. Supposons que $n \geq 5$. Quitte à réordonner, on peut supposer que H contient le 5-cycle $c = (1, 2, 3, 4, 5)$. Ce cycle est conjugué au cycle $c' = (1, 3, 5, 4, 2)$ par la permutation

$$\begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 4 & 3, & 1, & 5, & 2 \end{pmatrix}$$

qui est paire. Le cycle c' est donc dans H . Enfin cc' est le 3-cycle $(2, 5, 3)$. On conclut comme plus haut.

Considérons enfin le dernier cas. On observe facilement que, si $n \geq 3$, toutes les transpositions sont conjugués par un élément de $\mathcal{A}_n$. En fait, il nous suffit de savoir que, dans $\mathfrak{S}_3$, toutes les transpositions sont conjuguées par les éléments de $\mathcal{A}_3$. Soit alors $n \geq 5$. De ce qui précède on déduit que si le sous-groupe H contient un élément de la forme $\tau_{a,b}\tau_{i,j}$ avec a, b, i, j étant deux à deux distincts et si c est distinct des ceux-ci, il contient aussi $\tau_{a,c}\tau_{i,j}$. Donc, il contient leur produit qui est le 3-cycle (a, b, c) .

Ceci achève la démonstration du théorème. $\square$

Exercices

Les exercices sont rangés par thèmes correspondant dans une certaine mesure aux sections du chapitre. Néanmoins, il peut être nécessaire pour certains exercices d'utiliser des résultats de base des sections ultérieures. Enfin, certains exercices proposés dépendent de l'algèbre linéaire, plus précisément de la réduction des endomorphismes et de la structure du groupe linéaire. Il vaut donc mieux les faire après avoir lu le chapitre 5.

I. Actions de groupes, équation des classes, théorèmes de Sylow

1. Soit E un ensemble sur lequel agit un groupe G . Soient A et B des parties de E . Calculer le fixateur et le stabilisateur de la réunion de A et B .
2. Déterminer dans le groupe $\mathrm{GL}_3(\mathbb{R})$ le centralisateur des matrices suivantes :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Soit G un groupe fini, et soit p le plus petit nombre premier divisant l'ordre de G . Soit H un sous-groupe tel que $\#G = p(\#H)$. Montrer que H est distingué. On pourra utiliser la proposition 2 de la section 1.
4. Les équations suivantes peuvent-elles être l'équation des classes d'un groupe d'ordre 12 ?

$$12 = 3 + 5 + 4,$$

$$12 = 4 + 4 + 4,$$

le premier terme désignant l'ordre du centre.

5. Trouver tous les groupes finis ayant 2, 3, 4 classes de conjugaisons.
6. Montrer que dans un groupe d'ordre 12, il y a toujours soit un groupe d'ordre 4, soit un groupe d'ordre 3 distingué. On utilisera les théorèmes de Sylow. En déduire qu'un groupe d'ordre 12 est toujours le produit semi-direct d'un groupe d'ordre 4 par un groupe d'ordre 3 ou inversement.
7. a) Montrer que si un groupe d'ordre 12 contient $\mathbb{Z}/4\mathbb{Z}$ comme 2-sous-groupe de Sylow, et que celui-ci est distingué, le groupe est abélien.
b) Montrer que, si un groupe H d'ordre 12 contient $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ comme 2-sous-groupe de Sylow, et que celui-ci est distingué, alors l'une des deux alternatives suivantes a lieu :
 - H est abélien,
 - H est isomorphe à $\mathcal{A}_4$.

Pour démontrer cela, on considérera l'action d'un 3-sous-groupe de Sylow sur le 2-sous-groupe de Sylow (qui est distingué).

c) Montrer que si un groupe H d'ordre 12 a un 3-sous-groupe de Sylow distingué, alors l'une des alternatives suivantes a lieu :

- H est abélien,
- si le 2-sous-groupe de Sylow est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ alors H est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathfrak{S}_3$,
- si le 2-sous-groupe de Sylow est isomorphe à $\mathbb{Z}/4\mathbb{Z}$, alors H a pour générateurs deux éléments a et b , le premier d'ordre 3, le second d'ordre 4 et on a $bab^{-1} = a^2$.

8. Soit T un tétraèdre régulier de l'espace vectoriel euclidien E de dimension 3, centré en l'origine.

a) Montrer que l'ensemble des isométries de E laissant T stable est un sous-groupe H .

b) En considérant l'action sur les sommets montrer que ce groupe est isomorphe à $\mathfrak{S}_4$.

c) Montrer que le sous-groupe de H des isométries directes est isomorphe à $\mathfrak{A}_4$.

9. Soit C un cube de l'espace vectoriel euclidien E de dimension 3, centré en l'origine.

a) Montrer que l'ensemble des isométries de E laissant C stable est un sous-groupe H .

b) En considérant les diagonales du cube joignant des sommets opposés construire un homomorphisme de H dans $\mathfrak{S}_4$.

c) Montrer que cet homomorphisme est surjectif. En particulier, décrire les transformations géométriques correspondant aux transpositions, cycles d'ordre 3, 4, aux produits de 2 transpositions à supports disjoints.

d) Étudier le noyau de l'homomorphisme et décrire H .

10. Soit p un nombre premier, et soient n et h des entiers avec $h < n$. Montrer en détail que tout groupe d'ordre p^n a au moins un sous-groupe d'ordre p^h .

11. Soit H un groupe d'ordre 18, montrer qu'il contient un 3-sous-groupe de Sylow distingué. Généraliser aux groupes d'ordre $2p^2$ avec p premier impair. En déduire que ce sont tous des produits semi-directs.

12. Soit G un groupe fini dont tous les sous-groupes de Sylow sont distingués. Montrer que G est isomorphe au produit de ses sous-groupes de Sylow.

13. a) (Formule de Burnside) Soit G un groupe agissant sur un ensemble fini E . Pour $g \in G$ on note E^g l'ensemble des $e \in E$ tels que $ge = e$. Pour $e \in E$ on note G_e le fixateur de e dans G , c'est-à-dire le sous-groupe des éléments g tels que $ge = e$.

Montrer que :

$$\sum_{e \in E} \#(G_e) = \sum_{g \in G} \#(E^g).$$

- b) En déduire que si ν désigne le nombre des orbites de l'action de G sur E :

$$\nu \#(G) = \sum_{g \in G} \#(E^g).$$

14. Soit G un groupe fini, S un p -sous-groupe de Sylow, N son normalisateur (le sous-groupe des $g \in G$ tels que $gSg^{-1} = S$). Montrer que le normalisateur de N est égal à N .

II. Groupes symétriques

1. Construire un homomorphisme surjectif de $\mathfrak{S}_4$ vers $\mathfrak{S}_3$, en considérant une *action transitive* de $\mathfrak{S}_4$ sur un ensemble à 3 éléments.
2. Montrer que le groupe alterné $\mathcal{A}_4$ n'a pas de sous-groupe d'ordre 6. Interpréter le groupe $\mathcal{A}_4$ comme un produit semi-direct.
3. Soit n un entier, calculer le nombre de n -cycles dans $\mathfrak{S}_n$. Soit p un nombre premier, calculer le nombre de p -cycles.
4. a) Soit n un entier, et p et q deux entiers tels que $p + q = n$. Calculer le nombre d'éléments de $\mathfrak{S}_n$ qui sont produit d'un p -cycle et d'un q -cycle à supports disjoints. Calculer l'ordre du centralisateur d'un tel élément.
b) Plus généralement on calculera le nombre d'éléments dans la classe de conjugaison d'un élément à l'aide de la partition qui lui est associée.
5. Écrire l'équation des classes pour les groupes $\mathfrak{S}_5$, $\mathfrak{S}_6$, $\mathcal{A}_5$, $\mathcal{A}_6$. On donnera le nombre d'éléments dans chaque classe de conjugaison.
6. Montrer que, dans le groupe symétrique, tout élément est conjugué à son inverse.
7. Montrer que le seul sous-groupe d'indice 2 de $\mathfrak{S}_n$ est $\mathcal{A}_n$.
8. a) Montrer que $\mathfrak{S}_{n-1}$ est un sous-groupe maximal de $\mathfrak{S}_n$, c'est-à-dire que si on a un sous-groupe H tel que $\mathfrak{S}_{n-1} \subset H \subset \mathfrak{S}_n$, alors soit $H = \mathfrak{S}_{n-1}$ soit $H = \mathfrak{S}_n$.
b) Soient deux entiers p et q tels que $p + q = n$. On considère le sous-groupe stabilisateur du sous-ensemble $\{1, \dots, p\} \subset \{1, \dots, n\}$. Montrer que ce sous-groupe est isomorphe au groupe $\mathfrak{S}_p \times \mathfrak{S}_q$. Montrer que ce sous-groupe est un sous-groupe maximal (au sens précédent) de $\mathfrak{S}_n$.

9. Caractériser les partitions correspondant à des éléments du groupe alterné.
10. Soit $s \in \mathfrak{S}_n$ et soit λ la partition associée à sa classe de conjugaison. Montrer que l'ordre de s ne dépend que de λ , puis le calculer en fonction de λ .
11. Soit $\ell_i(s)$ le nombre de cycles d'ordre i dans la décomposition en cycles d'un élément s du groupe alterné. Montrer que l'alternative suivante a lieu. Soit les classes de conjugaison de s dans le groupe alterné et dans le groupe symétrique sont égales, soit il existe un entier i tel que $\ell_{2i}(s) > 0$ ou que $\ell_{2i+1}(s) > 1$.
12. a) (Cauchy) Soit $n \geq 5$ et p le plus grand nombre premier inférieur ou égal à n . Montrer que le groupe symétrique $\mathfrak{S}_n$ ne contient pas de sous-groupe d'indice compris entre 3 et $p - 1$. On commencera par montrer que le plus petit sous-groupe de $\mathfrak{S}_n$ contenant tous les p -cycles est le groupe alterné.
- b) Soit un sous-groupe de $\mathfrak{S}_n$, et soit H un sous-groupe d'indice ℓ , $3 \leq \ell \leq p - 1$. Montrer que l'action par translation d'un élément s d'ordre p sur l'espace quotient $\mathfrak{S}_n/H$ est triviale.
- c) Conclure.
13. Montrer que le groupe d'automorphismes de $\mathfrak{S}_3$, respectivement de $\mathfrak{S}_4$, est isomorphe au groupe des automorphismes intérieurs de $\mathfrak{S}_3$, respectivement de $\mathfrak{S}_4$, soit à $\mathfrak{S}_3$, respectivement à $\mathfrak{S}_4$, lui-même.
14. ** Montrer que si $n \neq 6$ le groupe d'automorphismes de $\mathfrak{S}_n$, est isomorphe au groupe des automorphismes intérieurs de $\mathfrak{S}_n$, soit à $\mathfrak{S}_n$ lui-même. On commencera par montrer que l'image d'un élément d'ordre 2 par un automorphisme est encore d'ordre 2. On décrira tous les éléments d'ordre 2 dans $\mathfrak{S}_n$, et on calculera l'ordre de leurs centralisateurs. On montrera que l'image d'une transposition est une transposition.

III. Classification de petits groupes finis

1. Montrer qu'il y a deux groupes, et seulement deux, d'ordre 10 non-isomorphes, l'un abélien, l'autre non.
2. Montrer qu'il y a deux groupes, et seulement 2, d'ordre 21 non-isomorphes, l'un abélien, l'autre non. Même question pour les groupes d'ordre 22 et 26.
3. Soit p un nombre premier impair de la forme $3k + 2$. Montrer que tout groupe d'ordre $3p^2$ est abélien.
4. Montrer qu'un groupe d'ordre 255 est abélien. On montrera qu'il y a un sous-groupe distingué d'ordre 85.

5. Montrer qu'un groupe d'ordre 30 a nécessairement un 5-sous-groupe de Sylow ou un 3-sous-groupe de Sylow distingué. En déduire que le groupe a un sous-groupe d'ordre 15. Montrer que le groupe est soit abélien, soit produit semi-direct non-abélien de $\mathbb{Z}/15\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z}$.
6. a) ** Décrivez les automorphismes de $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$.
b) ** Montrer qu'il existe un groupe d'ordre 27, non-abélien, dont tous les éléments sont d'ordre 3.
7. a) Montrer qu'il n'y a pas de groupe simple d'ordre 56.
b) ** Y a-t-il un groupe simple 224 ?

■ Quelques réponses ou indications ■

I. Actions de groupes, équation des classes, théorèmes de Sylow

1. C'est l'intersection des fixateurs (ou des stabilisateurs) de A et B .
3. On construit un homomorphisme de G vers le groupe symétrique $\mathfrak{S}_p$, en considérant l'action par translation de G sur G/H qui a p éléments. Puis on utilise la condition de primalité : le noyau de cet homomorphisme est l'intersection des conjugués de H , mais à cause des conditions arithmétiques l'ordre de l'image divise p et est en fait égal à p , donc l'ordre du noyau est égal à celui de H .
4. Le premier cas est impossible, car 5 ne divise pas 12. Le second l'est aussi car le groupe quotient de G par son centre serait $\mathbb{Z}/3\mathbb{Z}$ et G serait abélien.
5. Dans le premier cas il n'y a que $\mathbb{Z}/2\mathbb{Z}$, car on remarque que en dehors de l'élément neutre tous les éléments doivent avoir le même ordre, donc que le groupe est un p -groupe pour un nombre premier p . Ensuite on observera que si on a un homomorphisme surjectif d'un groupe vers un autre on en déduit une surjection au niveau des classes de conjugaison. On utilise alors le fait qu'un p -groupe a un centre non-trivial. On procède de manière analogue pour les cas suivants en observant que le cardinal est divisible par au plus 2 ou 3 nombres premiers suivant le cas.
6. On compte le nombre d'éléments appartenant à un 3-sous-groupe de Sylow. Si il n'y a pas de 3-sous-groupe de Sylow distingué, comme il y a au moins 4 3-sous-groupes de Sylow, il y a au moins 8 éléments d'ordre 3. Avec l'élément neutre cela fait 9 éléments. Les 3 éléments forment nécessairement avec l'élément neutre un 2-sous-groupe de Sylow.
7. a) Se référer au chapitre 1 pour la structure des groupes abéliens et des produits semi-directs. On observe que sous les hypothèses le groupe est nécessairement un produit semi-direct et qu'il n'y a pas d'homomorphismes non trivial de $\mathbb{Z}/3\mathbb{Z}$ dans les automorphismes de $\mathbb{Z}/4\mathbb{Z}$.
b) Comme dans le cas précédent, et pour les mêmes raisons, on a un produit semi-direct. Mais dans ce cas il y a un homomorphisme non-trivial de $\mathbb{Z}/3\mathbb{Z}$ dans le groupe des automorphismes de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ dont on rappelle qu'il s'identifie à $\mathfrak{S}_3$.

- c) On raisonne de manière analogue, on a des produits semi-directs, mais cette fois ci on doit considérer les homomorphismes d'un 2-sous-groupe de Sylow dans le groupe des automorphismes de $\mathbb{Z}/3\mathbb{Z}$ soit $\mathbb{Z}/2\mathbb{Z}$.
8. Montrer que toute permutation des sommets est induite par une isométrie. On peut se contenter de le montrer pour les transpositions, puisqu'elles engendrent le groupe symétrique. Or les symétries autour d'un plan contenant une arête et le milieu de l'arête opposée réalisent les transpositions. On peut aussi réaliser les cycles d'ordre 3 à l'aide des rotations d'angle $2\pi/3$ autour des hauteurs du tétraèdre.
 9. Les diagonales du cube forment un ensemble à 4 éléments laissé fixe par les isométries fixant le cube. Ceci détermine donc un homomorphisme du groupe du cube vers le groupe des permutations de ces diagonales. Les rotations d'angle $\frac{2\pi}{3}$ et d'axe une diagonale réalisent les cycles d'ordre 3. Les symétries autour des plans contenant deux arêtes opposées du cube réalisent les transpositions. Le noyau de cet homomorphisme est constitué de l'identité et de l'opposée de l'identité.
 10. Faire une récurrence.
 11. Compter les éléments dont l'ordre est 9.
 12. Commencer par montrer que, si un groupe G a deux sous-groupes distingués H et K d'intersection réduite à l'élément neutre et tels que $G = HK$, alors G est isomorphe au produit $H \times K$.

II. Groupes symétriques

1. Considérer l'action du groupe sur ses classes de conjugaison, les produits de transpositions à support disjoints forment une classe à 3 éléments.
2. On commencera par montrer qu'il y a 3 éléments d'ordre 2 dans $\mathcal{A}_4$, les produits de transpositions à supports disjoints. Puis un groupe d'ordre 6 contient au moins un élément d'ordre 3, et un élément d'ordre 2, donc contient un produit de transpositions à supports disjoint. En conjuguant par un élément d'ordre 3 en déduire qu'il contiendrait tous les éléments d'ordre 2, conclure.
3. La première réponse est $(n-1)!$: partant de 1 on a $n-1$ choix pour son image par le n -cycle $(2, \dots, n)$, puis on a $n-2$ choix pour celui de la seconde image... La seconde réponse est $C_n^p (p-1)!$ si $p \leq n$.
4. a) Distinguer suivant le cas où $p = q$ ou non. Dénombrer les partitions d'un ensemble à n éléments en un sous-ensemble à p éléments et un sous-ensemble à q éléments. On trouve $\frac{n!}{p^q}$ dans le premier cas, et $\frac{n!}{2p^2}$ dans le second.
 b) Supposons que la partition de n associée comporte μ_i fois l'entier i . Autrement dit, on a $n = \mu_1 + 2\mu_2 + \dots + n\mu_n$. On comptera le nombre de partitions de $\{1, \dots, n\}$ en μ_1 sous-ensembles à 1 élément, μ_2 sous-ensembles à 2 éléments. La réponse est alors

$$\frac{n!}{\mu_1! \cdots \mu_n! 1^{\mu_1} 2^{\mu_2} \cdots n^{\mu_n}}.$$

5. Pour le groupe $\mathcal{A}_5$:

partition	ordre	cardinal de la classe
(1, 1, 1, 1, 1)	1	1
(2, 2, 1)	2	15
(3, 1, 1)	3	20
(5)	5	24

Pour le groupe $\mathfrak{S}_6$:

partition	ordre	cardinal de la classe
(1, 1, 1, 1, 1, 1)	1	1
(2, 1, 1, 1, 1)	2	15
(2, 2, 1, 1)	2	45
(2, 2, 2)	2	15
(3, 1, 1, 1)	3	40
(3, 2, 1)	6	120
(3, 3)	3	5
(4, 1, 1)	4	30
(4, 2)	4	30
(5, 1)	5	144
(6)	6	120

6. Il suffit de démontrer la propriété pour les cycles.

7. Voir le cours.

8. a) Montrer que si on a un sous-groupe H qui contient $\mathfrak{S}_{n-1}$ et en est distinct, il contient la transposition $\tau_{n-1,n}$. On a alors dans le sous-groupe H un système générateur pour $\mathfrak{S}_n$. Pour montrer la première propriété on choisit un élément $s \in H$ mais tel que $s \notin \mathfrak{S}_{n-1}$. Quitte à remplacer s par st où $t \in \mathfrak{S}_{n-1}$ montrer que l'on peut supposer que $s(n) = n-1$, puis que $s(n-1) = n$. Modifier alors par un élément de $\mathfrak{S}_{n-2}$ pour conclure.

b) Montrer que si on a un sous-groupe H qui contient $\mathfrak{S}_p \times \mathfrak{S}_q$ et en est distinct, il contient la transposition $\tau_{p,p+1}$.

11. Montrer que la classe de conjugaison, d'un élément s du groupe alterné dans le groupe symétrique est réunion de deux classes de conjugaison dans le groupe alterné si et seulement

si la décomposition en cycles de s ne fait apparaître que des cycles non conjugués (c'est-à-dire que les cardinaux des supports sont deux à deux distincts) de longueur impaire.

12. a) b) c) Utiliser le cours, en particulier les arguments dans la démonstration de la simplicité de $\mathcal{A}_n$ si $n \geq 5$. Puis utiliser le fait que le cardinal de l'ensemble quotient $\mathfrak{S}_n/H$ est premier à p . Enfin si l'action est triviale cela veut dire que si s désigne l'élément d'ordre p considéré on $sH = H$ et donc $s \in H$.
13. Étudier l'action d'un automorphisme sur l'ensemble des éléments d'ordre 2.
En particulier, montrer qu'une transposition est transformée en un élément d'ordre 2. Conclure dans le cas $n = 3$.
Dans le cas $n = 4$, on observera qu'outre les transpositions, les seuls autres éléments d'ordre 2 sont les produits de transpositions à supports disjoints. En comptant le nombre de transpositions, et celui de produits de transpositions à supports disjoints, on montrera que l'image d'une transposition est encore une transposition. On utilisera le fait que les images de toutes les transpositions sont dans la même classe de conjugaison.
14. Généraliser l'argument de l'exercice 13, en observant que si f est un automorphisme de $\mathfrak{S}_n$ et si τ est une transposition alors $f(\tau)$ est un élément d'ordre 2. Montrer que $f(\tau)$ est une transposition en observant que les cardinaux des centralisateurs de τ et $f(\tau)$ sont égaux. Puis en calculant les cardinaux des centralisateurs des éléments d'ordre 2 dans $\mathfrak{S}_n$.

III. Classifications de petits groupes finis

- 1 et 2. Rechercher les sous-groupes de Sylow. Montrer que l'on a soit un groupe abélien, soit un produit semi-direct non-abélien.
3. Montrer qu'il y a un seul p -Sylow, qui est donc distingué. Puis montrer que l'on a nécessairement un produit semi-direct et que l'homomorphisme le définissant est trivial. On pourra se référer au chapitre 5.
4. Montrer qu'il n'y a qu'un 17-sous-groupe de Sylow, et qu'il est donc distingué.
5. Pour démontrer l'existence d'un sous-groupe d'ordre 15, on suppose, par exemple, que le 5-Sylow, soit H , est distingué. Soit K un 3-Sylow. On montre que KH est un sous-groupe d'ordre 15. On se servira de la structure des groupes d'ordre 15.
6. a) b) Il s'agit des matrices $(2, 2)$ inversibles à coefficients dans $\mathbb{Z}/3\mathbb{Z}$. Il y a 48 éléments. Montrer qu'un groupe quelconque d'ordre 27 contient $\mathbb{Z}/9\mathbb{Z}$ ou $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ comme sous-groupe distingué. Dans le second cas, et si on suppose que tous les éléments sont d'ordre 3 dans le groupe, montrer que le groupe est produit semi-direct de $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ et $\mathbb{Z}/3\mathbb{Z}$. Puis à l'aide d'un homomorphisme de $\mathbb{Z}/3\mathbb{Z}$ dans les automorphismes de $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ construire l'exemple cherché.
7. a) b) En appliquant les théorèmes de Sylow montrer qu'un tel groupe simple d'ordre 56 aurait 7 2-sous-groupes de Sylow et 8 7-sous-groupes de Sylow. Montrer qu'il y aurait 48 éléments d'ordre 7. Compter les éléments d'ordre une puissance de 2.
Soit G un groupe simple à 224 éléments si il existe. Vérifier qu'il a 7 2-Sylow. En déduire un homomorphisme non-trivial de G dans $\mathfrak{S}_7$. Montrer, en considérant les ordres, que cet homomorphisme n'est pas injectif.

Chapitre 3

Anneaux, idéaux, polynômes et séries formelles

Dans ce chapitre on rappelle dans la première section les définitions de base concernant les anneaux : sous-anneaux, homomorphismes, idéaux, anneaux quotients, corps. La seconde est consacrée aux anneaux principaux, la troisième section à une étude détaillée de l'anneau des classes de congruence modulo n , et la quatrième à une brève introduction à l'étude de la cryptographie.

Les deux sections suivantes sont consacrées aux anneaux de polynômes, de séries formelles, et aux anneaux factoriels. On en profite pour faire l'étude des polynômes symétriques. L'avant dernière section est consacrée aux zéros des polynômes et à la théorie du résultant. Enfin, diverses propriétés des séries formelles sont présentées dans la neuvième et dernière section.

1. Anneaux, idéaux, anneaux quotients

Définition 1

Un *anneau* est un ensemble A muni de deux lois internes. La première notée additivement est appelée l'addition, la seconde notée multiplicativement est appelée la multiplication. L'addition, détermine sur A une loi de groupe abélien sur A . La multiplication détermine une loi associative sur A . La multiplication est distributive par rapport à l'addition. C'est-à-dire que pour tous $x, y, z \in A$ on a :

$$x(y + z) = xy + xz \quad \text{et} \quad (y + z)x = yx + zx.$$

Enfin il y a un élément unité, noté 1 , tel que pour tout $a \in A$ on ait $1 \times a = a \times 1 = a$.

Les anneaux que nous considérons sont donc avec élément unité pour la multiplication, cet élément sera toujours noté 1. L'élément 0 est supposé distinct de 1.

Notons que pour tout élément x de l'anneau on a $0 \times x = x \times 0 = 0$. En effet, en utilisant la distributivité, on obtient

$$x = 1 \times x = (1 + 0) \times x = x + (0 \times x),$$

donc $x = x + (0 \times x)$, donc $0 \times x = 0$. Le même argument donne l'autre identité.

Définition 2

L'anneau est dit commutatif si la multiplication est commutative.

Dans un anneau commutatif quelconque on a la formule du binôme de Newton :

Proposition 1. Soit A un anneau commutatif. Pour tous a et b dans l'anneau A on a :

$$(a + b)^n = \sum_{i=0, \dots, n} C_n^i a^i b^{n-i}.$$

Démonstration. La démonstration peut se faire par récurrence sur l'entier n ou par dénombrement. Faisons la par dénombrement. On considère le produit

$$\underbrace{(a + b) \times \dots \times (a + b)}_{n \text{ fois}}.$$

Dans le développement du produit le terme $a^i b^{n-i}$ apparaît chaque fois que l'on a sélectionné i parenthèses parmi les n successives, que l'on développe suivant a les parenthèses sélectionnées, et suivant b les $n - i$ autres. Le coefficient de $a^i b^{n-i}$ est donc le nombre de sous-ensemble de $\{1, \dots, n\}$ à i éléments, c'est-à-dire C_n^i . ■

Remarque

On a plus généralement pour $a_1, \dots, a_p \in A$ la formule suivante :

$$(a_1 + \dots + a_p)^n = \sum_{(u_1, \dots, u_p) | u_1 + \dots + u_p = n} \frac{n!}{u_1! \dots u_p!} a_1^{u_1} \dots a_p^{u_p}.$$

Définition 3

On dira qu'un élément a d'un anneau A divise un élément non nul b , et on notera $a \mid b$, s'il existe un élément c de A tel que $b = ac$.

On comparera à la définition de diviseurs de zéro (définition 6).

Définition 4

Soit A un anneau, si pour tout élément $a \in A$ différent de zéro il existe un élément a' , appelé l'inverse, tel que $aa' = a'a = 1$ on dit que A est un *corps*. Autrement dit A un corps si la multiplication détermine une loi interne sur $A - \{0\}$ qui en fait un groupe.

La dernière phrase de la définition est en fait une proposition : si A est un corps la multiplication détermine une loi interne sur $A - \{0\}$. En effet, supposons donné deux éléments x et y différents de zéro dans le corps et montrons que leur produit xy ne peut être nul. Soient en effet x' , respectivement y' , un inverse pour x , respectivement pour y . Si on a $xy = 0$ on a $x'xyy' = x' \times 0 \times y' = 0$ et $x'xyy' = 1 \times 1 = 1$ ce qui est impossible.

Remarque

On remarquera que l'inverse est unique. En effet si on a $xy = 1$ et $xy' = 1$ alors on a $xyy' = y$ et $xyy' = y'$.

Dans un anneau quelconque un élément peut avoir ou ne pas avoir d'inverse, l'inverse s'il existe est unique.

Les deux premiers exemples d'anneaux sont l'anneau $\mathbb{Z}$ des entiers relatifs et le corps $\mathbb{Q}$ des nombres rationnels.

Un sous-ensemble B d'un anneau A est un sous-anneau si :

- B est un sous-groupe du groupe $(A, +)$,
- $1 \in B$,
- si x, y sont éléments de B , alors xy est élément de B .

Par exemple $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$. On laisse au lecteur le soin de vérifier que l'intersection d'une famille de sous-anneaux est un sous-anneau.

Définition 5

On appelle *homomorphisme* d'anneaux

$$\varphi : A \longrightarrow A'$$

toute application de A dans A' qui est un homomorphisme pour l'addition et la multiplication.

Ceci signifie que :

- $\varphi(x + y) = \varphi(x) + \varphi(y)$ pour tous x et y dans A ,
- $\varphi(1_A) = 1_{A'}$,
- $\varphi(xy) = \varphi(x)\varphi(y)$ pour tous x et y dans A .

On définit l'image d'un homomorphisme comme dans le cas des groupes. C'est un sous-anneau, on le note $\text{Im}(\varphi)$.

De même que l'on définit le produit d'une famille de groupes, on définit le produit d'une famille d'anneaux A_i , l'indice i décrivant un ensemble d'indices I fini ou infini. L'anneau $\prod_{i \in I} A_i$ est constitué des familles d'éléments (x_i) , $i \in I$. La somme et le produit de deux familles sont évidemment faites termes à termes :

$$(x_i) + (x'_i) = (x_i + x'_i)$$

et

$$(x_i) \times (x'_i) = (x_i \times x'_i).$$

La famille constituée uniquement de 1 est élément neutre.

On prendra garde à ce que l'on n'a pas une notion de somme directe comme pour les groupes. En effet on pourrait considérer, comme pour les groupes, les familles d'éléments tous nuls, sauf pour un nombre fini d'entre eux. On peut définir somme et produit par les mêmes règles que plus haut. Mais si l'ensemble d'indices est infini il n'y a pas d'élément neutre : la famille constituée uniquement par des 1 n'est pas dans l'ensemble des suites constituées d'éléments presque tous nuls.

Soit A un anneau, voici d'autres définitions :

Définition 6

Un élément non nul x dans A est un diviseur de zéro s'il existe un élément non nul y tel que $xy = 0$.

Définition 7

Un anneau est un domaine d'intégrité (ou est un *anneau intègre*) s'il n'a pas de diviseurs de zéro.

Remarque

On vérifiera en exercice que le produit d'au moins deux anneaux non triviaux n'est jamais intègre.

Tout anneau intègre peut être considéré comme sous-anneau d'un corps, son corps des quotients :

Proposition 2. Soit A un anneau commutatif intègre. Il existe un corps K tel que :

- A s'identifie à un sous-anneau de K ,
- pour tout élément $x \in K$ il existe un élément $u \in A$ tel que $ux \in A$.

Démonstration. La construction est la même que pour $\mathbb{Q}$.

On considère l'ensemble $A \times A^*$ et on le munit de la relation suivante : $(a, b) \cong (u, v)$ si et seulement si $av = bu$. C'est une relation d'équivalence. On laisse au lecteur le soin de vérifier la réflexivité et la symétrie.

En utilisant la commutativité et l'intégrité on vérifie la transitivité. En effet si $(a, b) \cong (u, v)$ et $(u, v) \cong (m, n)$ on a $av = bu$ et $un = vm$. Donc $avn = bun = bvm$, soit $v(an - bm) = 0$ et donc $an = bm$.

On définit deux lois internes, c'est-à-dire des applications

$$(A \times A^*) \times (A \times A^*) \longrightarrow A \times A^*$$

sur l'ensemble $A \times A^*$, appelées addition et multiplication, par :

- $(a, b) + (u, v) = (av + bu, bv)$,
- $(a, b) \times (u, v) = (au, bv)$.

Notons que ce sont des lois interne. En effet, A étant intègre, bv ne peut être nul sans que b ou v ne soit nul.

On vérifie aisément que la classe d'équivalence de la somme et du produit de deux éléments ne dépend que des classes d'équivalence des deux éléments. Ceci permet de définir l'addition et la multiplication sur l'ensemble quotient. La classe de $(0, 1)$ est élément neutre pour l'addition, celle de $(1, 1)$ l'est pour la multiplication. Les deux lois sont évidemment commutatives.

La multiplication est distributive par rapport à l'addition. C'est-à-dire que :

$$(u, v) \times ((a, b) + (a', b')) = (u, v) \times (a, b) + (u, v) \times (a', b').$$

Par ailleurs, l'ensemble des classes d'éléments de la forme (ka, a) est isomorphe à A . En fait, l'application de A dans K qui à k associe la classe de $(k, 1)$ est un isomorphisme de A sur son image. La classe de l'élément $(1, k)$, $k \neq 0$, est un inverse pour la multiplication de la classe de $(k, 1)$.

Le corps $\mathbb{K}$ est l'ensemble de ces classes d'équivalence muni de ces deux lois internes. C'est le corps des quotients de A . Par exemple, le corps des quotients des polynômes à coefficients dans un corps k est le corps $k(X)$ des fractions rationnelles à coefficients dans k . ■

Idéaux et anneaux quotients

Passons maintenant à l'étude des structures quotients. Soit un homomorphisme d'anneaux $\varphi : A \rightarrow A'$. Le noyau de l'homomorphisme φ est l'ensemble $\mathcal{I}$ des éléments $x \in A$ tels que $\varphi(x) = 0$. Il est noté $\text{Ker}(\varphi)$. C'est évidemment un sous-groupe de A pour l'addition et on a :

Proposition 3. *L'ensemble $\mathcal{I}$ est tel que pour tout $y \in A$ et tout $x \in \mathcal{I}$ on a $yx \in \mathcal{I}$ et $xy \in \mathcal{I}$.*

Démonstration. En effet si on a $\varphi(x) = 0$ on a $\varphi(yx) = \varphi(y)\varphi(x) = 0$ et $\varphi(xy) = \varphi(x)\varphi(y) = 0$. ■

Définition 8

Un sous-ensemble $\mathcal{I}$ d'un anneau A qui est un sous-groupe pour l'addition et tel que si $a \in A$ et $x \in \mathcal{I}$ alors $xa \in \mathcal{I}$ et $ax \in \mathcal{I}$ est appelé un *idéal bilatère* de A .

Donc le noyau d'un homomorphisme d'anneaux est un idéal bilatère de A . On peut définir la notion d'idéal à gauche et d'idéal à droite :

Définition 9

Un sous-ensemble $\mathcal{I}$ d'un anneau A qui est un sous-groupe pour la loi additive et qui est tel que si $a \in A$ et $x \in \mathcal{I}$ alors $ax \in \mathcal{I}$, respectivement $xa \in \mathcal{I}$ est appelé un *idéal à gauche*, respectivement à droite de A .

Dans la suite on ne considérera des idéaux à gauche et à droite que dans les exercices. Si bien que la mention « idéal » vaudra pour « idéal bilatère ». Soit donc A un anneau et $\mathcal{I}$ un idéal.

On vérifie sans peine que l'intersection d'une famille d'idéaux d'un anneau A est encore un idéal. Ce résultat vaut que l'on se place dans le cadre des idéaux à gauche, à droite ou bilatère. Soit par exemple une famille d'idéaux bilatères $\mathcal{I}_\alpha$, leur intersection est déjà un sous-groupe. Maintenant si x est dans l'intersection et si $a \in A$ ax et xa sont dans chacun des idéaux de la famille donc dans l'intersection.

Définition 10

On appelle idéal engendré par une partie de A l'intersection de tous les idéaux contenant cette partie. Autrement dit c'est le plus petit idéal contenant la partie considérée.

En voici une description plus concrète. On se place dans le cas où A est un anneau commutatif A . Soit $x_1, \dots, x_n$ une famille d'éléments de A .

L'ensemble des éléments de la forme $\lambda_1 x_1 + \dots + \lambda_n x_n$, où les λ_i décrivent A est un idéal de A . En effet la somme de deux éléments de ce type est encore du même type, et, si on multiplie par un élément de l'anneau on obtient encore un élément du même type.

On note cet idéal $(x_1, \dots, x_n)$, c'est l'idéal engendré par la famille considérée.

On étend le résultat à l'idéal engendré par une famille infinie de x_i prenant les sommes ou seuls un nombre fini de λ_i sont différents de zéro.

Remarque

La remarque suivante est souvent très utile. Soit $x \in A$ un élément et (x) l'idéal engendré. Alors (x) est égal à A si et seulement si x est inversible. En effet (x) est égal à A si et seulement si $1 \in (x)$; cette dernière condition a lieu si et seulement si il existe $y \in A$ tel que $xy = 1$, le résultat suit.

L'idéal $\mathcal{I}$ est un sous-groupe du groupe abélien A . On peut donc considérer le groupe quotient $A/\mathcal{I}$, on va le munir d'une multiplication et en faire un anneau.

Les éléments de $A/\mathcal{I}$ sont des classes de la forme $a + \mathcal{I} = \{a + x \mid x \in \mathcal{I}\}$. Considérons deux classes dans $A/\mathcal{I}$, soient $\mathcal{J}_1 = a + \mathcal{I}$ et $\mathcal{J}_2 = b + \mathcal{I}$, considérons l'ensemble $\mathcal{J}_1 \mathcal{J}_2$, c'est l'ensemble des éléments

$$\{ab + ay + xb + xy \mid x, y \in \mathcal{I}\}.$$

Comme $\mathcal{I}$ est un idéal cet ensemble est contenu dans la classe $ab + \mathcal{I}$. Définissons donc un produit, noté provisoirement $*$, sur $A/\mathcal{I}$ par la formule :

$$\mathcal{J}_1 * \mathcal{J}_2 = ab + \mathcal{I}.$$

Cette loi est bien définie par construction.

On vérifie que cette loi est associative et distributive par rapport à la loi additive. Les démonstrations sont formellement identiques à celles faites dans le cas des groupes et laissées au lecteur. En particulier, la classe $1 + \mathcal{I}$ est élément neutre pour la multiplication. Enfin l'application :

$$\begin{aligned} A &\longrightarrow A/\mathcal{I}, \\ a &\longmapsto a + \mathcal{I}, \end{aligned}$$

est par construction un homomorphisme d'anneaux.

On va étudier le comportement des idéaux vis à vis des homomorphismes.

Proposition 4

- Soit φ un homomorphisme d'un anneau A dans un anneau A' , et soit $\mathcal{I}$ un idéal de A . Alors l'ensemble $\varphi^{-1}(\mathcal{I})$ est un idéal de A .
- Soit φ un homomorphisme surjectif d'un anneau A dans un anneau A' , et soit $\mathcal{I}$ un idéal de A . Alors l'ensemble $\varphi(\mathcal{I})$ est un idéal de A' .
- Soit φ un homomorphisme d'un anneau A dans un anneau A' , et soit $\mathcal{I}$ un idéal de A . Supposons que $\mathcal{I}$ soit contenu dans $\text{Ker}(\varphi)$. Alors il existe un unique homomorphisme ψ de $A/\mathcal{I}$ dans A' tel que $\varphi = \psi \circ p$, où p est l'homomorphisme canonique de A dans $A/\mathcal{I}$. Autrement dit, on a un diagramme commutatif :

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & A' \\ & \searrow p & \nearrow \psi \\ & A/\mathcal{I} & \end{array}$$

Démonstration. On démontre le troisième point, les autres sont laissés en exercices. On définit ψ par la formule

$$\psi(x + \mathcal{I}) = \varphi(x).$$

Cette formule a un sens. En effet si $x + \mathcal{I} = x' + \mathcal{I}$ on a, par hypothèse sur φ et $\mathcal{I}$, la relation $\varphi(x) = \varphi(x')$. Il reste à vérifier que ψ est un homomorphisme. On a

$$\psi(x + \mathcal{I} + x' + \mathcal{I}) = \psi(x + x' + \mathcal{I}) = \varphi(x + x') = \varphi(x) + \varphi(x') = \psi(x + \mathcal{I}) + \psi(x' + \mathcal{I}),$$

et

$$\psi((x + \mathcal{I})(x' + \mathcal{I})) = \psi(xx' + \mathcal{I}) = \varphi(xx') = \varphi(x)\varphi(x') = \psi(x + \mathcal{I})\psi(x' + \mathcal{I}),$$

la première identité vient de ce que $\mathcal{I}$ est un idéal. Quant à l'unicité de ψ elle résulte de la surjectivité de p . ■

Soient un anneau A , un idéal $\mathcal{I}$ de A et un élément $x \in A$. L'ensemble $x + \mathcal{I}$ est la classe de x modulo $\mathcal{I}$ et sera notée le plus souvent $\bar{x}$.

Nous allons donner maintenant, comme dans le cas des groupes, un certain nombre de conséquences classiques de cette construction : la décomposition canonique des homomorphismes et les théorèmes d'isomorphismes.

Corollaire. Soit $\varphi : A \rightarrow A'$ un homomorphisme et soient p l'application canonique de A dans $A/\text{Ker}(\varphi)$, i l'inclusion canonique de $\text{Im}(\varphi)$ dans A' . Il existe un unique isomorphisme ϕ' de $A/\text{Ker}(\varphi)$ dans $\text{Im}(\varphi)$ tel que $\phi = i \circ \phi' \circ p$. Autrement dit, on a

une factorisation et un isomorphisme :

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A' \\ \downarrow p & & \uparrow i \\ A/\text{Ker}(\phi) & \xrightarrow[\cong]{\phi'} & \text{Im}(\phi) \end{array}$$

Démonstration. On applique le résultat précédent avec $\mathcal{I} = \text{Ker}(\phi)$, et avec ϕ' l'homomorphisme de $A/\mathcal{I}$ dans $\text{Im}(\phi)$ défini par

$$\phi'(x + \mathcal{I}) = \phi(x).$$

C'est un isomorphisme car il est injectif et surjectif par construction. Pour finir on compose à gauche par l'inclusion i de l'image de ϕ dans A' . ■

• **Théorème 1.** Soient A un anneau, et soient $\mathcal{I}$ et $\mathcal{J}$ deux idéaux de A tels que $\mathcal{J} \subset \mathcal{I}$. L'ensemble $\mathcal{I}/\mathcal{J}$ est un idéal $\mathcal{I}'$ de l'anneau quotient $A/\mathcal{J}$. L'anneau quotient $A/\mathcal{I}$ est isomorphe à l'anneau quotient $(A/\mathcal{J})/\mathcal{I}'$. Autrement dit, on a une factorisation et un isomorphisme :

$$\begin{array}{ccc} A & \longrightarrow & A/\mathcal{I} \\ \downarrow & & \uparrow \cong \\ A/\mathcal{J} & \longrightarrow & (A/\mathcal{J})/\mathcal{I}' \end{array}$$

Par abus de notation $\mathcal{I}'$ sera noté $\mathcal{I}/\mathcal{J}$.

Démonstration. On considère l'homomorphisme canonique $p : A \rightarrow A/\mathcal{I}$. Son noyau contient l'idéal $\mathcal{J}$. Il existe donc un homomorphisme surjectif de $A/\mathcal{J}$ vers $A/\mathcal{I}$, il associe à la classe $x + \mathcal{J}$ la classe $x + \mathcal{I}$. Son noyau $\mathcal{I}'$ consiste en l'ensemble des classes

$$x + \mathcal{J}$$

telles que $x + \mathcal{I}$ est égal $\mathcal{I}$, soit l'ensemble des classes $x + \mathcal{J}$ telles que $x \in \mathcal{I}$. ■

Soient A un anneau, et soient $\mathcal{I}$ et $\mathcal{J}$ deux idéaux de A . On notera $\mathcal{I}\mathcal{J}$ le plus petit idéal de A contenant tous les éléments xy pour $x \in \mathcal{I}$ et $y \in \mathcal{J}$. Vérifions que

$$\mathcal{I}\mathcal{J} = \left\{ \sum_i x_i y_i \mid x_i \in \mathcal{I}, y_i \in \mathcal{J} \right\}.$$

Précisons la formule, on considère l'ensemble de toutes les sommes (finies!) de la forme $\sum_i x_i y_i$ avec les conditions imposées. Il est clair que la somme et la différence de deux tels éléments est encore du même type. Si on multiplie par $a \in A$ on obtient $\sum_i a x_i y_i = \sum_i (a x_i) y_i$, qui est encore un élément du même type car $a x_i \in \mathcal{I}$.

Voici une variante du lemme chinois :

Théorème 2. Soient A un anneau, et soient $\mathcal{I}$ et $\mathcal{J}$ deux idéaux de A . Supposons que $\mathcal{I} + \mathcal{J} = \{x + y \mid x \in \mathcal{I}, y \in \mathcal{J}\}$ soit égal à A . L'anneau quotient $A/\mathcal{I}\mathcal{J}$ est alors isomorphe au produit $(A/\mathcal{I}) \times (A/\mathcal{J})$.

Démonstration. On considère l'homomorphisme

$$p : A \longrightarrow (A/\mathcal{I}) \times (A/\mathcal{J}), \quad a \longmapsto (a + \mathcal{I}, a + \mathcal{J}).$$

Cet homomorphisme a pour noyau $\mathcal{I} \cap \mathcal{J}$.

Lemme 1. Sous les hypothèses du théorème on a $\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}$.

Démonstration. L'inclusion $\mathcal{I}\mathcal{J} \subset \mathcal{I} \cap \mathcal{J}$ est claire : soient $x \in \mathcal{I}$ et $y \in \mathcal{J}$ alors $xy \in \mathcal{I} \cap \mathcal{J}$.

Réciproquement l'hypothèse $\mathcal{I} + \mathcal{J} = A$ montre que l'élément $1 \in A$ peut s'écrire $u + v$ où $u \in \mathcal{I}$ et $v \in \mathcal{J}$. Soit alors $z \in \mathcal{I} \cap \mathcal{J}$, on a $z = zu + zv$ et chacun des deux termes du membre de droite est dans $\mathcal{I}\mathcal{J}$. D'où le résultat. $\square$

Revenons à la démonstration du théorème. Nous venons de démontrer que le noyau de l'homomorphisme est $\mathcal{I}\mathcal{J}$. Nous en déduisons donc, par le corollaire à la proposition 3, un homomorphisme injectif

$$A/\mathcal{I}\mathcal{J} \longrightarrow (A/\mathcal{I}) \times (A/\mathcal{J}).$$

Il reste à démontrer que cet homomorphisme est surjectif. Il suffit évidemment de montrer que l'homomorphisme initial p l'est.

Cela résulte de l'argument donné dans la démonstration du lemme. Écrivons 1 comme somme d'un élément $u \in \mathcal{I}$ et $v \in \mathcal{J}$: $1 = u + v$ et soient $a + \mathcal{I} \in A/\mathcal{I}$ et $b + \mathcal{J} \in A/\mathcal{J}$. Alors l'homomorphisme p envoie l'élément $av + bu \in A$ vers $(a + \mathcal{I}, b + \mathcal{J}) \in (A/\mathcal{I}) \times (A/\mathcal{J})$. D'où le résultat. $\blacksquare$

On étend à une famille d'idéaux :

Théorème 3. Soient A un anneau, et soient $\mathcal{I}_\alpha$ une famille finie d'idéaux de A . Supposons que $\mathcal{I}_\alpha + \mathcal{I}_\beta = A$ dès que $\alpha \neq \beta$. Alors on a un isomorphisme

$$A / \prod_{\alpha} \mathcal{I}_{\alpha} \cong \prod_{\alpha} (A / \mathcal{I}_{\alpha}).$$

Démonstration. On fait la démonstration par récurrence sur le nombre d'idéaux dans la famille. Soit n ce nombre. Isolons un des idéaux de la famille, soit $\mathcal{J}$. On

va appliquer le théorème 2 à $\mathcal{J}$ et $\prod_{\mathcal{I}_\alpha \neq \mathcal{J}} \mathcal{I}_\alpha$. Pour cela on doit montrer que :

$$\mathcal{J} + \prod_{\mathcal{I}_\alpha \neq \mathcal{J}} \mathcal{I}_\alpha = A.$$

Ceci se démontre comme suit. Pour tout $\mathcal{I}_\alpha$ distinct de $\mathcal{J}$ on peut, par hypothèse, trouver $x_\alpha \in \mathcal{J}$ et $y_\alpha \in \mathcal{I}_\alpha$ tels que $x_\alpha + y_\alpha = 1$. Faisons le produit sur α de toutes ces identités. Quand on développe le produit tous les termes, sauf le produit des y_α , sont dans $\mathcal{J}$. On a donc bien une relation de la forme $x + y = 1$, avec $x \in \mathcal{J}$ et $y \in \prod_{\mathcal{I}_\alpha \neq \mathcal{J}} \mathcal{I}_\alpha$, ce qui implique la condition.

Le théorème 2 implique alors que :

$$A/\mathcal{J} \times A/\prod_{\mathcal{I}_\alpha \neq \mathcal{J}} \mathcal{I}_\alpha \cong A/\prod_{\mathcal{I}_\alpha} \mathcal{I}_\alpha.$$

Mais l'hypothèse de récurrence donne :

$$A/\prod_{\mathcal{I}_\alpha \neq \mathcal{J}} \mathcal{I}_\alpha \cong \prod_{\mathcal{I}_\alpha \neq \mathcal{J}} A/\mathcal{I}_\alpha.$$

Le résultat suit en combinant ces deux isomorphismes. ■

Définition 11

Soit A un anneau, et soit $\mathcal{J}$ un idéal. On dit que $\mathcal{J}$ est un *idéal premier* si et seulement si pour tous $x, y \in A$ la condition $xy \in \mathcal{J}$ implique $x \in \mathcal{J}$ ou $y \in \mathcal{J}$.

Voici une caractérisation des idéaux premiers à l'aide des anneaux quotients :

Proposition 5. *Un idéal $\mathcal{J}$ d'un anneau A est premier si et seulement si l'anneau quotient $A/\mathcal{J}$ est un domaine d'intégrité.*

Démonstration. Supposons que le quotient est intègre, et soient $x, y \in A$ tels que $\bar{x}\bar{y} = 0$ dans $A/\mathcal{J}$. Cela veut dire que $xy \in \mathcal{J}$. Comme l'anneau quotient est intègre on a par exemple $\bar{x} = 0$, soit $x \in \mathcal{J}$. D'où le résultat.

Inversement si on a pour deux éléments $\bar{x}$ et $\bar{y}$ de l'anneau quotient $\bar{x}\bar{y} = 0$ cela veut dire que $xy \in \mathcal{J}$. Donc $\mathcal{J}$ étant premier cela implique que x ou y , par exemple x , est dans $\mathcal{J}$. Ce qui implique que $\bar{x} = 0$. Donc le quotient est intègre. ■

Définition 12

Soit A un anneau, et soit $\mathcal{J}$ un idéal. On dit que $\mathcal{J}$ est un *idéal maximal* si et seulement si il n'existe pas d'idéal $\mathcal{J}'$ distinct de A et $\mathcal{J}$ tel que $\mathcal{J} \subset \mathcal{J}'$.

Voici une caractérisation des idéaux maximaux à l'aide des anneaux quotients :

Proposition 6. *Un idéal $\mathcal{I}$ d'un anneau A est maximal si et seulement si l'anneau quotient $A/\mathcal{I}$ est un corps.*

Démonstration. Supposons que le quotient soit un corps. Soit $\mathcal{J} \supset \mathcal{I}$, et distinct de $\mathcal{I}$. Soit $x \in \mathcal{J}$ tel que $x \notin \mathcal{I}$. L'élément $\bar{x}$ n'est pas nul dans le quotient et y est donc inversible. Il existe donc y tel que $\bar{x}\bar{y} = 1$, soit $xy = 1 + u$ avec $u \in \mathcal{I}$. Donc $1 = xy - u$ appartient à $\mathcal{J}$ qui est donc égal à A . D'où le résultat.

Réciproquement supposons $\mathcal{I}$ maximal, et soit $\bar{x}$ un élément non nul du quotient $A/\mathcal{I}$, x n'est donc pas dans $\mathcal{I}$. $\perp$

Le plus petit idéal contenant x et $\mathcal{I}$ est constitué par les éléments de la forme $xa + u$, $a \in A$ et $u \in \mathcal{I}$. Il contient strictement $\mathcal{I}$. Comme ce dernier est maximal il est égal à A .

Par conséquent il existe $y \in A$ et $u \in \mathcal{I}$ tels que $1 = xy + u$, et on a dans l'anneau quotient $\bar{x}\bar{y} = 1$. On a donc trouvé un inverse pour $\bar{x}$ et $A/\mathcal{I}$ est un corps. $\blacksquare$

Par exemple les idéaux maximaux de $\mathbb{Z}$ sont ceux qui sont constitués par les multiples d'un nombre premier. En effet il faut montrer que si un idéal contient un nombre premier p et un entier non divisible par p il contient 1. Mais cela résulte de l'identité de Bézout. Ceci sera généralisé dans la prochaine section. Soit p un nombre premier, le corps quotient de $\mathbb{Z}$ par l'idéal $p\mathbb{Z}$ sera noté $\mathbb{F}_p$.

Théorème 4 (Krull). *Tout idéal est contenu dans un idéal maximal.*

La démonstration de ce théorème repose sur l'axiome du choix, nous l'admettrons. Comme conséquence on notera qu'un élément est inversible si et seulement si il n'est contenu dans aucun idéal maximal. En effet si x n'est pas inversible, l'idéal (x) est distinct de A et est donc contenu dans un idéal maximal. La réciproque est claire.

2. Anneaux principaux

On va maintenant étudier une classe d'anneaux qui est particulièrement importante, celle des anneaux principaux. Tous les anneaux que l'on va considérer sont commutatifs.

Définition 1

Un anneau commutatif A est un *anneau principal* si :

- il est intègre,
- tout idéal $\mathcal{I}$ est de la forme

$$(x) = \{xa \mid a \in A\}$$

pour un certain $x \in A$

L'anneau $\mathbb{Z}$ est principal, en effet il est intègre.

Par ailleurs un idéal de $\mathbb{Z}$ n'est autre qu'un sous-groupe de $\mathbb{Z}$. On a démontré au chapitre 1 qu'un sous-groupe de $\mathbb{Z}$ est constitué par les multiples d'un entier n bien défini au signe près, c'est-à-dire est égal à l'idéal (n) .

Rappelons que l'on note A^* l'ensemble des éléments de A qui admettent un inverse pour la multiplication. Notons que si on considère un élément $x \in A$ et un élément $u \in A^*$ les idéaux (x) et (ux) coïncident. En particulier on rappelle que si u est inversible on a $(u) = A$.

Revenons à un anneau quelconque pour définir la notion d'éléments irréductibles :

Définition 2

Soit A un anneau. On dit qu'un élément différent de zéro x est dit un *élément irréductible* si, dès que l'on a une égalité $x = uv$, l'un des éléments, et un seul, u ou v est inversible.

La condition «et un seul» impose à un élément irréductible de ne pas être inversible. Le résultat suivant caractérise les éléments irréductibles dans un anneau principal :

Proposition 1. Soit A un anneau principal, et soit x un élément différent de zéro. Les conditions suivantes sont équivalentes :

- l'élément x est irréductible,
- l'idéal (x) est premier,
- l'idéal (x) est maximal.

Démonstration. Il est clair que la troisième condition implique la seconde, un idéal maximal étant premier car un corps est intègre.

La seconde implique la première. En effet supposons que l'on ait $x = uv$. En particulier $uv \in (x)$, on a donc par exemple, l'idéal (x) étant premier, $u \in (x)$. Donc l'élément u est de la forme xa pour un certain $a \in A$. On a donc $x = xav$, comme A est intègre on obtient $av = 1$ donc v est inversible.

Il reste à montrer que la première condition implique la troisième. Pour ce faire on considère l'anneau quotient $A/(x)$. On doit montrer que c'est un corps, c'est-à-dire que tout élément non nul est inversible.

Considérons donc un élément non nul $\bar{y} \in A/(x)$, et soit y un élément de A dont la classe dans $A/(x)$ est $\bar{y}$. Comme ce dernier élément est non nul cela veut dire que y n'est pas dans l'idéal engendré par x . Considérons alors l'idéal engendré par x et y , soit

$$(x, y) = \{ax + by \mid a, b \in A\}.$$

L'anneau étant principal cet idéal est de la forme (z) pour un certain élément $z \in A$, montrons que $(z) = A$.

D'abord il existe deux éléments u et v tels que $x = zu$ et $y = zv$, puisque x et y sont dans (z) . L'élément x étant irréductible on a soit z , soit u inversible. Dans le premier cas, on a bien $(z) = A$. Dans le second cas les idéaux (x) et (z) coïncident, c'est impossible par hypothèse. On en conclut que 1 appartient à l'idéal (x, y) et qu'il existe $a, b \in A$ tels que $ax + by = 1$. Réduisons cette équation modulo l'idéal (x) on obtient dans l'anneau quotient $A/(x)$ l'identité $\bar{y}\bar{b} = 1$, donc $\bar{y}$ est inversible. ■

L'une des propriétés fondamentales des anneaux principaux est la factorisation unique des éléments en produit de facteurs irréductibles. Ainsi qu'on le verra, ceci ne caractérise pas les anneaux principaux, mais les anneaux factoriels.

Soit A un anneau principal. Commençons par introduire un système de représentants des éléments irréductibles. Précisons ce que l'on entend par là : d'abord deux éléments irréductibles x et y sont dans la même classe s'il existe un élément inversible u tel que $x = yu$. Ceci définit une relation d'équivalence sur les éléments irréductibles. Un système de représentants est un ensemble d'éléments irréductibles x_i , avec $i \in I$, tel que :

- tout élément irréductible x est de la forme $x_i u$ pour un certain indice i et un certain élément inversible u ,
- par ailleurs si $i \neq j$ x_i et x_j ne sont pas en relation, c'est-à-dire qu'il n'existe pas d'élément inversible u tels que $x_i = x_j u$.

Théorème 1. *Soit A un anneau principal. Tout élément $x \in A$, différent de zéro, peut s'écrire sous la forme*

$$u \times \prod_{i \in I} x_i^{\alpha_i}, \quad u \in A^*, \quad \alpha_i \in \mathbb{N},$$

où seuls un nombre fini d'entiers α_i sont non nuls. Si bien que le produit précédent est fini, ce qui donne un sens à la formule. Par ailleurs si on a une égalité

$$u \times \prod_{i \in I} x_i^{\alpha_i} = v \times \prod_{i \in I} x_i^{\beta_i},$$

alors $u = v$ et $\alpha_i = \beta_i$ pour tout $i \in I$.

Démonstration. Démontrons d'abord l'existence de la décomposition. On utilise le lemme suivant :

Lemme 1. *Considérons une suite croissante d'idéaux dans un anneau principal A :*

$$\mathcal{I}_1 \subset \mathcal{I}_2 \subset \cdots \subset \mathcal{I}_i \subset \mathcal{I}_{i+1} \subset \cdots .$$

Il existe un entier N tel que si $i \geq N$ on a $\mathcal{I}_i = \mathcal{I}_{i+1}$.

Démonstration du lemme. On considère l'ensemble $\mathcal{J} = \bigcup_i \mathcal{I}_i$. C'est un idéal. En effet, soient $x, y \in \mathcal{J}$, on a $x \in \mathcal{I}_m$ et $y \in \mathcal{I}_n$ pour certains entiers m et n . Donc x et y appartiennent à $\mathcal{I}_{\sup(m,n)} \subset \mathcal{J}$. Donc l'élément $x - y$ est dans $\mathcal{I}_{\sup(m,n)} \subset \mathcal{J}$. De même si $x \in \mathcal{J}$ et $a \in A$ on a $ax \in \mathcal{I}_u$, pour un certain u , donc $ax \in \mathcal{J}$. L'idéal $\mathcal{J}$ est égal à l'idéal (x) pour un certain $x \in A$, l'élément $x \in \mathcal{J}$ appartient à $\mathcal{I}_N$ pour un certain entier N . On a alors pour tout $i \geq N$:

$$\mathcal{J} = (x) \subset \mathcal{I}_N \subset \mathcal{I}_i \subset \mathcal{J} .$$

donc on a $\mathcal{I}_i = \mathcal{J}$, d'où le résultat. $\square$

Revenons à la démonstration du théorème. On notera qu'il n'est pas clair qu'il existe des éléments irréductibles. La démonstration qui suit démontre tout à la fois qu'il en existe et que l'on peut décomposer un élément quelconque en produit d'éléments irréductibles.

On raisonne par l'absurde et supposons que l'ensemble S des éléments non nuls et non inversibles qui ne peuvent s'écrire comme produit d'éléments irréductibles soit non vide. Soit l'élément x un élément de cet ensemble.

Puisque l'élément n'est pas irréductible il s'écrit $x_1 y_1$ où x_1 et y_1 ne sont pas inversibles. On a une inclusion stricte $(x) \subset (x_1)$, l'inclusion résulte de ce que x_1 divise x , le fait qu'elle est stricte de ce que y_1 n'est pas inversible. Nécessairement l'un des deux éléments x_1 ou y_1 est dans S , sinon x ne serait pas dans S . En effet si x_1 et y_1 s'écrivent comme produit d'éléments irréductibles il suffit de multiplier ces deux produits pour obtenir x comme produit d'éléments irréductibles.

Supposons que $x_1 \in S$. On peut alors écrire $x_1 = x_2 y_2$ avec les mêmes propriétés, x_2 et y_2 ne sont pas inversibles et $x_2 \in S$. On a alors des inclusions strictes $(x) \subset (x_1) \subset (x_2)$. On itère le processus et on obtient une suite *strictement* croissante d'idéaux :

$$(x) \subset (x_1) \subset \cdots \subset (x_n) \subset \cdots ,$$

ce qui est en contradiction avec le lemme 1. Cela démontre l'existence d'une décomposition en produit d'éléments irréductibles, et par conséquent l'existence d'éléments irréductibles.

Passons maintenant à l'unicité, on la démontre par récurrence. Supposons que l'on ait avec les notations du théorème :

$$u \times \prod_{i \in I} x_i^{\alpha_i} = v \times \prod_{i \in I} x_i^{\beta_i},$$

bien entendu seuls un nombre fini d'exposants sont non nuls. On fait la récurrence sur l'entier $\sup(\sum_i \alpha_i, \sum_i \beta_i)$. Insistons bien sur le fait que ces quantités sont finies puisque seuls un nombre fini d'exposants sont différents de zéro. Considérons un des éléments irréductibles apparaissant à gauche avec un exposant non nul, soit x_k . Le terme de droite est dans l'idéal engendré par x_k .

Rappelons que si x est élément d'un anneau commutatif on dit que x divise $a \in A$ et on note $x \mid a$ si cet élément est dans l'idéal engendré par x , donc si $a = xu$ pour un certain $u \in A$.

Le lemme qui suit montre alors que x_k apparaît aussi dans le membre de droite.

Lemme 2 (Lemme d'Euclide). *Soit A un anneau principal, soit x un élément irréductible. Supposons que x divise le produit ab , $a, b \in A$. Alors x divise a ou b .*

Démonstration du lemme. Raisonnons par l'absurde, supposons que x ne divise ni a ni b . La proposition 1 implique que $A/(x)$ est un corps. par conséquent les classes de a et b sont inversibles dans ce corps et il existe u, v et u', v' tels que $ux + va = 1$ et $u'x + v'b = 1$.

Faisons le produit de ces deux égalités, on obtient :

$$uu'x^2 + uv'xb + vu'ax + vv'ab = 1,$$

x divise le membre de gauche, et divise donc 1, donc est inversible, c'est impossible car x est irréductible. $\square$

Achevons la démonstration du théorème. Le lemme implique que x_k apparaît à droite avec un exposant différent de zéro. On peut donc diviser à gauche et à droite par x_k , A étant intègre, et donc se ramener au pas précédent de la récurrence. ■

On observera qu'un élément x divise un élément y si et seulement si l'exposant de tout élément irréductible dans la décomposition de x est inférieur ou égal à celui de la décomposition de y .

Le lemme d'Euclide se généralise en :

Lemme 3 (lemme de Gauss). *Soit A un anneau principal, soit $x \in A$. Supposons que x divise le produit ab , $a, b \in A$ et que le pgcd de x et a soit 1. Alors x divise b .*

La démonstration est laissée en exercice.

Pour définir le pgcd de deux éléments x et y différents de zéro on va procéder à l'envers de la démarche standard. Celle-ci consiste à le définir comme étant un générateur de l'idéal engendré par x et y et à en décrire après la décomposition en nombres premiers. On le définit avec le ppcm par les formules qui suivent. Soient $x = u \times \prod_{i \in I} x_i^{\alpha_i}$ et $y = v \times \prod_{i \in I} x_i^{\beta_i}$ les décompositions de x et y . Alors par définition :

$$\text{pgcd}(x, y) = \prod_{i \in I} x_i^{\inf(\alpha_i, \beta_i)}, \quad \text{et} \quad \text{ppcm}(x, y) = \prod_{i \in I} x_i^{\sup(\alpha_i, \beta_i)}.$$

Plus généralement on définit le pgcd et le ppcm d'une famille finie d'éléments non nuls

$$y_m = u_m \times \prod_{i \in I} x_i^{\alpha_{m,i}},$$

$1 \leq m \leq n$, par :

$$\text{pgcd}(y_1, \dots, y_n) = \prod_{i \in I} x_i^{\inf(\alpha_{m,i})_{1 \leq m \leq n}}.$$

Le pgcd d'une famille finie d'éléments quelconques est le pgcd de la sous-famille des éléments non nuls.

Il est commode d'étendre la définition précédente en disant qu'un élément est un pgcd d'une famille d'éléments s'il est égal à l'élément précédent, à multiplication près par un élément inversible. Cet abus allégera souvent les énoncés.

On notera aussi, avec les notations précédentes pour x et y , que x divise y si et seulement si pour tout i on a $\alpha_i \leq \beta_i$. Cela a déjà été signalé plus haut. On a :

Proposition 3 (identité de Bézout). Soient x et y des éléments différents de zéro d'un anneau principal A , et soit v leur pgcd. Il existe des éléments a et b de A tels que

$$ax + by = v.$$

En particulier deux éléments différents de zéro x et y sont premiers entre eux, c'est-à-dire n'ont pas de diviseurs non trivial commun si et seulement si il existe u et v tels que

$$ux + vy = 1.$$

Soit y_i une famille d'éléments différents de zéro d'un anneau principal A , et soit v leur pgcd. Il existe des éléments $a_i \in A$ tels que

$$\sum_i a_i y_i = v.$$

En particulier des éléments différents de zéro y_i sont premiers entre eux, c'est-à-dire n'ont pas de diviseurs non trivial commun si et seulement si il existe $a_i \in A$ tels que

$$\sum_i a_i y_i = 1.$$

Démonstration. Considérons le cas de deux éléments. Soit v un générateur de l'idéal (x, y) . Le pgcd de x et y divisant x et y , il divise v .

Inversement montrons que le pgcd de x et y est divisible par v . Comme v est générateur de l'idéal (x, y) , il divise x et il divise y . Si bien que, si on considère les décompositions en produit de facteurs irréductibles de x , y et v on obtient $x = u_0 \times \prod_{i \in I} x_i^{\alpha_i}$, $y = u_1 \times \prod_{i \in I} x_i^{\beta_i}$ et $v = u_2 \times \prod_{i \in I} x_i^{\gamma_i}$ on a $\gamma_i \leq \inf(\alpha_i, \beta_i)$. Le résultat suit. ■

Le résultat s'étend à une famille finie d'éléments sans difficultés.

Anneaux euclidiens

Pour conclure cette section décrivons une classe particulière d'anneaux principaux.

Définition 3

Un anneau A commutatif intègre est un *anneau euclidien* s'il est muni d'une application $s : A - \{0\} \rightarrow \mathbb{N} - \{0\}$ telle que :

- pour tous $x, y \in A$, $y \neq 0$, il existe $q, r \in A$ tels que $x = yq + r$ et $r = 0$ ou $s(r) < s(y)$,
- r est appelé le reste de la division de x par y .

Par exemple $\mathbb{Z}$ est euclidien en prenant $s(x) = |x|$.

Proposition 4. *Un anneau euclidien est principal.*

Démonstration. C'est la même que dans $\mathbb{Z}$. Soit $\mathcal{I}$ un idéal non réduit à 0. Soit y un élément différent de zéro de l'idéal tel que $s(y)$ soit minimal. Soit maintenant x quelconque dans l'idéal, on a $x = yq + r$, donc $r = x - yq$ est dans l'idéal. Mais $s(r) < s(y)$, par définition de y la seule possibilité est $r = 0$, ce qui montre que $\mathcal{I} = (y)$. ■

Dans un anneau euclidien on a un algorithme d'Euclide pour calculer le pgcd exactement comme dans le cas de $\mathbb{Z}$. On peut aussi l'utiliser pour calculer les coefficients de la formule de Bézout. Rappelons le brièvement.

On a d'abord :

Lemme 4. Soient $x, y \in A$, où A est un anneau principal. Soit $x = yq + r$ la division euclidienne de x par y . Alors le pgcd de x et y est égal au pgcd de y et r si $r \neq 0$, à y sinon.

Démonstration. La démonstration est laissée au lecteur. □

Passons à la description de l'algorithme. Posons $r = r_0$. On itère le processus, on écrit la division de y par r_0 , qui donne un reste r_1 . Le pgcd de x et y est alors égal à celui de r_0 et r_1 . Ceci pourvu que r_0 ne divise pas y , il est égal à r_0 sinon. Dans le premier cas on itère le processus et on définit ainsi r_2 . Plus généralement supposons avoir défini r_{n-1} et r_n . Alors, soit r_n divise r_{n-1} et le pgcd de x et y qui est le pgcd r_{n-1} et r_n est r_n . Soit on définit r_{n+1} comme le reste de la division de r_{n-1} par r_n , et le pgcd de x et y est le pgcd de r_n et r_{n+1} . Comme la suite $s(r_n)$ est une suite d'entiers positifs strictement décroissante ce processus s'achève après un nombre fini d'itérations. Le pgcd de x et y est le dernier reste non nul dans cet algorithme.

Cet algorithme permet aussi de trouver les coefficients de la formule de Bézout. En effet, partons du pgcd qui est le dernier reste non nul r_n . On a $r_n = -r_{n-1}q_{n-1} + r_{n-2}$. À l'étape précédente on avait $r_{n-1} = -r_{n-2}q_{n-1} + r_{n-3}$. Dans la première équation on peut donc remplacer r_{n-1} par sa valeur en fonction de r_{n-2} et r_{n-3} . On obtient donc une expression de r_n en fonction de r_{n-2} et r_{n-3} . À l'étape suivante on exprime r_{n-2} en fonction de r_{n-3} et r_{n-4} , et on substitue. De proche en proche on obtient une expression de r_n en fonction de x et y . ■

À titre d'exercice le lecteur pourra écrire le lemme chinois dans le contexte présent.

3. L'anneau $\mathbb{Z}/n\mathbb{Z}$

Un idéal de $\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$ et réciproquement. La proposition 5 de la section 1 du chapitre 1 classe donc les idéaux de $\mathbb{Z}$ ainsi qu'on l'a déjà dit plus haut.

Proposition 1. Soit $\mathcal{I}$ un idéal de $\mathbb{Z}$ non réduit à zéro. Il existe un unique entier positif non nul a tel que $\mathcal{I} = a\mathbb{Z}$.

Rappelons la définition de la fonction, ou indicateur, d'Euler :

Définition 1

La fonction d'Euler φ de $\mathbb{N}$ dans $\mathbb{N}$ est définie par :

$$\varphi(n) = \#\{i \mid 1 \leq i \leq n \text{ et } \text{pgcd}(i, n) = 1\}.$$

Notons que $\varphi(1) = 1$.

Montrons que $\varphi(n)$ est aussi le cardinal de $(\mathbb{Z}/n\mathbb{Z})^*$.

Les éléments non nuls de $\mathbb{Z}/n\mathbb{Z}$ sont en bijection avec les entiers compris entre 1 et $n - 1$. Par ailleurs si un tel entier a est premier à n (si son pgcd avec n est égal à 1) l'identité de Bézout implique l'existence de u et v tels que $ua + vn = 1$. Réduisant modulo n on obtient $\bar{a} \bar{u} = 1$, donc $\bar{a}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$.

Inversement si la classe d'un élément $\bar{x}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$, il existe $\bar{k}$ tel que $\bar{x} \bar{k} = 1$. Soit $xk = 1 - un$ pour un certain entier u . Ce qui implique que x et n sont premiers entre eux par Bézout.

La formule suivante sera utile plus loin :

Proposition 2

$$n = \sum_{d|n} \varphi(d) = \sum_{d|n} \varphi\left(\frac{n}{d}\right).$$

Démonstration. Soit d un diviseur de n . On définit un sous-ensemble C_d de $\{1, \dots, n\}$ par $C_d = \{i \mid \text{pgcd}(i, n) = \frac{n}{d}\}$. Les sous-ensembles C_d constituent une partition de $\{1, \dots, n\}$ quand d décrit les diviseurs de n . Donc, on a

$$n = \sum_{d|n} \# C_d.$$

D'autre part on a $\# C_d = \varphi(d)$, en effet un entier u dans C_d s'écrit $\frac{n}{d} \times v$. L'entier v devant être premier à n donc à d et compris entre 1 et d . Il y a bien $\varphi(d)$ entiers satisfaisant à ces deux conditions. Le résultat suit en substituant dans la formule ci-dessus. ■

Calculons maintenant la fonction φ .

Proposition 3. Si m et n sont premiers entre eux on a la relation :

$$\varphi(mn) = \varphi(m)\varphi(n).$$

Démonstration. Ceci résulte du lemme chinois (théorème 2 section 1). Comme m et n sont premiers entre eux $\mathbb{Z}/mn\mathbb{Z}$ est isomorphe au produit $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. L'ensemble des éléments inversibles de $\mathbb{Z}/mn\mathbb{Z}$, qui est de cardinal $\varphi(mn)$ est donc isomorphe à $(\mathbb{Z}/m\mathbb{Z})^* \times (\mathbb{Z}/n\mathbb{Z})^*$, qui est de cardinal $\varphi(m)\varphi(n)$. ■

Il en résulte que :

Corollaire. Soit n un entier positif non nul, et soit $p_1^{\alpha_1} \dots p_k^{\alpha_k}$ la décomposition de cet entier en puissances de nombres premiers. On a :

$$\varphi(n) = \prod_{i=1, \dots, k} \varphi(p_i^{\alpha_i}).$$

Il nous reste à calculer la fonction d'Euler sur les puissances des nombres premiers :

Théorème 1. Soit p un nombre premier impair, on a :

$$\varphi(p^n) = (p-1)p^{n-1}$$

si $n \geq 1$.

Ce théorème se démontre en comptant les entiers inférieurs ou égaux à p^n premier à p , c'est-à-dire qui ne sont pas divisibles par p . On compte le nombre d'éléments dans le complémentaire, c'est-à-dire le nombre d'entiers divisible par p . Ceux-ci sont de la forme pu avec $1 \leq u \leq p^{n-1}$, il y en a p^{n-1} , il y a donc $p^n - p^{n-1}$ éléments premiers avec p .

En fait ce résultat est un corollaire du théorème plus précis suivant :

Théorème 2

- Le groupe $(\mathbb{Z}/2^n\mathbb{Z})^*$ est isomorphe à $\mathbb{Z}/2\mathbb{Z}$ si $n = 2$ et à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}$ si $n > 2$,
- le groupe $(\mathbb{Z}/p^n\mathbb{Z})^*$ est isomorphe à $\mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{n-1}\mathbb{Z}$ si $n \geq 1$.

La démonstration est proposée en exercice.

4. Cryptographie à clé révélée

Dans cette section on va donner une brève introduction à la *cryptographie*. Plus précisément, on va donner une description rapide et très simplifiée du *système RSA* (pour R. Rivest, A. Shamir et L. Adleman) et de celui du « sac à dos ». Le but de la cryptographie est de rendre des messages inintelligibles à toute personne qui ne possède pas la clé pour les traduire. Donnons un exemple classique (utilisé par J. César !) pour commencer. Considérons les 26 lettres de l'alphabet et une permutation σ quelconque de cet ensemble. Par exemple la permutation cyclique qui envoie a sur b , b sur c , ..., z sur a . Si on applique cette transformation à un mot, c'est-à-dire à chaque lettre substituer son image par la transformation, on obtient par exemple pour le mot *oui*, le mot *pvj*. Pour déchiffrer le message on applique évidemment la transformation inverse. Le problème avec cet exemple classique est que la fonction σ et son inverse peuvent être facilement déduites de quelques messages. Il suffit en effet de connaître la fréquence d'apparition des lettres dans un texte d'une langue donnée. À partir d'un échantillon de messages suffisant le décryptage est rapide. Le problème est donc de rendre pratiquement impossible la détermination de l'application inverse.

Le système RSA fournit une solution à ce problème, on renvoie à [De] et [Ko] pour compléter la description qui suit.

On commence par compliquer le cryptage en procédant comme suit. Au lieu de remplacer une lettre par une autre on peut remplacer une suite, disons de k lettres

ou symboles successifs (éventuellement des blancs) par un autre symbole. Cela veut dire construire une application de A^k , où A est notre alphabet usuel auquel on a adjoint un certain nombre de symboles (virgule, point, blanc, ...) vers un autre ensemble B . On donnera un exemple plus loin.

Le système RSA

On passe maintenant au système RSA proprement dit. On considère deux personnes appelées IV et OA par exemple, souhaitant communiquer entre elles en toute confidentialité. Elles choisissent chacun une paire de nombres premiers « grands », p_x et q_x pour IV, p_y et q_y pour OA. On pose $n_x = p_x q_x$ et $n_y = p_y q_y$. Puis IV choisit également un entier c_x premier avec $\varphi(n_x) = (p_x - 1)(q_x - 1)$ et OA un entier c_y premier avec $\varphi(n_y) = (p_y - 1)(q_y - 1)$.

La classe de congruence de l'entier c_x admet un inverse d_x dans le groupe multiplicatif $(\mathbb{Z}/n_x\mathbb{Z})^*$. Cette classe est facile à calculer rapidement dans la mesure où l'on connaît la valeur de $\varphi(n_x)$. De même la classe de congruence de l'entier c_y admet un inverse d_y dans le groupe multiplicatif $(\mathbb{Z}/n_y\mathbb{Z})^*$, elle est rapide à calculer dans la mesure où l'on connaît la valeur de $\varphi(n_y)$. Nous ne rentrerons pas dans les détails mais ce calcul, *en temps*, est court.

Mais si on ne connaît que les valeurs n_x et c_x , respectivement n_y et c_y , le calcul devient très difficile, c'est-à-dire très long en temps. La seule façon connue de procéder en général est de chercher la décomposition en nombres premiers de n_x et n_y , puis d'en déduire les valeurs de la fonction d'Euler.

La décomposition en nombres premiers est facile théoriquement, on divise successivement par tous les nombres premiers, par contre, elle est extrêmement longue du point de vue du temps de calcul, au point de devenir inaccessible dès que les nombres en jeu sont très grands. Pratiquement le calcul de d_x , respectivement d_y est impossible à partir de n_x et c_x , respectivement de n_y et c_y .

Les messages sont alors codés de la façon suivante. D'abord les lettres des alphabets B introduit plus haut sont identifiées à des classes de congruences inversibles modulo n_x d'une part, à des classes de congruences modulo n_y d'autre part. On entend par là que l'on se donne une application injective de

$$\alpha : B \longrightarrow (\mathbb{Z}/n_x\mathbb{Z})^*$$

et une application injective

$$\beta : B \longrightarrow (\mathbb{Z}/n_y\mathbb{Z})^* .$$

Ces identifications sont publiques et connues des correspondants.

On ne considère donc maintenant que des suites de symboles qui sont des classes de congruences.

Définition

Les entiers n_x, c_x et n_y, c_y sont les clés publiques. Les entiers p_x, q_x, d_x et p_y, q_y, d_y sont les clés privées.

Ces dernières ne sont connues que de IV, et respectivement de OA. Bien entendu les valeurs de $\varphi(n_x) = (p_x - 1)(q_x - 1)$ et $\varphi(n_y) = (p_y - 1)(q_y - 1)$ sont aussi secrètes (privées).

Supposons alors que IV veuille coder un message destiné à OA et que ce message soit exprimé comme une suite de symboles appartenant à $(\mathbb{Z}/n_y\mathbb{Z})^*$. Il applique à cette suite de symboles, soit $(a_1, a_2, \dots, a_k, \dots)$ la transformation :

$$(a_1, a_2, \dots, a_k, \dots) \mapsto (a_1^{c_y}, a_2^{c_y}, \dots, a_k^{c_y}, \dots),$$

puis il envoie le message. Le calcul effectif n'est pas prohibitif en temps.

Quand OA reçoit le message il applique la transformation inverse :

$$(b_1, b_2, \dots, b_k, \dots) \mapsto (b_1^{d_y}, b_2^{d_y}, \dots, b_k^{d_y}, \dots).$$

Ce qu'il est seul à pouvoir faire dans la mesure où il est le seul à connaître d_y . Ceci restitue le message émis. En effet, le k -ième terme de la suite est $a_k^{c_y d_y}$. Comme $c_y d_y$ est égal à 1 modulo $\varphi(n_y)$ on a $a_k^{c_y d_y} = a_k$. Rappelons en effet que a_k appartient au groupe multiplicatif $(\mathbb{Z}/n_y\mathbb{Z})^*$ qui est d'ordre $\varphi(n_y)$. Encore une fois le calcul effectif n'est pas prohibitif en temps.

En fait on peut faire mieux. On peut garantir l'origine du message, c'est-à-dire que OA pourra être sûr que c'est IV qui lui a envoyé le message. On va se placer dans un cas particulier et supposer que $\varphi(n_x) < \varphi(n_y)$. Pour faire cela IV va procéder comme suit. Il commence par considérer son message comme une suite d'éléments non pas de $(\mathbb{Z}/n_y\mathbb{Z})^*$ mais de $(\mathbb{Z}/n_x\mathbb{Z})^*$. Il applique alors la transformation

$$(a_1, a_2, \dots, a_k, \dots) \mapsto (a_1^{d_x}, a_2^{d_x}, \dots, a_k^{d_x}, \dots),$$

qu'il est seul à pouvoir effectuer. Il considère alors la suite résultante, soit $(b_1, b_2, \dots, b_k, \dots)$, comme une suite d'éléments de $\mathbb{Z}/n_y\mathbb{Z}$. Pour faire ce pas il utilise une fonction injective (c'est ici que l'on utilise l'hypothèse $\varphi(n_x) < \varphi(n_y)$) de $(\mathbb{Z}/n_x\mathbb{Z})^*$ dans $(\mathbb{Z}/n_y\mathbb{Z})^*$ convenue d'avance avec OA et donc connue de OA. Il applique alors la transformation

$$(b_1, b_2, \dots, b_k, \dots) \mapsto (b_1^{c_y}, b_2^{c_y}, \dots, b_k^{c_y}, \dots).$$

et transmet le message.

Pour décoder le message OA procède donc essentiellement en deux étapes. D'abord il applique

$$(f_1, f_2, \dots, f_k, \dots) \mapsto (f_1^{d_y}, f_2^{d_y}, \dots, f_k^{d_y}, \dots).$$

qu'il est seul à pouvoir effectuer. Il obtient donc ainsi un message composé d'éléments de $\mathbb{Z}/n_y\mathbb{Z}$ qu'il réinterprète comme un message composé d'une suite d'éléments $(g_1, g_2, \dots, g_k, \dots)$ dans $\mathbb{Z}/n_x\mathbb{Z}$. Il applique alors la transformation publique :

$$(g_1, g_2, \dots, g_k, \dots) \mapsto (g_1^{c_x}, g_2^{c_x}, \dots, g_k^{c_x}, \dots).$$

Si ce message n'a pas été transmis par le correspondant x , mais par un autre correspondant utilisant une autre clé de cryptage, le résultat sera incompréhensible. Il ne peut être compréhensible que s'il a été transmis par IV.

On parlera alors de message signé.

Décrivons maintenant le processus sur un exemple. Supposons que l'on veuille transmettre le mot *cryptage*. On commence par le diviser en paquets de 2 lettres successives. On obtient donc la suite *cr, yp, ta, ge*. Choisissons $p_y = 19$ et $q_y = 43$. On a $n_y = 817$ et $\varphi(n_y) = 756$.

Supposons avoir choisi une application β de A^2 , où A est l'alphabet usuel dans $(\mathbb{Z}/817\mathbb{Z})^*$. C'est possible, il suffit de comparer les cardinaux. Supposons avoir $\beta(cr) = 2$, $\beta(yp) = 11$, $\beta(ta) = 3$, et $\beta(ge) = 8$.

OA choisit maintenant $c_y = 11$ et à l'aide de l'algorithme d'Euclide, trouve $d_y = 275$.

IV connaît la valeur de c_y et donc considère la suite suivante

$$(2^{11}, 11^{11}, 3^{11}, 8^{11}),$$

il la réduit modulo 817, et obtient

$$(414, 64, 675, 677)$$

et transmet.

OA applique alors la transformation suivante :

$$(414, 64, 675, 677) \mapsto (414^{275}, 64^{275}, 675^{275}, 677^{275}),$$

il est le seul à pouvoir la faire, étant le seul à connaître d_y . Il retrouve $(2, 11, 3, 8)$.

Le lecteur vérifiera à titre d'exercice les calculs et calculera également le message signé, au sens défini plus haut, avec $p_x = 23$, $q_x = 37$, $c_x = 5$, $d_x = 317$. Il choisira une fonction α .

Le système du sac à dos

On passe maintenant au problème dit du sac à dos. Ce problème donne un autre moyen de crypter des informations. On va dans ce cas comparer brièvement les nombres d'opérations à effectuer pour décrypter un message selon que l'on dispose de la clé ou non. Pour nous, « nombre d'opérations » signifiera nombre d'additions, multiplications, divisions, et non d'opérations sur un ordinateur. Ceci n'est donc qu'une approximation très grossière (et dangereuse) du problème, car chaque opération joue elle-même un grand rôle selon l'algorithme qui sert à l'exécuter sur ordinateur. Il convient d'ailleurs de rappeler que l'on effectue les calculs en base 2. On renvoie à [Ko] pour une analyse plus détaillée ainsi qu'à l'exercice 24, section 1. Nous la faisons simplement pour attirer l'attention sur son importance pratique.

On se donne une suite d'entiers positifs ou nuls $\{v_1, v_2, \dots, v_n\}$ ainsi qu'un entier positif V . Le problème du sac à dos est de savoir si on peut trouver un sous-ensemble de $\{v_1, v_2, \dots, v_n\}$ tel que la somme des entiers appartenant au sous-ensemble est égale à V . D'une manière plus formelle il s'agit de déterminer une suite (si elle existe) $\{e_1, e_2, \dots, e_n\}$ d'entiers égaux à 0 ou 1 telle que $\sum_i e_i v_i = V$. Comme il y a 2^n telles suites, pour résoudre le problème on doit tester toutes les possibilités et donc effectuer un nombre d'opérations arithmétiques (2^n sans tenir compte du nombre d'additions à effectuer dans chaque cas) qui croît au moins exponentiellement avec n . Sauf, bien entendu, si on dispose d'un algorithme qui permet d'éviter de faire tous les tests. Des cas particuliers de ce problème sont plus rapides à résoudre. Le problème « super croissant », que l'on notera **sc**, est de résoudre le problème précédent dans le cas où on suppose que pour tout i on a $\sum_{1 \leq j \leq i-1} v_j < v_i$.

Proposition. *Le problème **sc** peut être résolu en un nombre d'opérations arithmétiques qui croît polynomialement avec n*

En effet, on commence par choisir le plus grand entier v_{a_1} inférieur ou égal à V . Comme la somme $\sum_{i=1, \dots, a_1-1} v_i$ est strictement inférieure à v_{a_1} , une somme partielle des v_j égale à V doit contenir ce terme. On doit donc faire n -opérations au plus, puis on applique le même processus à $V - v_{a_1}$ et ainsi de suite, on aboutit donc au plus $\frac{n(n+1)}{2}$ opérations arithmétiques.

Venons en au processus de cryptage. On choisit un entier C tel que $\sum_i v_i < C$, et un entier $\alpha < C$ premier avec C . On calcule l'inverse modulo C β de α , plus précisément son représentant qui est strictement inférieur à C . On calcule de même les entiers positifs w_i inférieurs strictement à C qui sont congrus à αv_i modulo C .

Définition

La clé publique est la donnée des entiers w_i , $1 \leq i \leq n$. La clé privée les entiers v_i , $1 \leq i \leq n$, α , β , C

Si l'on veut transmettre au détenteur des clés publiques et privées un nombre binaire $\{e_i, 1 \leq i \leq n\}$ de longueur k , on procédera donc comme suit : on calcule l'entier $\ell = \sum_i e_i w_i$ et on le transmet. Le destinataire du message connaît les entiers C , α et β . Il peut donc multiplier ℓ par β qui est égal à $\sum_i e_i \alpha \beta w_i = \sum_i e_i \alpha \beta w_i$ modulo C et calculer le reste de la division par C . À cause de la condition $\sum_i v_i < C$ ce reste est égal à $\sum_i e_i v_i$. Il peut alors résoudre le problème **sc** rapidement puisque le problème auquel il fait face est **sc**.

Ceci ne peut être fait avec les w_i qui ne vérifient pas la condition requise. Cet algorithme est dû à M. E. Hellman et R. C. Merkle. Il convient cependant d'ajouter qu'il existe (A. Shamir 1982) une méthode pour décrypter ces messages rapidement, sans connaître C et α , cette méthode n'est donc finalement pas fiable.

5. Anneaux de polynômes et de séries formelles

Soit A un anneau, il sera dans toute cette section supposé commutatif. Soit $\mathbb{N}$ l'ensemble des entiers naturels, introduisons l'ensemble des suites à coefficients dans A , c'est-à-dire l'ensemble des applications de $\mathbb{N}$ dans A . Une telle suite sera pour l'instant notée $(a_0, a_1, \dots, a_i, \dots)$, l'élément a_i désignant l'image de l'entier i .

Définissons une addition sur cet ensemble par la formule

$$(a_0, a_1, \dots, a_i, \dots) + (a'_0, a'_1, \dots, a'_i, \dots) = (a_0 + a'_0, a_1 + a'_1, \dots, a_i + a'_i, \dots).$$

On additionne termes à termes. Cette loi est associative et commutative. Elle admet l'élément $(0, \dots, 0, \dots)$ pour élément neutre. L'élément $(-a_0, -a_1, \dots, -a_i, \dots)$ est l'inverse de $(a_0, a_1, \dots, a_i, \dots)$.

Définissons une multiplication par la formule

$$\begin{aligned} (a_0, a_1, \dots, a_i, \dots) \times (a'_0, a'_1, \dots, a'_i, \dots) &= \dots \\ \dots &= (a_0 a'_0, a_0 a'_1 + a_1 a'_0, \dots, a_i a'_0 + a_{i-1} a'_1 + \dots + a_0 a'_i, \dots). \end{aligned}$$

Cette loi admet l'élément $(1, 0, \dots, 0, \dots)$ comme élément neutre. Elle est associative, considérons le i -ième terme du produit

$$((a_0, a_1, \dots, a_i, \dots)(b_0, b_1, \dots, b_i, \dots))(c_0, c_1, \dots, c_i, \dots),$$

tous calculs faits il est égal à :

$$\sum_{(k,l,m) \mid k+l+m=i} a_k b_l c_m.$$

On obtient le même résultat en calculant

$$(a_0, a_1, \dots, a_i, \dots)((b_0, b_1, \dots, b_i, \dots)(c_0, c_1, \dots, c_i, \dots)).$$

Enfin, la multiplication est distributive par rapport à l'addition. En effet, on a :

$$\begin{aligned} (a_0, a_1, \dots, a_i, \dots)((b_0, b_1, \dots, b_i, \dots) + (c_0, c_1, \dots, c_i, \dots)) &= \\ (a_0, a_1, \dots, a_i, \dots)(b_0 + c_0, b_1 + c_1, \dots, b_i + c_i, \dots) &= \\ (a_0 b_0 + a_0 c_0, \dots, a_0(b_i + c_i) + \dots + a_i(b_0 + c_0), \dots), \end{aligned}$$

d'un côté. De l'autre on a :

$$\begin{aligned} (a_0, \dots, a_i, \dots)(b_0, \dots, b_i, \dots) + (a_0, \dots, a_i, \dots)(c_0, \dots, c_i, \dots) &= \\ (a_0 b_0, \dots, a_i b_0 + \dots + a_0 b_i, \dots) + (a_0 c_0, \dots, a_i c_0 + \dots + a_0 c_i, \dots) &= \\ (a_0 b_0 + a_0 c_0, \dots, a_i b_0 + \dots + a_0 b_i + a_i c_0 + \dots + a_0 c_i, \dots), \end{aligned}$$

d'où le résultat.

L'ensemble des suites d'éléments de A possède donc une structure d'anneau. La suite $(a_0, a_1, \dots, a_n, \dots)$ sera notée

$$\sum_{i \geq 0} a_i X^i$$

dans la suite et sera appelée une *série formelle*, X est appelée l'*indéterminée*.

Définition 1

L'ensemble des suites d'éléments à coefficients dans A muni de l'addition et de la multiplication définie plus haut est par définition l'anneau des séries formelles en X à coefficients dans A . Il est noté $A[[X]]$.

Les constructions pédantes faites ci-dessus ne doivent pas masquer que l'on additionne et que l'on multiplie les séries formelles comme on le fait pour les séries ordinaires, à savoir :

$$\sum_{i \geq 0} a_i X^i + \sum_{i \geq 0} b_i X^i = \sum_{i \geq 0} (a_i + b_i) X^i,$$

et

$$\left(\sum_{i \geq 0} a_i X^i \right) \left(\sum_{i \geq 0} b_i X^i \right) = \sum_{i \geq 0} \left(\sum_{k+l=i} a_k b_l \right) X^i.$$

Les a_i sont appelés les coefficients de la série formelle.

Supposons maintenant que dans la suite tous les termes sauf un nombre fini d'entre eux soient nuls, *i.e.* que presque tous les coefficients soient nuls :

Proposition 1. *Le sous-ensemble de $A[[X]]$ constitué par les séries formelles dont presque tous les coefficients sont nuls est un sous-anneau de $A[[X]]$.*

Démonstration. En effet, soient deux séries formelles

$$\sum_{i \geq 0} a_i X^i, \quad \sum_{i \geq 0} b_i X^i$$

dont presque tous les coefficients sont nuls. Les coefficients a_i sont donc tous nuls dès que i est assez grand, disons $i \geq C$, de même les coefficients b_j sont nuls dès que $j \geq D$. Les coefficients de $\sum_{i \geq 0} a_i X^i + \sum_{i \geq 0} b_i X^i$ sont nuls dès que $i \geq \sup(C, D)$. Ceux de

$$\left(\sum_{i \geq 0} a_i X^i \right) \left(\sum_{i \geq 0} b_i X^i \right)$$

le sont dès que $i \geq C + D - 1$.

Ceci montre que l'addition et la multiplication sont des lois internes sur ce sous-ensemble, et en fait donc un sous-anneau de $A[[X]]$. ■

Définition 2

Le sous-anneau de $A[[X]]$ constitué par les séries formelles dont les coefficients sont presque tous nuls est par définition l'anneau des polynômes en X à coefficients dans A . Il est noté $A[X]$.

Évidemment un *polynôme* sera noté comme à l'ordinaire

$$\sum_0^n a_i X^i,$$

avec les règles usuelles d'addition et de multiplication.

Si tous les termes, sauf peut être a_0 , sont nuls le polynôme est dit constant.

Voici une propriété facile mais fondamentale des anneaux de polynômes sur un corps k . Elle sera utilisée fréquemment. Définissons d'abord une algèbre sur le corps k .

Définition 3

On dit qu'un anneau B est une algèbre sur un corps k s'il existe un homomorphisme i (nécessairement injectif) de k dans B .

Ceci implique que B est un espace vectoriel sur k . L'anneau $k[X]$ en particulier est une algèbre sur k . L'homomorphisme $k \rightarrow k[X]$ est donné par $a \mapsto (a, 0, \dots)$. Autrement dit on associe à l'élément a le polynôme $aX^0 = a$. Rappelons que ces polynômes sont dits constants (ou de degré zéro, voir plus loin). On comparera au chapitre 4 où la structure d'espace vectoriel d'un corps sur un sous-corps est étudiée en détail.

Proposition 2. Soit B une algèbre sur k , et soit $\alpha \in B$. Il existe un homomorphisme d'anneaux et un seul $f : k[X] \rightarrow B$ tel que :

- $f(X) = \alpha$
- f restreint aux polynômes constants s'identifie à l'homomorphisme $i : k \rightarrow B$ déterminant la structure d'algèbre.

Démonstration. On identifie dans la suite un élément de k et son image dans B pour alléger les notations. Soit $P = \sum_i a_i X^i$, $a_i \in k$. Posons alors $f(P) = \sum_i a_i \alpha^i$. C'est un homomorphisme, en effet $f(P+Q) = f(P) + f(Q)$ et $f(PQ) = f(P)f(Q)$. Il satisfait aux conditions requises. D'autre part si un homomorphisme φ satisfait aux conditions requises, il est forcément donné par cette formule car :

$$\varphi(P) = \varphi\left(\sum_i a_i X^i\right) = \sum_i a_i \varphi(X^i) = \sum_i a_i \alpha^i. \quad \blacksquare$$

On dira que l'on évalue le polynôme en x , ou que l'on remplace l'indéterminée par x .

Définition 4

On dit qu'un polynôme non nul $P = \sum_0^N a_i X^i$ est de *degré* inférieur ou égal à n si ses coefficients a_i sont nuls pour $i > n$. Il est alors dit de degré n si $a_n \neq 0$. Le degré de P sera noté $|P|$ ou encore $\deg(P)$.

Il est souvent commode d'interpréter le polynôme nul comme étant de degré $-\infty$. Notons que l'on a la formule

$$|P + Q| \leq \sup(|P|, |Q|),$$

pour tous P, Q dans $A[X]$.

Si le polynôme P est de degré n le scalaire a_n sera appelé le coefficient du terme de degré dominant. Si le coefficient du terme de degré dominant est égal à 1 le polynôme sera dit unitaire.

Définition 5

Une série formelle $\sum_{i \geq 0} a_i X^i$ est dite de *valuation* supérieure ou égale à n si ses coefficients a_i sont nuls pour $i < n$. Elle est dite de valuation n si $a_n \neq 0$.

Proposition 3. Si l'anneau A est un domaine d'intégrité, les anneaux $A[X]$ et $A[[X]]$ sont intègres.

Ceci est en particulier vrai si A est un corps k .

Démonstration. En effet considérons le cas de deux polynômes non nuls $P = \sum_0^m a_i X^i$ et $Q = \sum_0^n b_i X^i$, de degré respectif m et n . Alors le produit est de degré $m + n$, le coefficient du terme de degré dominant est $a_m b_n$. Ce coefficient est non nul car a_m et b_n le sont. Donc le produit PQ est non nul.

Considérons le cas de deux séries formelles $f = \sum_{i \geq u} a_i X^i$ et $g = \sum_{i \geq v} b_i X^i$, de valuation respectivement u et v . Alors le produit est de valuation $u + v$. Les coefficients c_j du produit fg sont nuls pour $j < uv$. En effet $c_j = \sum_{k+l=j} a_k b_l$, et si $j < u + v$ soit $k < u$ soit $l < v$, donc soit $a_k = 0$ soit $b_l = 0$. Le coefficient $a_u b_v$ est non nul car a_u et b_v le sont. ■

Division euclidienne

On va maintenant passer à l'étude de la *division euclidienne* pour un anneau de polynômes à coefficients dans un corps.

Théorème 1. Soit P un polynôme non nul à coefficients dans un corps k , et soit D un polynôme non nul à coefficients dans k . Alors il existe un polynôme Q et un polynôme R tels que :

- $P = QD + R$,
- $|R| < |D|$.

Les polynômes Q et R sont uniquement déterminés par ces conditions.

Démonstration. Montrons d'abord l'unicité, admettant que le problème a au moins une solution. Soit $P = QD + R$ et soit $P = Q'D + R'$. On a donc $D(Q - Q') = R' - R$; si $Q \neq Q'$ le degré du terme de gauche est supérieur ou égal à celui de D , donc strictement supérieur à celui de $R' - R$. C'est impossible, donc $Q = Q'$ et par voie de conséquence $R = R'$.

La démonstration de l'existence se fait par récurrence sur le degré de $|P|$. Le début de la récurrence est immédiat. Supposons la faite en degré inférieure ou égal à $n - 1$ et passons au degré n . Supposons aussi que $n \geq |D|$, sinon $Q = 0$ et $R = P$ conviennent.

Le polynôme P est donc de degré n , écrivons le sous la forme $aX^n + T$ avec $|T| \leq n - 1$.

Considérons maintenant aX^n , soit $D = b_k X^k + \dots + b_0$, où b_k est le coefficient du terme de degré dominant (et est non nul). On pose alors $P' = P - \frac{a}{b_k} X^{n-k} D$ ($n \geq k$), ce polynôme est de degré inférieur ou égal à $n - 1$. Par récurrence il existe donc U et S tels que $P' = UD + S$ avec $|S| < |D|$. On obtient alors :

$$P = P' + \frac{a}{b_k} X^{n-k} D = (U + \frac{a}{b_k} X^{n-k})D + S,$$

qui donne la décomposition souhaitée car $|S| < |D|$. ■

Le degré détermine une structure d'anneau euclidien sur l'anneau des polynômes à coefficients dans un corps. Il résulte alors de la section 2 que :

Corollaire 1. L'anneau des polynômes à coefficients dans un corps k est principal. En conséquence tout idéal non trivial (non réduit à 0) est de la forme (P) pour un polynôme non nul P bien déterminé à un facteur scalaire près. En fait ce polynôme est un polynôme non nul de plus bas degré dans l'idéal.

En conséquence on peut appliquer tous les résultats concernant les anneaux principaux. On a une notion de pgcd pour les polynômes, l'identité de Bézout a lieu.

Corollaire 2. Soient P et Q deux polynômes à coefficients dans un corps k dont l'un au moins est non nul. Soit W leur pgcd. Il existe U et V tels que $UP + VQ = W$. En particulier P et Q sont premiers entre eux si et seulement si il existe U et V tels que $UP + VQ = 1$.

La démonstration est identique à celle donnée plus haut dans le cas général, ou se transcrit à partir de celle pour $\mathbb{Z}$. Rappelons que l'on a l'algorithme d'Euclide pour calculer le pgcd et qu'on peut en déduire les coefficients de la formule de Bézout.

Le fait que l'anneau soit euclidien permet de renforcer ces énoncés, comme dans le cas de $\mathbb{Z}$. Gardons les notations du corollaire, alors :

Proposition 4. Supposons que l'un des polynômes P et Q ne soit pas constant. On peut trouver des polynômes U et V tels que $|U| < \left|\frac{Q}{W}\right|$ et $|V| < \left|\frac{P}{W}\right|$. De plus U et V sont alors uniquement déterminés.

Démonstration. Supposons d'abord que les polynômes P et Q soient premiers entre eux, donc que $W = 1$. On choisit U et V satisfaisant aux conditions du corollaire. Supposons que Q ne soit pas constant. Faisons la division de U par Q , on a $U = QH + R$. Donc on obtient en remplaçant dans l'identité de Bézout

$$RP + (V + HP)Q = 1.$$

Le degré de R est, par construction, strictement inférieur à celui de Q . Montrons que le degré de $V + HP$ est strictement inférieur à celui de P . Éliminons le cas où $V + HP = 0$ où il n'y a rien à vérifier.

Le polynôme $(V + HP)Q$ est de degré strictement positif. Son terme de plus haut degré doit donc s'annuler avec celui de RP car leur somme est égale à 1. On a donc $0 < |Q| \leq |(V + H)Q|$, et $|RP| = |R| + |P| = |(V + H)Q| = |V + H| + |Q|$. Or on sait que $|R| < |Q|$, donc $|Q| + |P| > |R| + |P| = |V + H| + |Q|$, soit $|P| > |V + H|$.

Pour ce qui est l'unicité, supposons avoir deux solutions U, V et U', V' . En soustrayant termes à termes on obtient $(U - U')P = (V' - V)Q$. Donc P divise $V' - V$, comme $|V' - V| < |P|$ ce n'est possible que si $V = V'$. De même on obtient $U = U'$.

Si le pgcd est W on remplace P par $P' = \frac{P}{W}$ et Q par $Q' = \frac{Q}{W}$. Les polynômes P' et Q' sont premiers entre eux, on peut leur appliquer le résultat. ■

Le théorème 1 peut se généraliser dans certains cas à celui des polynômes à coefficients dans un anneau.

Plus précisément supposons que A soit un domaine d'intégrité, alors :

Théorème 2. Soit P un polynôme non nul à coefficients dans un anneau intègre A , et soit D un polynôme non nul à coefficients dans A . Supposons que le coefficient du terme de degré dominant de D soit inversible dans A . Alors il existe un polynôme Q et un polynôme R à coefficients dans A tels que :

- $P = QD + R$,
- $|R| < |D|$.

Les polynômes Q et R sont uniquement déterminés par ces conditions.

Démonstration. La démonstration de l'unicité est exactement la même que la précédente dans la mesure où l'anneau de polynômes est un domaine d'intégrité. Pour ce qui est de l'existence on constate en relisant la démonstration du théorème précédent que la seule chose utilisée de la structure de corps est le fait que le coefficient du terme de degré dominant de D a un inverse. C'est exactement l'hypothèse que l'on rajoute dans le théorème. ■

Par exemple si on considère des polynômes à coefficients dans $\mathbb{Z}$ on peut toujours effectuer la division par un polynôme unitaire.

6. Polynômes en plusieurs indéterminées, polynômes symétriques

On peut itérer les constructions données dans la section précédente. Plaçons nous dans le cas des polynômes. On se donne un anneau commutatif A , l'anneau de polynômes en deux indéterminées à coefficients dans A est par définition l'anneau des polynômes à une indéterminée dans $A[X]$. On donne un autre nom à la nouvelle indéterminée, par exemple Y , ou on rebaptise les indéterminées X_1 et X_2 . L'anneau obtenu, soit $A[X][Y]$, est noté $A[X, Y]$ (ou $A[X_1, X_2]$). Les règles de calcul sont bien entendues les règles ordinaires. De façon plus générale :

Définition 1

L'anneau des polynômes à n indéterminées à coefficients dans A est défini par récurrence comme étant l'anneau des polynômes en une indéterminée X_n sur l'anneau des polynômes à $n - 1$ indéterminées, soit $A[X_1, \dots, X_{n-1}][X_n]$. Il est noté $A[X_1, \dots, X_n]$.

Un polynôme en n indéterminées s'écrit :

$$\sum_{a_1, \dots, a_n} c_{a_1, \dots, a_n} X_1^{a_1} \dots X_n^{a_n} ,$$

où $a_1, \dots, a_n \in \mathbb{N}$ et $c_{a_1, \dots, a_n} \in A$. La somme est bien entendue finie. Un terme de la forme $X_1^{a_1} \dots X_n^{a_n}$ est appelé un monôme en les indéterminées X_i . Si A est un corps k l'anneau $k[X_1, \dots, X_n]$ est un espace vectoriel sur k et les monômes en constituent une base.

Notons que si A est intègre

$$A[X_1, \dots, X_{n-1}][X_n]$$

l'est aussi à cause de la section précédente. Dans la suite on supposera A intègre.

On peut évidemment définir le degré par rapport à chacune des indéterminées. Ce n'est qu'un cas particulier de ce qui a été fait plus haut. On peut aussi définir un degré total :

Définition 2

Le polynôme $\sum_{a_1, \dots, a_n} c_{a_1, \dots, a_n} X_1^{a_1} \dots X_n^{a_n}$ est dit de degré inférieur ou égal à d si $c_{a_1, \dots, a_n} = 0$ dès que $a_1 + \dots + a_n > d$. Il est alors dit de degré total d s'il existe $a_1, \dots, a_n$ tel que $a_1 + \dots + a_n = d$ et $c_{a_1, \dots, a_n} \neq 0$.

Par exemple le polynôme $X_1^7 + X_1^3 X_2^2 + X_2^9$ est de degré total 10, de degré 8 en X_1 , 9 en X_2 .

Définition 3

Un polynôme est dit homogène de degré n s'il est de degré total n et si les coefficients des monômes de degré total strictement inférieur à n sont nuls.

Supposons que A soit un corps k et soit B une algèbre sur k . Soient $\alpha_1, \dots, \alpha_n \in B$. On a :

Proposition 1. *Il existe un homomorphisme d'anneaux et un seul*

$$f : k[X_1, \dots, X_n] \longrightarrow B$$

tel que :

- $f(X_i) = \alpha_i$,
- f restreint aux polynômes de degré zéro s'identifie à l'homomorphisme $i : k \rightarrow B$ déterminant la structure d'algèbre.

La démonstration est identique à celle du cas d'une indéterminée. Si on a :

$$P = \sum_{a_1, \dots, a_n} c_{a_1, \dots, a_n} X_1^{a_1} \dots X_n^{a_n}$$

on a :

$$f(P) = \sum_{a_1, \dots, a_n} c_{a_1, \dots, a_n} \alpha_1^{a_1} \dots \alpha_n^{a_n}.$$

Nous passons maintenant à l'étude des polynômes symétriques. Soit A un anneau, et soit $A[X_1, \dots, X_n]$ l'anneau des polynômes à coefficients dans A en n indéterminées. Définissons une action du groupe symétrique $\mathfrak{S}_n$ (ainsi qu'il a été fait au chapitre 2) par :

$$(\sigma P)(X_1, \dots, X_n) = P(X_{\sigma(1)}, \dots, X_{\sigma(n)}).$$

Ceci définit bien une action car :

$$(\sigma(\nu P))(X_1, \dots, X_n) = P(X_{\sigma\nu(1)}, \dots, X_{\sigma\nu(n)}) = ((\sigma\nu)P)(X_1, \dots, X_n).$$

On se propose de rechercher les éléments fixes pour cette action, que l'on appellera *polynômes symétriques*. Considérons le cas particulier où $n = 2$, les polynômes $X_1 + X_2$ et $X_1 X_2$ sont symétriques. En fait on montre que pour tout polynôme symétrique P il existe un et un seul polynôme Q à coefficients dans A tel que $P(X_1, X_2) = Q(X_1 + X_2, X_1 X_2)$.

Considérons maintenant les polynômes en n indéterminées $X_1, \dots, X_n$ et introduisons les polynômes symétriques élémentaires Σ_i où $1 \leq i \leq n$. Comme dans la démonstration du théorème ci-dessous on doit considérer les polynômes symétriques élémentaires en n indéterminées ainsi qu'en $n - 1$ indéterminées on précise la notation en $\Sigma_{i,n}$ et en $\Sigma_{i,n-1}$. Quand il n'y a pas d'ambiguïtés le second indice est omis.

Définition 4

Le polynôme $\Sigma_{i,n}$ défini par :

$$\Sigma_{i,n} = \sum_{(a_1, \dots, a_i) | a_1 < a_2 < \dots < a_i} X_{a_1} \cdots X_{a_i}$$

est symétrique. Plus généralement si Q est un polynôme en n indéterminées à coefficients dans A le polynôme $Q(\Sigma_1, \dots, \Sigma_n)$ est symétrique.

Par abus on a inclus dans la définition un énoncé. La démonstration procède comme suit. Appliquer une permutation revient à modifier l'ordre des termes et des facteurs dans les termes. Cela ne change rien puisque l'anneau est commutatif. Pour la seconde affirmation il suffit de vérifier que la somme et le produit de deux polynômes symétriques le sont encore. Mais si s est dans $\mathfrak{S}_n$ on a $s(P + Q) = s(P) + s(Q)$ et $s(PQ) = s(P)s(Q)$. D'où le résultat.

Par exemple on a (omettant le second indice) :

- $\Sigma_1 = X_1 + \dots + X_n$,
- $\Sigma_n = X_1 X_2 \cdots X_n$,
- si $n=3$ on a $\Sigma_2 = X_1 X_2 + X_1 X_3 + X_2 X_3$.

Observons que

$$\Sigma_{i,n}(X_1, \dots, X_{n-1}, 0) = \Sigma_{i,n-1}$$

si $i \leq n-1$ et que

$$\Sigma_{n,n}(X_1, \dots, X_{n-1}, 0) = 0.$$

Voici le théorème fondamental de cette section :

Théorème 1. Soit $P \in A[X_1, \dots, X_n]$. Supposons que P soit symétrique. Alors il existe un polynôme Q en n indéterminées à coefficients dans A et un seul tel que :

$$P(X_1, \dots, X_n) = Q(\Sigma_1, \dots, \Sigma_n).$$

Démonstration. La démonstration de l'existence se fait par une double récurrence. D'abord sur le nombre d'indéterminées, dans le cas $n=1$ il n'y a rien à démontrer, puis sur le degré total du polynôme. Nous supposons donc le théorème démontré jusqu'au degré $k-1$ pour les polynômes en n indéterminées et passons au degré k .

Soit donc P un polynôme symétrique de degré k . Considérons le polynôme en $n-1$ indéterminées $P(X_1, \dots, X_{n-1}, 0)$. Il est symétrique, pour l'action de $\mathfrak{S}_{n-1}$. Il existe donc par hypothèse de récurrence un polynôme Q en $n-1$ indéterminées tel que :

$$P(X_1, \dots, X_{n-1}, 0) = Q(\Sigma_{1,n-1}, \dots, \Sigma_{n-1,n-1}).$$

Considérons alors le polynôme en n indéterminées :

$$P' = P(X_1, \dots, X_n) - Q(\Sigma_{1,n}, \dots, \Sigma_{n-1,n}).$$

Il est symétrique en tant que polynôme en n indéterminées.

Si, à la variable X_n , on substitue zéro le polynôme obtenu est nul. Ceci implique que le coefficient d'un monôme où X_n n'apparaît pas est nul.

Soit $i < n$, et soit τ la transposition échangeant i et n . On a $\tau P' = P'$, donc

$$\tau P'(X_1, \dots, X_{n-1}, 0) = 0$$

par définition de τ , $P'(X_1, \dots, X_{i-1}, 0, X_{i+1}, \dots, X_n) = 0$. Ce qui implique que le coefficient d'un monôme où X_i n'apparaît pas est nul.

Puisque ceci est vrai pour tout i ($i = 1, \dots, n$) le coefficient d'un monôme où tous les X_i n'apparaissent pas (c'est-à-dire qu'au moins l'un d'entre eux a pour exposant zéro) est nul. Donc chaque monôme ayant un coefficient différent de zéro est divisible par $\Sigma_{n,n}$.

On peut donc poser $P' = \Sigma_{n,n} R$, le degré de R est strictement inférieur à celui de P' et donc de P . Il est symétrique ; en effet, pour une permutation quelconque s

on a

$$P' = s(P') = s(\Sigma_{n,n} R) = s(\Sigma_{n,n})s(R) = \Sigma_{n,n}s(R).$$

Soit

$$\Sigma_{n,n}(R - s(R)) = 0.$$

Si l'anneau A est intègre le résultat suit, l'anneau de polynômes en n indéterminées l'étant aussi. Le résultat est vrai sans hypothèses sur A , l'équation ci-dessus impliquant l'égalité des coefficients des polynômes R et $s(R)$.

En effet, si on a $H = \sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} X_1^{i_1} \dots X_n^{i_n}$ l'équation

$$\Sigma_{n,n} H = 0,$$

équivalait à

$$\sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} X_1^{i_1+1} \dots X_n^{i_n+1} = 0.$$

Soit $a_{i_1, \dots, i_n} = 0$ pour tout multi-indice, soit $H = 0$.

Il nous faut maintenant démontrer l'unicité de Q . On doit donc montrer qu'un polynôme H en n indéterminées, tel que $H(\Sigma_{1,n}, \dots, \Sigma_{n,n}) = 0$, est nul. Faisons une récurrence sur le nombre d'indéterminées. Pour $n = 1$ la propriété est vraie. Admettons le résultat pour $n - 1$ indéterminées. Puis faisons une récurrence sur le degré total.

À l'indéterminée X_n substituons 0. Le polynôme $H(\Sigma_{1,n-1}, \dots, \Sigma_{n,n-1}, 0)$, est nul. Donc par récurrence le polynôme en $n - 1$ indéterminées

$$H(X_1, \dots, X_{i-1}, \dots, X_{n-1}, 0)$$

est nul, autrement dit, $H(\Sigma_{1,n}, \dots, \Sigma_{n,n}) = \Sigma_{n,n} H'(\Sigma_{1,n}, \dots, \Sigma_{n,n})$.

Comme le degré total de H' est strictement inférieur à celui de H on peut conclure par une récurrence sur le degré total. ■

Formules de Newton

On va maintenant démontrer des relations qui expriment certains polynômes symétriques à l'aide des polynômes symétriques élémentaires. Il s'agit des polynômes :

$$N_k = \sum_i X_i^k.$$

Le plan de la démonstration est donnée, les détails sont laissés en exercice au lecteur.

Théorème 2. On a les formules suivantes :

► si $0 \leq i \leq n$

$$N_i = N_{i-1}\Sigma_{1,n} - N_{i-2}\Sigma_{2,n} + \cdots + (-1)^{i-1}N_1\Sigma_{i-1,n} + (-1)^i\Sigma_i,$$

► si $i > n$,

$$N_i = N_{i-1}\Sigma_{1,n} - N_{i-2}\Sigma_{2,n} + \cdots + (-1)^{n-2}N_{i+1-n}\Sigma_{n-1,n} + (-1)^{n-1}N_{i-n}\Sigma_{n,n},$$

Démonstration. Dans la démonstration on utilisera la dérivation formelle des polynômes en plusieurs variables qui est définie dans la section 8. Le lecteur est invité à s'y reporter. En fait, il suffit de savoir que les propriétés algébriques standards sont vérifiées.

Lemme 1. Soit A un anneau et considérons l'anneau de polynômes en une indéterminée T sur $A[X_1, \dots, X_n]$. Soient $\Sigma_{i,n}$ les polynômes symétriques élémentaires par rapport aux indéterminées X_i . On a :

$$f(T) = \prod_{i=1, \dots, n} (T - X_i) = \sum_{j=0, \dots, n} (-1)^j \Sigma_{j,n} T^{n-j}.$$

Démonstration. Par convention $\Sigma_{0,n} = 1$. La démonstration consiste à vérifier la définition des polynômes symétriques élémentaires. $\square$

On dérive alors cette relation par rapport à la variable T et on obtient la formule suivante :

$$\sum_i \prod_{j \neq i} (T - X_j) = \sum_{j=0, \dots, n} (-1)^j (n-j) \Sigma_{j,n} T^{n-j-1}.$$

On a $(T - X_i) \times \prod_{j \neq i} (T - X_j) = f(T)$. Faisons la division euclidienne de f par $T - X_i$ on obtient la formule :

$$\prod_{j \neq i} (T - X_j) = \sum_{j=0, \dots, n-1} \left(\sum_{k=0, \dots, j} (-1)^k \Sigma_{k,n} X_i^{j-k} \right) T^{n-j-1}.$$

Puis on somme sur l'indice i , on obtient :

$$\sum_{i=1, \dots, n} \prod_{j \neq i} (T - X_j) = \sum_{j=0, \dots, n-1} \left(\sum_{k=0, \dots, j} (-1)^k \Sigma_{k,n} N_{j-k} \right) T^{n-j-1}.$$

En comparant à la formule obtenue par dérivation, on identifie les coefficients de T^j , $0 \leq j \leq n-1$ de chaque côté, et on obtient la première partie du théorème.

Pour ce qui est de la seconde, on fait simplement $T = X_i$ dans l'équation du lemme 1. On multiplie l'équation obtenue par X_i^k et on somme sur i de 1 à n . Les termes se regroupent et font apparaître les N_i pour $k \leq i \leq k+n$. $\blacksquare$

7. Anneaux factoriels et applications

Nous passons maintenant à l'étude des anneaux factoriels. On suppose donné un anneau A commutatif, et soit $x_i, i \in I$ un système de représentants des éléments irréductibles (définition 2 et suite, section 2).

Rappelons encore une fois ce que l'on entend par là : d'abord deux éléments irréductibles x et y sont dans la même classe d'équivalence si et seulement il existe un élément inversible u tel que $x = yu$. Alors un système de représentants est un ensemble d'éléments irréductibles x_i , avec $i \in I$, auquel appartient un élément de chaque classe et un seul. Il n'y a pas bien évidemment un seul système de représentants. On passe d'un système de représentant à un autre en multipliant chaque élément par un élément inversible approprié.

Définition 1

On dit que l'anneau A est un *anneau factoriel* si :

- il est intègre,
- tout élément non nul $x \in A$ peut s'écrire sous la forme :

$$x = u \times \prod_i x_i^{\alpha_i}, \quad \alpha_i \in \mathbb{N},$$

où $u \in A^*$ et les entiers α_i sont presque tous nuls,

- cette décomposition est unique.

L'unicité de la décomposition signifie la même chose que dans le cas des anneaux principaux. Si on a :

$$u \times \prod_i x_i^{\alpha_i} = v \times \prod_i x_i^{\beta_i},$$

où seuls un nombre fini d'exposants sont non nuls, alors $u = v$ et $\alpha_i = \beta_i$ pour tout $i \in I$.

Ainsi qu'on l'a vu plus haut, les anneaux principaux sont factoriels. La réciproque n'est pas vraie.

Dans un anneau factoriel on a, comme d'habitude, la notion de divisibilité d'un élément par un autre. On rappelle donc que a divise b s'il existe $c \in A$ tel que $b = ac$ et qu'on le note $a \mid b$.

On peut déterminer, à partir des décompositions en produit d'éléments irréductibles, quand un élément non nul x divise un élément non nul y . Soit $x = u \times \prod_i x_i^{\alpha_i}$ et $y = v \times \prod_i x_i^{\beta_i}$, $u, v \in A^*$, alors :

Proposition 1. *L'élément x divise l'élément y si et seulement si pour tout $i \in I$ on a $\alpha_i \leq \beta_i$.*

Démonstration. Commençons par rappeler que dans les formules précédentes seuls un nombre fini d'exposants sont non nuls. Si on a $x = u \times \prod_i x_i^{\alpha_i}$ et $y = v \times \prod_i x_i^{\beta_i}$ et si $y = xz$, alors nécessairement on a $z = vu^{-1} \times \prod_i x_i^{\beta_i - \alpha_i}$, et les exposants $\beta_i - \alpha_i$ sont positifs ou nuls.

Inversement si on a $\beta_i \geq \alpha_i$ posons $z = vu^{-1} \times \prod_i x_i^{\beta_i - \alpha_i}$, on a bien $y = xz$. ■

On définit le pgcd et le ppcm de deux éléments non nuls x et y par les mêmes formules que dans un anneau principal. Soient $x = u \times \prod_{i \in I} x_i^{\alpha_i}$ et $y = v \times \prod_{i \in I} x_i^{\beta_i}$ les décompositions de x et y . Alors, par définition :

$$\text{pgcd}(x, y) = \prod_{i \in I} x_i^{\inf(\alpha_i, \beta_i)},$$

par définition le pgcd d'un élément non nul x et de zéro est x .

$$\text{ppcm}(x, y) = \prod_{i \in I} x_i^{\sup(\alpha_i, \beta_i)}.$$

En particulier, si le pgcd de deux éléments est 1 on dit qu'ils sont premiers entre eux. Les formules ci-dessus s'étendent au pgcd et au ppcm de n éléments non nuls. Dans le cas du pgcd on peut étendre à une famille d'éléments quelconques, par définition leur pgcd est celui des éléments non nuls de la famille. On a pour une famille d'éléments non nuls :

$$\text{pgcd}(a_1, \dots, a_n) = \prod_{i \in I} x_i^{\inf(\alpha_{i,j})},$$

où

$$a_j = u_j \times \prod_{i \in I} x_i^{\alpha_{i,j}}.$$

On remarquera que le pgcd et le ppcm sont déterminés par le choix du système d'éléments irréductibles. En modifiant ce système on modifie pgcd et ppcm par un élément inversible de l'anneau. Si bien que la modification n'est pas significative.

La différence avec les anneaux principaux est que l'on n'a pas l'identité de Bézout. Par contre on a le lemme de Gauss :

Lemme 1. Soit A un anneau factoriel et soient $a, b, c \in A$. Supposons que $a \mid bc$ et que a soit premier à b alors a divise c .

Démonstration. Elle résulte de la proposition 1 en écrivant les décompositions en facteurs irréductibles des éléments considérés. ■

Le principal théorème de cette section est le suivant :

Théorème 1. Soit A un anneau factoriel, alors l'anneau $A[X]$ est factoriel.

Le corollaire suivant est immédiat :

Corollaire (Gauss). Soit A un anneau factoriel, par exemple un corps ou un anneau principal. L'anneau des polynômes en n indéterminées $A[X_1, \dots, X_n]$ est factoriel.

Introduisons une définition :

Définition 2

Soit $P = \sum_{i=0, \dots, n} a_i X^i \in A[X]$, le contenu de P , noté $\text{cont}(P)$, est par définition $\text{pgcd}(a_0, \dots, a_n)$.

Venons en à la démonstration du théorème, précisons l'énoncé par un lemme :

Lemme 2. Les éléments irréductibles de $A[X]$ sont :

- les éléments irréductibles de A ,
- les polynômes à coefficients dans A qui sont irréductibles comme polynômes à coefficients dans le corps des fractions K de A , et qui sont de contenu égal à 1.

Lemme 3 (Gauss). Soient $P, Q \in A[X]$, on a

$$\text{cont}(PQ) = \text{cont}(P) \text{cont}(Q).$$

Démonstration. Commençons par supposer que $\text{cont}(P) = \text{cont}(Q) = 1$ et montrons que $\text{cont}(PQ)$ est inversible. Supposons qu'un élément irréductible p divise $\text{cont}(PQ)$. Posons $P = a_0 + a_1 X + \dots + a_m X^m$ et $Q = b_0 + b_1 X + \dots + b_n X^n$. Comme le contenu de P est 1 on peut supposer que p divise $a_0, a_1, \dots, a_k$ mais pas a_{k+1} (qui est donc différent de zéro). De même, comme le contenu de Q est 1, on peut supposer que p divise $b_0, a_1, \dots, b_\ell$ mais pas $b_{\ell+1}$ (qui est donc différent de zéro). On sait que p divise tous les coefficients du polynôme PQ . Mais ce n'est pas possible, supposons en effet qu'il divise le coefficient du terme de degré $k + \ell + 2$ qui est

$$a_0 b_{k+\ell+2} + \dots + a_k b_{\ell+2} + a_{k+1} b_{\ell+1} + a_{k+2} b_\ell + \dots + a_{k+\ell+2} b_0.$$

Tous les termes, à l'exception du terme central, sont divisibles par p à cause des hypothèses. Le terme central $a_{k+1} b_{\ell+1}$ devrait donc l'être aussi, or c'est contraire aux hypothèses et au lemme d'Euclide. Le contenu de PQ n'étant divisible par aucun élément irréductible est nécessairement inversible (rappelons qu'un élément inversible n'est pas irréductible). La définition du pgcd force alors le contenu à être égal à 1.

Faisons la démonstration dans le cas général. Écrivons P et Q sous la forme $\text{cont}(P)P'$ et $\text{cont}(Q)Q'$. Les polynômes P' et Q' sont à coefficients dans A car le contenu de P (resp. de Q) divise tous les coefficients de P (resp. de Q). Les polynômes P' et Q' sont par construction de contenu 1.

On obtient donc $PQ = \text{cont}(P)\text{cont}(Q)P'Q'$. Le résultat suit en calculant le contenu du membre de droite, c'est $\text{cont}(P)\text{cont}(Q)$ car on vient de montrer que le pgcd des coefficients de $P'Q'$ est 1, et le pgcd des coefficients de $\text{cont}(P)\text{cont}(Q)P'Q'$ est donc $\text{cont}(P)\text{cont}(Q)$. ■

Démonstration du lemme 2. Un élément irréductible dans l'anneau A l'est encore dans $A[X]$ car A est intègre : le produit de deux polynômes dont l'un, au moins, est de degré différent de zéro est un polynôme de degré différent de zéro et ne peut donc s'identifier à un élément de A .

Considérons maintenant un polynôme de degré différent de zéro et supposons qu'il soit irréductible. Son contenu doit être égal à 1. En effet on peut l'écrire $\text{cont}(P)P'$ et il est réductible si son contenu n'est pas égal à 1. Supposons qu'il soit réductible en tant que polynôme à coefficients dans le corps des fractions K de A . Soit $P = UV$, où *a priori* U et V sont à coefficients dans le corps des fractions K . Multiplions par le ppcm r des dénominateurs des coefficients de U et par le ppcm s des dénominateurs des coefficients de V . On obtient

$$rsP = (rU)(sV) = \text{cont}(rU)\text{cont}(sV) \frac{rU}{\text{cont}(rU)} \frac{sV}{\text{cont}(sV)}.$$

Les polynômes rU et sV sont à coefficients dans A . Les polynômes

$$\frac{rU}{\text{cont}(rU)}$$

et

$$\frac{sV}{\text{cont}(sV)}$$

sont à coefficients dans A et de contenu 1 par construction. Calculons le contenu des deux côtés, il est $\text{cont}(rU)\text{cont}(sV)$ à droite et rs à gauche, ces deux quantités sont égales, on peut diviser à gauche et droite et obtenir :

$$P = \frac{rU}{\text{cont}(rU)} \frac{sV}{\text{cont}(sV)}.$$

Donc P est réductible comme polynôme à coefficients dans A , ce qui est contraire à l'hypothèse.

Inversement, un polynôme P à coefficients dans A de contenu 1 et irréductible comme polynôme à coefficients dans K l'est comme polynôme à coefficients dans A . En effet, la seule possibilité serait qu'il s'écrive UV avec, par exemple, U de degré zéro donc constant, donc égal à un élément de A non-inversible. Mais cela serait en contradiction avec le fait qu'il est de contenu 1 si cet élément n'est pas inversible. ■

Démonstration du théorème 1. Choisissons donc un système de représentants d'éléments irréductibles de $A[X]$. Il est constitué par un système d'éléments

irréductibles de A , soit a_i , $i \in I$ et une famille de polynômes P_j , $j \in J$ satisfaisant aux propriétés du lemme 2. Les polynômes P_j constituent un système de représentants des éléments irréductibles de $K[X]$.

Soit donc $P \in A[X]$ et soit

$$a \times \prod_{j \in J} P_j^{\alpha_j}$$

sa décomposition en éléments irréductibles dans $K[X]$, où $a \in K$. Écrivons a sous la forme $\frac{p}{q}$ avec $p, q \in A$. On a donc

$$qP = p \prod_{j \in J} P_j^{\alpha_j},$$

ces deux polynômes sont à coefficients dans A . En calculant le contenu des deux côtés on obtient $q \text{ cont}(P) = pu$, avec $u \in A^*$. Donc l'élément $a = \frac{pu}{q}$ de K est, en fait, élément de A . Si $a = v \times \prod_i a_i^{\beta_i}$, $v \in A^*$, est la décomposition de a on a

$$P = v \times \prod_i a_i^{\beta_i} \times \prod_{j \in J} P_j^{\alpha_j}.$$

Il reste à démontrer l'unicité de la décomposition. Mais si on a deux écritures $P = v \times \prod_i a_i^{\beta_i} \times \prod_{j \in J} P_j^{\alpha_j}$, et $P = v' \times \prod_i a_i^{\beta'_i} \times \prod_{j \in J} P_j^{\alpha'_j}$ en écrivant l'égalité des contenus on obtient $\prod_i a_i^{\beta_i} = \prod_i a_i^{\beta'_i}$. Soit $\beta_i = \beta'_i$ pour tout indice i . Il reste $P = v \times \prod_{j \in J} P_j^{\alpha_j} = v' \times \prod_{j \in J} P_j^{\alpha'_j}$. Considérant cette identité dans $K[X]$ on en déduit que $\alpha_j = \alpha'_j$ pour tout indice j . D'où le résultat. ■

Pour terminer cette section nous donnons un critère qui permet de déterminer si certains polynômes sont irréductibles. Il s'agit du *critère d'Eisenstein*.

Proposition 2 (critère d'Eisenstein). Soit A un anneau factoriel et soit p un élément irréductible de A . Soit $P = a_0 + a_1X + \dots + a_nX^n$, $a_n \neq 0$, un polynôme à coefficients dans A . Supposons que $\text{cont}(P) = 1$ et que p divise $a_0, \dots, a_{n-1}$ mais que p ne divise pas a_n et que p^2 ne divise pas a_0 . Alors le polynôme P est irréductible dans $A[X]$.

Démonstration. On raisonne par l'absurde et on écrit $P = QR$, $Q = b_0 + b_1X + \dots + b_pX^p$ et $R = c_0 + c_1X + \dots + c_qX^q$, avec b_p et c_q différents de zéro. On a $a_0 = b_0c_0$, par hypothèse p divise a_0 et p^2 ne divise pas a_0 donc p divise b_0 ou c_0 mais ne divise pas les deux. Supposons qu'il divise b_0 . Comme p divise aussi $a_1 = b_0c_1 + b_1c_0$ il divise b_1c_0 , et comme il ne divise pas c_0 , il divise b_1 . Supposons par récurrence qu'il divise $b_0, \dots, b_{k-1}$, il divise $a_k = b_kc_0 + \dots + b_0c_k$. Donc il divise b_kc_0 et donc b_k . On vient de démontrer par récurrence que p divise tous

les coefficients du polynôme Q . L'élément p divise donc le coefficient $a_n = b_p c_q$, ce qui est contraire à l'hypothèse. ■

Par exemple si p est un nombre premier le polynôme $X^n - p$ est irréductible dans $\mathbb{Z}[X]$.

Voici une application classique :

Proposition 3. Soit p un nombre premier. Le polynôme $1 + X + \dots + X^{p-1}$ est irréductible dans $\mathbb{Z}[X]$.

Démonstration. On applique le critère d'Eisenstein en faisant le changement de variables $X = Y + 1$. On a

$$1 + X + \dots + X^{p-1} = \frac{X^p - 1}{X - 1} = \frac{(Y + 1)^p - 1}{Y}$$

et

$$\frac{(Y + 1)^p - 1}{Y} = \sum_{i=0, \dots, p-1} C_p^i Y^{i-1}.$$

Le coefficient binomial C_p^i est égal à $\frac{p!}{i!(p-i)!}$ et est donc divisible par p si $0 < i < p$, par ailleurs $C_p^p = 1$ et $C_p^1 = p$ n'est pas divisible par p^2 . Le résultat suit. ■

8. Zéros des polynômes

Dans cette section on va étudier les zéros des polynômes et exposer la théorie du résultant.

Définition 1

Soit P un polynôme à coefficients dans un corps k . On dit que $\alpha \in k$ est zéro ou racine de P si $P(\alpha) = 0$.

Théorème 1. L'élément $\alpha \in k$ est zéro du polynôme P si et seulement si le polynôme $X - \alpha$ divise le polynôme P .

Démonstration. Soit P un polynôme quelconque. faisons la division de P par $X - \alpha$, on obtient $P = (X - \alpha)Q + c$, où c est un élément de k . Si dans cette formule, à l'indéterminée X , on substitue α on obtient $P(\alpha) = c$. Or α est zéro de P si et seulement si $P(\alpha) = 0$, soit si et seulement si $c = 0$, soit si et seulement si P est divisible par $X - \alpha$. ■

Définition 2

On dit que l'élément $\alpha \in k$ est zéro d'ordre ℓ de P , ou a pour multiplicité ℓ , $\ell > 0$, si $(X - \alpha)^\ell$ divise P mais $(X - \alpha)^{\ell+1}$ ne divise pas P .

Le résultat suivant est fondamental :

Théorème 2. *Un polynôme P de degré n a au plus n zéros, en tenant compte de leur multiplicité.*

Démonstration. Soit $P \in k[X]$. Supposons qu'il ait les éléments $a_i \in k$, $i = 1, \dots, h$, pour zéros et qu'ils aient pour ordre respectif α_i . Alors P est divisible par $(X - a_i)^{\alpha_i}$ pour tout i . Les zéros a_i sont distincts entre eux deux à deux, donc les polynômes $(X - a_i)^{\alpha_i}$ sont premiers entre eux deux à deux. Le polynôme P est donc divisible par leur produit. Celui-ci est de degré $\sum_i \alpha_i$, qui est le nombre de zéros, étant tenu compte de l'ordre. Ce degré est inférieur ou égal à celui de P , d'où le résultat. ■

Corollaire. *Un polynôme P de degré inférieur ou égal à n qui prend la même valeur c en $n + 1$ points distincts est constant.*

Démonstration. Il suffit pour le démontrer de considérer les zéros du polynôme $P - c$ et d'appliquer le théorème. ■

Supposons maintenant qu'un polynôme P unitaire de degré n ait n racines, compte tenu des multiplicités, dans le corps k . On dira qu'un tel polynôme est scindé dans k .

On a donc :

$$P = \prod_{i=1, \dots, n} (X - \alpha_i).$$

D'un autre côté on a :

$$P = X^n + a_{n-1}X^{n-1} + \dots + a_0.$$

On a :

Proposition 1. *On a la relation suivante entre les racines et les coefficients du polynôme :*

$$a_{n-i} = (-1)^i \Sigma_{i,n}(\alpha_1, \dots, \alpha_n).$$

Autrement dit les coefficients sont au signe près les fonctions symétriques des racines.

Démonstration. Il suffit d'effectuer le développement de la première expression et de comparer à la définition des fonctions symétriques.

On fera d'abord les cas de degré 2 et 3 comme exemples. ■

Dérivation formelle

Introduisons maintenant la *dérivation formelle* des polynômes et des séries formelles.

Définition 3

Soit $P = \sum_{i=0, \dots, n} a_i X^i \in k[X]$. On appelle *dérivée formelle* de $P = \sum_{i=0, \dots, n} a_i X^i$ le polynôme

$$P' = \sum_{i=1, \dots, n} i a_i X^{i-1}.$$

La dérivée d'une série formelle est donnée par la même formule. Si on a

$$f = \sum_{i \geq 0} a_i X^i,$$

alors sa dérivée formelle est la série f' donnée par

$$f' = \sum_{i \geq 1} i a_i X^{i-1}.$$

On définit la dérivée n -ième $P^{(n)}$ de P , respectivement $f^{(n)}$ de f , par la formule de récurrence $P^{(n)} = (P^{(n-1)})'$, respectivement $f^{(n)} = (f^{(n-1)})'$.

Les propriétés standards concernant la dérivée d'une somme et d'un produit ont lieu :

- ▀ $(P + Q)' = P' + Q'$ et $(f + g)' = f' + g'$,
- ▀ $(PQ)' = P'Q + PQ'$ et $(fg)' = f'g + fg'$,
- ▀ de même on a la formule de Leibniz :

$$(PQ)^{(n)} = \sum_{i=0, \dots, n} C_n^i P^i Q^{n-i},$$

$$(fg)^{(n)} = \sum_{i=0, \dots, n} C_n^i f^i g^{n-i}.$$

Pour démontrer la formule du produit on observe qu'il suffit de la démontrer pour $P = ax^m$ et $Q = bX^j$. Ceci résulte de la formule suivante :

$$(P(Q + R))' = (PQ)' + (PR)',$$

et du développement standard.

Pour ce qui est du cas considéré on a

$$(abX^{m+n})' = (m+n)abX^{m+n-1},$$

si $n+m \geq 1$, sinon la dérivée est nulle. Par ailleurs on a $(aX^m)'bX^n + aX^n(bx^m)' = mabX^{m+n-1} + nabX^{m+n-1}$, le résultat suit. La formule est vraie pour les séries formelles et est laissée en exercice au lecteur.

La *formule de Leibniz* se démontre par récurrence sur n , comme dans le cas ordinaire.

La dérivation est utile quand on étudie la décomposition en facteurs irréductibles d'un polynôme. La question étant de donner un critère permettant de déterminer quand un facteur irréductible H divise un polynôme P si le carré H^2 divise également P .

Théorème 3. *Soit P un polynôme à coefficients dans un corps k , supposons que P' soit non nul. Soit H un polynôme irréductible à coefficients dans k tel que H' soit non nul. Alors le polynôme H^2 divise P si et seulement si H divise le polynôme dérivé P' .*

Démonstration. Supposons que H^2 divise P , alors $P = H^2Q$, donc

$$P' = 2HH'Q + H^2Q',$$

donc H divise P' (non nul par hypothèse). Inversement montrons que si H^2 ne divise pas P alors H ne divise pas P' . Si on a $P = HQ$ avec H ne divisant pas Q on a $P' = H'Q + HQ'$. Si le polynôme H divise P' , il divise donc $H'Q$. Mais comme H est irréductible et ne divise pas H' (non nul) qui est de degré strictement inférieur, il divise Q ce qui est impossible. Le résultat suit. ■

On a le corollaire suivant :

Corollaire. *Soit P un polynôme à coefficients dans k , P' non nul, et soit α un zéro de P . Alors α est de multiplicité au moins 2 si et seulement si α est zéro de P' .*

En caractéristique zéro on peut préciser ces résultats :

Théorème 4. *Soit P un polynôme à coefficients dans un corps k de caractéristique zéro, et soit H un polynôme irréductible à coefficients dans k . Alors le polynôme H^ℓ divise P si et seulement si H divise les polynômes dérivés $P', P^{(2)}, \dots, P^{(n-1)}$.*

La démonstration est laissée au lecteur qui aura soin de remarquer que l'hypothèse de caractéristique zéro est nécessaire. On donnera aussi des contre-exemples. Nous reviendrons dans le prochain chapitre sur les propriétés de la dérivation en caractéristique p , on donnera des contre-exemples.

On a évidemment le corollaire suivant :

Corollaire. *Soit P un polynôme à coefficients dans k de caractéristique nulle, et soit α un zéro de P . Alors α est de multiplicité ℓ si et seulement si α est zéro de $P', P^{(2)}, \dots, P^{(\ell-1)}$.*

Résultant

Nous allons passer maintenant à la théorie du *résultant*.

La question est de savoir quand deux polynômes non nuls, à coefficients dans un corps k , P et Q , ont un facteur irréductible en commun. Soient donc $P = \sum_{i=0, \dots, m} a_i X^i$ de degré m et $Q = \sum_{j=0, \dots, n} b_j X^j$, de degré n . On veut donner

une condition qui s'exprime à l'aide des coefficients a_i et b_j . On suppose $a_m \neq 0$ et $b_n \neq 0$.

Supposons que les polynômes aient le facteur irréductible H en commun. On a alors $P = HR$ et $Q = HS$ donc on a la relation $PS - QR = 0$ avec $|R| < |P| = m$ et $|S| < |Q| = n$. Inversement supposons qu'il existe des polynômes U et V non nuls tels que $PV - QU = 0$ avec $|U| < |P|$ et $|V| < |Q|$. Montrons qu'alors P et Q ont un facteur irréductible en commun. Raisonnons par l'absurde et supposons qu'ils n'en aient pas, l'équation $PV = QU$ montre alors que P divise U , ce qui est impossible, à cause de la condition sur les degrés.

On obtient donc la condition nécessaire et suffisante suivante pour que les polynômes P et Q à coefficients dans K aient un facteur irréductible en commun :

Lemme. *Les polynômes non nuls P et Q ont un facteur irréductible en commun si et seulement si il existe des polynômes U et V non nuls tels que $PV - QU = 0$ avec $|U| < |P|$ et $|V| < |Q|$.*

Démonstration. Posons $U = \sum_{i=0, \dots, m-1} u_i X^i$ et $V = \sum_{j=0, \dots, n-1} v_j X^j$. Nous devons résoudre l'équation suivante :

$$\left(\sum_{i=0, \dots, m} a_i X^i \right) \left(\sum_{j=0, \dots, n-1} v_j X^j \right) - \left(\sum_{j=0, \dots, n} b_j X^j \right) \left(\sum_{i=0, \dots, m-1} u_i X^i \right) = 0,$$

où les a_i et les b_j sont donnés, et les u_i et les v_j sont les inconnues. Pour ce faire on écrit que tous les coefficients du polynôme obtenu sont nuls. Ce polynôme est de degré $m+n-1$, il y a donc $m+n$ équations linéaires en les $m+n$ inconnues u_i et v_j . pour qu'il y ait une solution non nulle il faut et suffit que le déterminant de ce système soit nul.

Le coefficient du terme de degré k , $0 \leq k \leq m+n-1$, est égal à :

$$\sum_{i+j=k, 0 \leq i \leq m, 0 \leq j \leq n-1} a_i v_j + \sum_{i+j=k, 0 \leq i \leq n, 0 \leq j \leq m-1} b_i u_j.$$

On écrit donc la nullité de ces coefficients pour $0 \leq k \leq m+n-1$. Pour obtenir une solution non nulle il faut et suffit que le déterminant du système soit nul. ■

Définition 4

On appelle ce déterminant le résultant des polynômes P et Q , on le note $\text{Res}(P, Q)$, c'est un déterminant $(m+n, m+n)$.

Avec $m \geq n$, il est égal à :

$$\begin{vmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ & a_1 & \ddots & 0 & b_1 & \ddots & \ddots & 0 \\ & & \ddots & a_0 & b_{n-1} & \ddots & \ddots & 0 \\ & & & a_1 & b_n & \ddots & \ddots & 0 \\ & & & & 0 & \ddots & \ddots & 0 \\ & & & & & \ddots & \ddots & b_0 \\ & & & & & & \ddots & b_1 \\ & & & & & & & \ddots \\ & & & & & & & b_{n-1} \\ & & & & & & & & b_n \\ 0 & \cdots & 0 & a_m & 0 & \cdots & 0 & 0 \end{vmatrix}.$$

On a donc :

Théorème 5. *Le résultant de deux polynômes P et Q à coefficients dans un corps K est nul si et seulement si P et Q ont au moins un facteur irréductible en commun.*

En particulier si P et Q ont toutes leurs racines (i.e. sont scindés) dans k et si leur résultant est nul ils ont une racine commune (au moins).

Un cas particulièrement important est celui où on prend pour couple P et Q le polynôme P et son polynôme dérivé P' . Le résultant de P et P' est appelé le *discriminant* de P . On le note $\text{Dis}(P)$.

On a :

Théorème 6. *Le discriminant d'un polynôme P à coefficients dans un corps k de caractéristique zéro est nul si et seulement si il a au moins un facteur irréductible commun avec son polynôme dérivé P' .*

On peut se contenter de supposer $P' \neq 0$ et d'autoriser toute caractéristique.

En particulier si P a toutes ses racines dans k et si son discriminant est nul il a une racine commune (au moins) avec son polynôme dérivé.

Voici des exemples.

Calculons le discriminant de $X^2 + bX + c$, c'est :

$$\begin{vmatrix} c & b & 0 \\ b & 2 & b \\ 1 & 0 & 2 \end{vmatrix}.$$

Soit tous calculs faits $4c - b^2$.

Si on considère le polynôme $X^3 + pX + q$ on obtient :

$$\begin{vmatrix} q & 0 & p & 0 & 0 \\ p & q & 0 & p & 0 \\ 0 & p & 3 & 0 & p \\ 1 & 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 & 3 \end{vmatrix}.$$

Ce qui tous calculs faits donne $4p^3 + 27q^2$.

9. Compléments sur les séries formelles

Dans ce paragraphe, nous allons donner quelques compléments sur les séries formelles et indiquer brièvement quelques possibilités d'applications. Nous nous contenterons de travailler avec des séries formelles à coefficients dans $\mathbb{Z}$.

Nous commençons par l'étude du groupe des éléments inversibles de l'anneau $\mathbb{Z}[[X]]$.

Théorème 1. *Un élément de $\mathbb{Z}[[X]]$ est inversible si et seulement si son coefficient constant a_0 est égal à 1 ou à -1 .*

Démonstration. Supposons donnée $f = \sum_{i \geq 0} a_i X^i$, et soit $g = \sum_{i \geq 0} b_i X^i$. Supposons avoir par hypothèse $fg = 1$, on a alors $a_0 b_0 = 1$. Cette équation ayant lieu dans $\mathbb{Z}$ on a $a_0 = 1$ ou $a_0 = -1$.

Inversement supposons par exemple que $a_0 = 1$ et cherchons à déterminer les coefficients b_j . L'équation $fg = 1$ est équivalente aux équations suivantes :

- $b_0 = 1$,
- $a_n b_0 + \dots + a_1 b_{n-1} + b_n = 1$.

Si on suppose par récurrence avoir déterminé $b_0, \dots, b_{n-1}$ l'équation ci-dessus permet de calculer b_n . Le résultat suit car nous avons déterminé une série formelle satisfaisant aux conditions requises. ■

Le lecteur vérifiera à titre d'exercice que ce théorème reste vrai si on remplace $\mathbb{Z}$ par un anneau quelconque et la condition $a_0 = 1$ ou $a_0 = -1$ par $a_0 \in A^*$. En particulier toute série formelle à coefficients dans un corps dont le terme constant est non nul est inversible.

Voici quelques exemples :

- l'inverse de la série $1 - X$ est $1 + X + \dots + X^i + \dots$, on la note $\frac{1}{1 - X}$,
- l'inverse de la série $1 - X^n$ est $1 + X^n + X^{2n} + \dots + X^{in} + \dots$, on la note $\frac{1}{1 - X^n}$,

- si on se place dans les séries formelles à coefficients dans $\mathbb{Q}$ l'inverse de la série formelle $\sum_{i \geq 0} \frac{X^i}{i!}$ est $\sum_{i \geq 0} (-1)^i \frac{X^i}{i!}$. Il s'agit là bien évidemment des séries exponentielles e^X et e^{-X} .

Nous allons maintenant étudier la composition des séries formelles. Commençons par l'observation suivante, soient $f(X)$ et $g(X)$ deux séries formelles dont le coefficient constant est nul, égales respectivement à $\sum_{i \geq 1} a_i X^i$ et à $\sum_{i \geq 1} b_i X^i$. Comme dans le cas des séries entières on peut former la série formelle composée $f(g(X))$. Elle est en effet égale à :

$$\sum_{i \geq 1} a_i \left(\sum_{j \geq 1} b_j X^j \right)^i.$$

Cette formule a un sens pour les raisons suivantes. D'abord chaque terme dans la somme est une puissance dans l'anneau des séries formelles, donc est bien défini. Ensuite, si on considère la suite infinie des termes de cette somme, on constate que seul un nombre fini d'entre eux ont un coefficient éventuellement différent de zéro en X^n . En effet les termes à partir du $(n+1)$ -ième sont divisibles par X^{n+1} à cause de l'hypothèse de nullité du terme constant.

Plus précisément notons α_n le coefficient de X^n dans $(\sum_{j \geq 1} b_j X^j)^n$. Le coefficient γ_n de la série $f(g(X))$ est égal à :

$$\gamma_n = \sum_{i=1, \dots, n} a_i \alpha_i.$$

Il est clair que la composée de deux séries formelles débutant par X (i.e. $a_1=b_1=1$) débute aussi par X .

Théorème 2. Les séries formelles à coefficients dans $\mathbb{Z}$ de la forme $X + \sum_{i \geq 2} a_i X^i$ forment un groupe pour la composition.

Démonstration. Ainsi qu'on l'a dit plus haut la composition est une loi interne. La série formelle X est évidemment élément neutre.

L'associativité est laissée au lecteur en exercice. On va démontrer l'existence d'un inverse. Soit donc $f = X + \sum_{i \geq 2} a_i X^i$. On cherche $g = X + \sum_{i \geq 2} b_i X^i$ telle que $f(g(X)) = g(f(X)) = X$. Notons γ_n le coefficient de la série composée $f(g(X))$. D'après la formule donnée plus haut, c'est un polynôme en les inconnues b_j et les a_i , qui eux sont connus. Plus précisément, on a :

- $\gamma_1 = 1$,
- si $n \geq 2$, $\gamma_n = b_n + P(a_2, \dots, a_n; b_2, \dots, b_{n-1}) = 0$, où P désigne un certain polynôme à coefficients entiers.

Gardons les notations introduites plus haut, et supposons avoir calculé par récurrence les inconnues $b_2, \dots, b_{n-1}$ l'équation ci-dessus permet de calculer b_n . Observons que la récurrence démarre par l'équation $b_2 + a_2 = 0$.

On trouve donc ainsi un inverse à droite. Par le même type d'argument on trouverait un inverse à gauche. Ils coïncident nécessairement. ■

On appellera cet inverse *la série réciproque*.

Notons que le groupe que l'on vient de considérer n'est pas commutatif, contrairement à celui des éléments inversibles, car la composition ne l'est pas. C'est ce qui oblige à considérer dans la démonstration des inverses à gauche et à droite.

Comme on l'a indiqué plus haut ce théorème se généralise sans peine à un anneau quelconque et en particulier à un corps. Par exemple la série réciproque de la série $e^X - 1$ est donnée par $\log(1 + X)$.

Il existe une formule générale, dite de Lagrange [Co], qui permet de calculer les coefficients de la série réciproque à partir de ceux de la série originelle.

Exercices

Tous les anneaux seront, sauf en I.2, supposés commutatifs.

I. Anneaux quotients, idéaux

1. Montrer qu'il existe une structure d'anneau et une seule sur un ensemble à 2 éléments, respectivement sur un ensemble à 3 éléments. Étudier le cas d'un ensemble à 4 éléments. Dire lesquels sont des corps.
2. On appelle *algèbre de Boole* un anneau B tel que pour tout $x \in B$ on a $x^2 = x$. On ne suppose pas que B est commutatif. On note $\text{Spec}(B)$ l'ensemble des idéaux maximaux de B . On supposera que B est un ensemble fini dans tout l'exercice. Montrer que $\text{Spec}(B)$ est fini et que B est commutatif.
 - a) Montrer que le quotient de B par un idéal maximal est toujours isomorphe à un corps à 2 éléments que l'on notera k .
 - b) En déduire que $\text{Spec}(B)$ est en bijection avec l'ensemble des homomorphismes de B dans k .
 - c) Soit S un ensemble fini. On considère l'ensemble des fonctions de S dans k . On définit la somme de deux fonctions f et g par

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

la somme dans le membre de droite étant prise dans k . On définit la multiplication de deux fonctions f et g par

$$(f \times g)(x) = f(x) \times g(x),$$

la multiplication dans le membre de droite étant prise dans k . Montrer que l'ensemble des fonctions de S dans k est une algèbre de Boole.

- d) Montrer que l'algèbre de Boole est isomorphe à l'ensemble des fonctions de $\text{Spec}(B)$ dans k , muni de sa structure d'algèbre de Boole.
3. Montrer qu'un anneau qui a un nombre fini d'éléments, et qui est un domaine d'intégrité est un corps. On pourra étudier la bijectivité de la multiplication par un élément.
4. Montrer qu'un anneau qui est une algèbre de dimension finie sur un corps et qui est un domaine d'intégrité est un corps. On pourra étudier la bijectivité de la multiplication par un élément donné, en montrant que cette application est linéaire.
5. a) On dit qu'un élément x d'un anneau A est nilpotent s'il existe un entier n tel que $x^n = 0$. Montrer qu'un élément nilpotent appartient à tout idéal premier de l'anneau.
- b) Montrer qu'un élément qui appartient à tous les idéaux premiers d'un anneau est nilpotent. Pour ce faire, on prendra un élément non-nilpotent x et on admettra l'existence d'un idéal $\mathcal{J}$, maximal pour l'inclusion, qui ne contiennent ni x ni aucune de ces puissances. Par maximal, on veut dire ici que tout idéal contenant strictement $\mathcal{J}$ contient nécessairement une puissance de x (l'existence de cet idéal est une conséquence du lemme de Zorn). On montrera que $\mathcal{J}$ est un idéal premier et on conclura.
6. Un idéal $\mathcal{J}$ est dit primaire si la condition $xy \in \mathcal{J}$ implique qu'il existe un entier n tel que soit $x^n \in \mathcal{J}$, soit $y^n \in \mathcal{J}$.
- a) Déterminer les idéaux premiers de $\mathbb{Z}$.
- b) Étant donné un idéal $\mathcal{J}$ on appelle radical de $\mathcal{J}$ et on note $\sqrt{\mathcal{J}}$ l'idéal engendré par les éléments x tel qu'il existe un entier n tel que $x^n \in \mathcal{J}$. Montrer que pour tout idéal $\mathcal{J}$ on a $\sqrt{\mathcal{J}} = \sqrt{\sqrt{\mathcal{J}}}$.
- c) Montrer que le radical d'un idéal primaire est premier.
7. a) Soit p un nombre premier impair. Montrer que la classe de $1+p$ est d'ordre p^{n-1} dans $\mathbb{Z}/p^n\mathbb{Z}$. On raisonnera par récurrence en montrant que $(1+p)^{p^n} = 1 + \lambda_n p^{n+1}$ avec λ_n premier à p .
- b) Démontrer le théorème 2 de la section 3.
- c) Démontrer le théorème 2 de la section 3 dans le cas où $p = 2$. On considérera, entre autres, la classe de 5 modulo 2^n , et on montrera que tout élément a un ordre qui divise 2^{n-2} .
8. Soit n un entier et soit a un entier premier avec n . Montrer que $a^{\varphi(n)}$ est congru à 1 modulo n .

9. a) Calculer le reste de la division de 2081^{213} par 42, puis celui de 173^{199651} par 70.
 b) Calculer le reste de la division de $2^{3n} + 7^{2n}$ par 9 en fonction de n .
10. a) On considère le sous-ensemble de $\mathbb{C}$ constitué par les éléments de la forme $a + bi$, où a et b sont des entiers. Montrer que c'est un sous-anneau. On l'appelle l'anneau des *entiers de Gauss* et on le note $\mathbb{Z}[i]$.
 b) Déterminer l'ensemble des éléments inversibles de cet anneau.
 c) Montrer que l'anneau des entiers de Gauss est euclidien. Pour définir la structure d'anneau euclidien on utilisera la fonction $s(a + bi) = a^2 + b^2$.
11. On traitera le même problème que précédemment en remplaçant i par la racine 3-ième de l'unité j .
12. Soit A un anneau commutatif. Démontrer que, pour tous $a, b, c, d \in A$ on a $(a^2 + b^2)(c^2 + d^2) = (ab + cd)^2 + (ad - bc)^2$.
13. Cet exercice fait suite à l'exercice 9. On se place donc dans l'anneau des entiers de Gauss dont on rappelle qu'il est principal.
 a) Soit $p \in \mathbb{Z}$ un entier premier. Montrer que, dans l'anneau des entiers de Gauss, p est, soit irréductible, soit produit de deux entiers de Gauss complexes conjugués, eux-mêmes irréductibles. On pourra observer que si un élément irréductible z de l'anneau des entiers de Gauss divise p il en est de même de son complexe conjugué $\bar{z}$. On pourra aussi étudier la décomposition induite du carré du module de p . Étudier les cas de 3, 5, 7, 11, ...
 b) On admettra le résultat suivant démontré dans le chapitre 5 : soit p un nombre premier impair alors il existe dans $\mathbb{Z}/p\mathbb{Z}$ un élément u tel que $u^2 = -1$ (cette équation a lieu dans $\mathbb{Z}/p\mathbb{Z}$) si et seulement si p est congru à 1 modulo 4. Montrer que si un nombre premier p s'écrit comme somme de deux carrés d'éléments de $\mathbb{Z}$ il est congru à 1 modulo 4.
 c) Montrer qu'un nombre premier est réductible dans les entiers de Gauss si et seulement si il est congru à 1 modulo 4. On utilisera le fait que p est irréductible dans $\mathbb{Z}[i]$ si et seulement si l'anneau quotient $\mathbb{Z}[i]$ est intègre.
14. Factoriser en éléments irréductibles dans $\mathbb{Z}[i]$, les entiers de Gauss suivants : $9 + i$, $11 + 2i$, $17 + 5i$.
15. Soit k un corps. Montrer que l'anneau quotient $k[X, Y]/(XY - 1)$ est principal.
16. À quelle condition sur l'anneau A , l'anneau $A[X]$ est-il principal ?

17. a) Un anneau A est dit local s'il a un seul idéal maximal. Montrer qu'un anneau A est local si et seulement si l'ensemble des éléments non-inversibles forme un idéal.

b) Soit p un nombre premier, et soit E_p le sous-ensemble des rationnels de la forme $\frac{a}{b}$ avec b premier à p . Montrer que c'est un sous-anneau de $\mathbb{Q}$. Puis montrer que c'est un anneau local.

18. Trouver toutes les solutions entières de l'équation de Fermat en degré 2 : $x^2 + y^2 = z^2$. On commencera par supposer que les entiers x, y, z sont premiers entre eux. Puis on écrira $y^2 = (z - x)(z + x)$, et on montrera que le pgcd de $z - x$ et $z + x$ est 2.

19. On veut déterminer les nombres premiers p pour lesquels l'équation $x^2 + 2y^2 = p$ a des solutions entières.

a) On commence par montrer que -2 doit être un carré modulo p . On considérera l'anneau

$$\mathbb{Z}[i\sqrt{2}] = \{a + bi\sqrt{2} | a, b \in \mathbb{Z}\}$$

on montrera qu'il est euclidien.

b) On conclura à l'aide d'une réduction modulo p .

20. Trouver les solutions entières de $x^2 - y^3 = -2$. On se placera dans $\mathbb{Z}[i\sqrt{2}]$ dont on a montré auparavant qu'il était euclidien et donc principal.

21. Soit d un entier strictement positif. On veut étudier l'équation de Pell-Fermat $x^2 - dy^2 = 1$ et déterminer ses solutions dans $\mathbb{Z}$.

a) On se place dans l'anneau

$$\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Z}\}.$$

Commencer par vérifier que c'est bien un anneau. Montrer que les éléments $x + y\sqrt{2}$ tels que $x^2 - 2y^2 = 1$ ou $x^2 - 2y^2 = -1$ sont les éléments inversibles de l'anneau. Remarquer que ces éléments se répartissent sur une hyperbole.

b) On considère la solution particulière $3 + 2\sqrt{2}$. Montrer que parmi les solutions $a + b\sqrt{2}$ avec $a, b > 0$ c'est celle pour laquelle a est minimum, puis que si on multiplie une solution $a + b\sqrt{2}$ par $3 - \sqrt{2}$ on obtient une solution $\alpha + \beta\sqrt{2}$ avec $\alpha > 0$ et $\alpha < a$. En déduire que toutes les solutions sont au signe près puissance de $3 + 2\sqrt{2}$.

c) En corollaire on décrira tous les entiers k pour lesquels $\frac{k(k+1)}{2}$ est un carré parfait.

- 22. a) **** Soit n un entier. Soit $n = \sum_{i=1, \dots, k} 2^{a_i}$ l'écriture de n en base 2. Calculer la plus grande puissance de 2 divisant $n!$ à l'aide des a_i . On commencera par se ramener au cas où $n = 2^a$, en démontrant que

$$\frac{n!}{\prod_i 2^{a_i}!}$$

est premier à 2.

b) Traiter le même problème pour un nombre premier p quelconque.

c) En déduire l'ordre d'un p -sous-groupe de Sylow du groupe symétrique $\mathfrak{S}_n$.

23. Acheter l'exemple donné dans la section sur la cryptographie. En donner d'autres.

24. Soit n un entier positif. Donner un équivalent, en fonction de n de la longueur de son écriture en base 2.

a) On appelle opération élémentaire pour sommer deux entiers en base 2 le fait d'exécuter : soit la somme de deux 0 apparaissant comme les coefficients d'une même puissance de 2 dans les écritures en base 2, soit faire la somme d'un 0 et d'un 1 apparaissant de manière analogue, soit encore et dans la même situation si on a la somme de deux 1 effectuer une retenue, soit enfin si on a la somme de deux 1 et une retenue effectuer une retenue conserver un 1. Pour une formulation plus précise on renvoie à [K].

Donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour sommer deux entiers en base 2 ayant des développements de longueur k et ℓ en base 2 (la longueur du développement est l'exposant de la puissance maximale de 2 qui y apparaît à laquelle on ajoute 1).

b) Donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour sommer n entiers (en base 2) ayant des développements de longueur k_i en base 2.

c) En utilisant les opérations précédentes donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour multiplier deux entiers ayant des développements de longueur k et ℓ .

d) Donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour calculer $n!$ en base 2.

e) Donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour calculer le pgcd de deux entiers a et b avec $a \geq b$, dont les développements en base 2 sont respectivement de longueur k et ℓ .

f) Donner une borne supérieure au nombre d'opérations élémentaires à effectuer pour écrire un nombre, donné par un développement en base 2, en base 10.

II. Polynômes

1. On considère l'anneau des polynômes à coefficients dans $\mathbb{Q}$. Pour un entier n donné on pose

$$E_n(X) = \frac{X(X-1)\cdots(X-n+1)}{n!}.$$

- a) On définit un sous-anneau E de $\mathbb{Q}[X]$ par $E = \{P \in \mathbb{Q}[X] \mid \forall k \in \mathbb{Z} P(k) \in \mathbb{Z}\}$. Vérifier que c'est un sous-anneau.
- b) Soit $P \in E$, montrer qu'il existe une unique famille finie d'entiers λ_i tels que $P = \sum_i \lambda_i E_i$.
- c) Sur l'ensemble des polynômes à coefficients dans $\mathbb{Q}$, on définit une opération Δ par $\Delta(P)(X) = P(X+1) - P(X)$. Calculer $\Delta(E_n)$.
2. ** On conserve les notations de l'exercice précédent. Soit p un nombre premier. On considère l'anneau quotient $E/(p)$. Montrer que cet anneau contient un sous-anneau F_n isomorphe à l'anneau quotient

$$\mathbb{F}_p[X_1, \dots, X_n] / (X_1^p - X_1, \dots, X_n^p - X_n),$$

et que $E/(p)$ est réunion des F_n .

3. a) Soit $P \in \mathbb{R}[x]$ de degré $n-1$. Démontrer que tout polynôme de degré $\leq n-1$ peut se mettre sous la forme

$$\alpha_0 P(X) + \alpha_1 P(X+1) + \cdots + \alpha_{n-1} P(X+n-1).$$

- b) En déduire que le déterminant suivant est nul :

$$\begin{vmatrix} P(X) & P(X+1) & \cdots & P(X+n) \\ P(X+1) & P(X+2) & \cdots & P(X+n+1) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ P(X+n) & P(X+n+1) & \cdots & P(X+2n) \end{vmatrix}$$

4. a) Montrer que les polynômes irréductibles dans $\mathbb{R}[X]$ sont de degré 1 ou 2, les caractériser. Décrire la factorisation en éléments irréductibles dans $\mathbb{R}[X]$.
- b) Montrer que tout polynôme à coefficients réels qui ne prend que des valeurs positives sur la droite réelle peut s'écrire $P^2 + Q^2$, où P et Q sont des polynômes à coefficients réels. On étudiera d'abord le cas des polynômes irréductibles.
5. ** ([J]) Soit P un polynôme à coefficients réels, et soit $[a, b]$ un segment. Soit $P = P_0, P_1, \dots, P_s$ une suite de polynômes telle que :
- $P(a)P(b) \neq 0$,
 - P_s n'a pas de zéros sur $[a, b]$,
 - si $c \in [a, b]$ est un zéro de P_j , $0 < j < s$ alors $P_{j-1}(c)P_{j+1}(c) < 0$,

- si $P(c)=0$, $c \in]a, b[$ il existe $c_1 < c$ et $c_2 > c$ tels que pour $x \in]c_1, c[$ $P_0(x)P_1(x) < 0$ et pour $x \in]c, c_2[$ $P_0(x)P_1(x) > 0$.

Une telle suite, si elle existe est appelée *une suite de Sturm* sur $[a, b]$ pour P .

a) Montrer que si P a une suite de Sturm sur $[a, b]$ il n'a que des zéros simples sur $[a, b]$.

b) Étant donnée une suite de réels différents de zéro $(t_1, t_2, \dots, t_k)$ on définit le nombre de changements de signe de la suite comme étant le nombre d'entiers i avec $1 \leq i \leq k-1$ et $t_i t_{i+1} < 0$. Étant donné $x \in [a, b]$, on note $V(x)$ le nombre de changements de signe de la suite obtenue à partir de la suite $P_0(x), P_1(x), \dots, P_s(x)$, en éliminant les termes nuls.

Montrer que le nombre de zéros de P dans $[a, b]$ est égal à $V(x) - V(y)$.

6. ([J]) Soit P un polynôme de degré s à coefficients réels qui n'a que des zéros simples dans $[a, b]$. Montrer que la suite définie par $P_0 = P$, $P_1 = P'$, et P_k est l'opposé du reste de la division de P_{k-2} par P_{k-1} est une suite de Sturm pour P sur $[a, b]$.

7. Appliquer le résultat précédent à la recherche du nombre de zéros entre 0 et 15 du polynôme $X^7 - X^5 + 2X^2 + X - 5$ et $X^9 + X^6 - 6X^5 + 3X^3 - X + 1$.

8. Soit A un anneau, on se place dans l'anneau $A[X]$. Démontrer la formule de Leibniz pour la dérivation des polynômes. Établir la formule de dérivation pour un polynôme de la forme $P(Q(X))$.

9. Soit p un nombre premier. Soit le corps $\mathbb{Z}/p\mathbb{Z}$, et soient u, v, w des éléments différents de zéro de $\mathbb{Z}/p\mathbb{Z}$. Montrer que l'équation en x, y : $ux^2 + vy^2 = w$ admet au moins une solution. On comptera le nombre d'éléments de la forme $ux^2 + w - vy^2$ quand x et y décrivent $\mathbb{Z}/p\mathbb{Z}$.

10. ** ([FS]) Montrer que le polynôme $\prod_j (X - m_j) - 1$ où les m_j sont des entiers deux à deux distincts est irréductible dans $\mathbb{Z}[X]$.

11. (Lemme de Lazard) Montrer que les polynômes homogènes de degré n , H en 2 variables X et Y , à coefficients dans $\mathbb{Q}$, tels que $H(X, Y) = H(Y, X)$, $H(X, 0) = 0$ et que

$$H(X, Y) + H(X + Y, Z) = H(Y, Z) + H(X, Y + Z),$$

sont de la forme $\gamma((X + Y)^n - X^n - Y^n)$, $\gamma \in \mathbb{Q}$.

12. ** (Lemme de Lazard pour $\mathbb{F}_p$) Soit p un nombre premier.
Montrer que les coefficients du polynôme (à coefficients entiers)

$$((X + Y)^{p^a} - X^{p^a} - Y^{p^a})$$

sont divisibles par p et pas par p^2 . Dans la suite, la notation $(\frac{1}{p}((X + Y)^{p^a} - X^{p^a} - Y^{p^a}))$ désignera la réduction modulo p du polynôme à coefficients entiers. Montrer que les polynômes homogènes de degré n , H en 2 variables X et Y , à coefficients dans $\mathbb{F}_p$, tels que $H(X, Y) = H(Y, X)$, $H(X, 0) = 0$ et que

$$H(X, Y) + H(X + Y, Z) = H(Y, Z) + H(X, Y + Z),$$

sont de la forme $\gamma((X + Y)^n - X^n - Y^n)$, $\gamma \in \mathbb{F}_p$ si n n'est pas une puissance de p , et de la forme $\gamma(\frac{1}{p}((X + Y)^{p^a} - X^{p^a} - Y^{p^a}))$ sinon.

13. (Formule d'Euler) Soit P un polynôme en n variables, $X_1, \dots, X_n$, ne comportant que des monômes de degré total m . Montrer que :

$$X_1 P'_{X_1} + \dots + X_n P'_{X_n} = mP.$$

14. Exprimer les polynômes suivants à l'aide des fonctions symétriques élémentaires :

- $X^3 Y^2 + X^2 Y^3 + X^2 Z^3 + X^3 Z^2 + Y^3 Z^2 + Y^2 Z^3,$
- $X^5 Y^3 + X^3 Y^5 + X^3 Z^5 + X^5 Z^3 + Y^5 Z^3 + Y^3 Z^5.$

15. a) Exprimer les polynômes symétriques suivants à l'aide des sommes de Newton :

- $\sum_{i \neq j} X_i^a X_j^b,$
- $\sum_{i \neq j \neq k \neq i} X_i^a X_j^b X_k^c,$
- $\sum_{i < j} (X_i + X_j)^k.$

b) Soit le déterminant

$$\Delta = \begin{vmatrix} 2XYZ & -Z^3 & -Y^3 & X \\ -Z^3 & -2XYZ & -X^3 & Y \\ -Y^3 & -X^3 & 2XYZ & Z \\ X & Y & Z & 0 \end{vmatrix}$$

Montrer que Δ est homogène de degré 8, puis qu'il est symétrique en X^2, Y^2, Z^2 . Le calculer. Δ est-il divisible par $X + Y + Z$?

16. Calculer le discriminant du polynôme $X^n + pX + q$.

17. Soit $P \in k[X]$ un polynôme de degré n . Supposons qu'il ait toutes ses racines dans k , soient $\lambda_1, \dots, \lambda_n$. Montrer que

$$\text{Dis}(P) = (-1)^n \prod_{i=1, \dots, n} P'(\lambda_i).$$

18. On garde les hypothèses de l'exercice précédent. Montrer que

$$\text{Dis}(P) = (-1)^n n^n a_n^n \prod_{i \neq j} (\lambda_i - \lambda_j).$$

19. (D'après agrégation 1993) On considère l'anneau $\mathbb{C}[X, Y]$. Soit f une fonction holomorphe sur le plan complexe. Déterminer pour $f = e^z$, $f = ze^z$, et $f = \sin(z)$ l'idéal de $\mathbb{C}[X, Y]$ constitué par les polynômes P tels que $P(f, f') = 0$.

On pose la même question en remplaçant $\mathbb{C}[X, Y]$ par $\mathbb{C}[X, Y, Z]$, pour $f = xe^z$, et $f = \cos(z)$, et les polynômes P tels que $P(f, f', f'') = 0$.

20. (théorème de Gauss-Lucas) Soit $P \in \mathbb{C}[X]$. Montrer que les zéros de P' sont dans l'enveloppe convexe des zéros de P .

21. Soit k un corps ayant un nombre infini d'éléments. Soit $P \in k[X_1, \dots, X_n]$.

a) Soit P qui prend la valeur zéro pour tout élément appartenant à un ensemble de la forme $A_1 \times \dots \times A_n$, où chaque A_i est un sous-ensemble de k de cardinal infini. Montrer que P est nul.

b) Soit $k = \mathbb{C}$ ou $k = \mathbb{R}$. Montrer qu'un polynôme $P \in k[X_1, \dots, X_n]$, qui s'annule sur un ouvert de k^n est nul.

22. Montrer qu'un polynôme homogène en 3 variables de degré 2 sur un corps fini a toujours une solution (α, β, γ) non triviale. Généraliser à 4 variables en degré 3.

III. Séries formelles

1. Montrer que l'anneau des séries formelles à coefficients dans un corps est un anneau local (pour la définition voir l'exercice 17 de la section I).

2. Notons $p(n)$ le nombre de partitions d'un entier naturel n donné. Donner un sens au produit infini suivant :

$$\prod_{n \geq 1} \frac{1}{1 - T^n}.$$

Montrer qu'il est égal à :

$$\phi(T) = \sum_n p(n) T^n.$$

3. Montrer que la série formelle :

$$1 + \sum_n \frac{T^n}{\prod_{i=1, \dots, n} (1 - T^i)}$$

est égale à $\Phi(T)$.

4. (Partitions p -régulières) On garde les notations de 2 et 3. Soit p un nombre premier et soit $\pi_p(n)$ le nombre de partitions de n en somme d'entiers premiers à p . Soit $\Phi_p(T) = \sum_n \pi_p(n)T^n$. Montrer que $\Phi_p(T) = \frac{\Phi(T^p)}{\Phi(T)}$.

5. a) (Formules de Newton)

Le but de cet exercice est de donner une seconde démonstration des formules de Newton. On se place dans l'anneau $k[X_1, \dots, X_n][[T]]$, k étant un corps. On utilisera la dérivation des séries formelles.

On considère la série formelle :

$$\sum_{i=1, \dots, n} \frac{X_i}{1 - X_i T}.$$

Montrer que :

$$\sum_{i=1, \dots, n} \frac{X_i}{1 - X_i T} = \sum_i N_{i+1} T^i.$$

b) Montrer que la série formelle précédente est égale à l'opposée de la dérivée par rapport à T de la série suivante :

$$\frac{d}{dT} \log \left(\prod_i (1 - X_i T) \right).$$

c) On note Σ_i les fonctions symétriques élémentaires des X_i . Montrer que :

$$\frac{d}{dT} \log \left(\prod_i (1 - X_i T) \right) = \frac{-\Sigma_1 + 2\Sigma_2 T + \dots + (-1)^n n \Sigma_n T^{n-1}}{1 - \Sigma_1 T + \dots + (-1)^n \Sigma_n T^n}.$$

d) En identifiant, établir les formules de Newton.

Quelques réponses ou indications

I. Anneaux quotients, idéaux

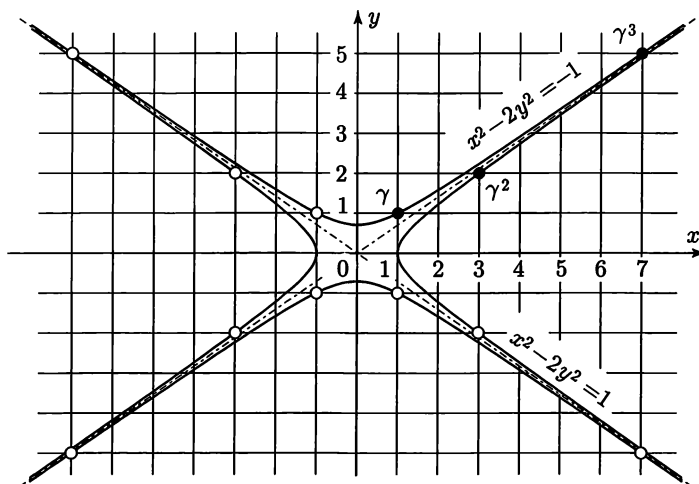
1. Dans les cas à 2 et 3 éléments on trouve $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/3\mathbb{Z}$. Dans le cas d'un ensemble à 4 éléments commencer par étudier la structure comme groupe additif. Puis discuter suivant que l'anneau est un corps ou non.

2. b) On montre la commutativité à l'aide de $(x+y)^2 = x+y$. Il y a un nombre fini d'idéaux maximaux car ce sont des sous-ensembles de B . Si $\mathcal{J}$ est maximal, l'anneau $B/\mathcal{J}$ est aussi une algèbre de Boole. Si c'est un corps commutatif, comme tout élément y est zéro de $X^2 - X$, il ne peut avoir que deux éléments 0 et 1.

d) Montrer que l'intersection des idéaux maximaux est réduite à $\{0\}$. Puis utiliser le Lemme Chinois et l'application de B dans l'ensemble des fonctions de $\text{Spec}(B)$ dans k qui à $a \in B$ associe $\hat{a} : \text{Spec}(B) \rightarrow k$ donnée par $\hat{a}(\mathcal{J}) = 0$ si $a \in \mathcal{J}$ et $\hat{a}(\mathcal{J}) = 1$ si $a \notin \mathcal{J}$.

3. Montrer que la multiplication par un élément donné différent de zéro détermine une application injective, donc bijective, car l'ensemble est fini.
4. Montrer que la multiplication par un élément donné différent de zéro détermine une application est linéaire injective donc bijective car l'espace vectoriel est de dimension finie.
5. b) On raisonne par l'absurde. On suppose que $\mathcal{I}$ n'est pas un idéal premier. Soit $ab \in \mathcal{I}$ avec $a \notin \mathcal{I}$ et $b \notin \mathcal{I}$. Alors $(\mathcal{I}, a)$ et $(\mathcal{I}, b)$ contiennent chacun une puissance de x . En conclure que $\mathcal{I}$ contient une puissance de x , ce qui est impossible.
7. a) Raisonner par récurrence sur n en tenant compte du fait que le coefficient binomial C_p^i est divisible par p (premier) si $0 < i < p$. On pourra redémontrer ce fait ou se reporter au chapitre 5.
8. Utiliser le fait que l'ordre d'un élément divise l'ordre du groupe et que $\phi(n)$ est l'ordre de $(\mathbb{Z}/n\mathbb{Z})^*$.
10. b) Les éléments inversibles sont $1, -1, i, -i$.
c) Étant donné $z, u \in \mathbb{Z}[i]$, $u \neq 0$, on considère $\frac{z}{u} \in \mathbb{C}$. On définit q comme étant l'un des points de $\mathbb{Z}[i]$ le plus proche de $\frac{z}{u}$. Montrer que $s(z - uq) < s(u)$.
13. a) Si on a $p = zu$, avec z irréductible on a aussi $p = \bar{z}\bar{u}$. Donc $\bar{z}$ divise aussi p . Maintenant, on a donc une décomposition du type $p = \prod_i z_i \bar{z}_i$, avec les z_i irréductibles. On écrit donc $p^2 = \prod_i |z_i|^2$. Discuter suivant que z_i est dans $\mathbb{Z}$ ou non.
Par exemple $5 = (2 + i)(2 - i)$.
b) Si on a $p = a^2 + b^2$, on obtient le résultat par réduction modulo p de cette formule et application du critère.
Puis on montre que $p \in \mathbb{Z}$ premier est irréductible dans $\mathbb{Z}[i]$ si et seulement si il n'est pas somme de deux carrés d'entiers.
c) Enfin on montre que $p \in \mathbb{Z}$ est irréductible, si et seulement si -1 est un carré modulo p , en démontrant que $\mathbb{Z}[i]/(p) \cong \mathbb{Z}[X]/(p, X^2 + 1) \cong \mathbb{F}_p[X]/(X^2 + 1)$. Puis, on utilise le fait que $p \in \mathbb{Z}$ premier est irréductible dans $\mathbb{Z}[i]$, si et seulement si $\mathbb{Z}[i]$ est intègre.
14. Par exemple chercher à exprimer le module de $9 + i$ au carré comme produit de somme de deux carrés. On a $82 = 2 \times 41 = (1 + 1) \times (25 + 16)$. Donc $9 + i = (1 + i)(5 - 4i)$. Il faut vérifier que les éléments sont irréductibles ou poursuivre la décomposition.
15. Identifier $k[X, Y]/(XY - 1)$ au sous-anneau de $k(X)$ des éléments de la forme $\frac{P}{X^k}$, où $P \in k[X]$ et $k \in \mathbb{Z}$ à l'aide de l'homomorphisme qui, à $P(X, Y)$, associe $P(X, \frac{1}{X})$. Puis étudier cet anneau, en utilisant le fait qu'il contient $k[X]$.
18. D'après la décomposition $y^2 = (z - x)(z + x)$ si un nombre premier impair divise donc $x - z$ et $x + z$, il divise $2x$ et $2z$, donc x et z , donc aussi y ce qui est exclu par hypothèse. De même si 4 divise $x - z$ et $x + z$ on déduit que 2 divise y . On pose alors $z - x = 2\alpha$ et $x + z = 2\beta$, α et β premiers entre eux. Alors l'équation $y^2 = (z - x)(z + x)$ montre que α et β sont des carrés.
En général si d est le pgcd de x, y et z on trouve $x = d(u^2 - v^2)$, $y = 2duv$ et $z = d(u^2 + v^2)$.

20. On montrera $i\sqrt{2}$ est premier dans $\mathbb{Z}[i\sqrt{2}]$, on procédera comme il a été fait pour les entiers de Gauss. Écrire alors une décomposition en facteurs premiers de l'identité $(x+i\sqrt{2})(x-i\sqrt{2})=y^3$. Montrer qu'un nombre premier (irréductible) de $\mathbb{Z}[i\sqrt{2}]$ qui divise $(x+i\sqrt{2})$ et $(x-i\sqrt{2})$ est nécessairement $i\sqrt{2}$. en déduire que $i\sqrt{2}$ ne divise pas x . En utilisant la décomposition en facteurs premiers dans $\mathbb{Z}[i\sqrt{2}]$ en conclure que $(x+i\sqrt{2})$ est un cube dans $\mathbb{Z}[i\sqrt{2}]$: $(a+bi\sqrt{2})^3 = x+i\sqrt{2}$ en identifiant montrer que $3a^2b-2b^2=1$.
21. a) b) c). On écrit $(a+b\sqrt{2})^n = (a_n+b_n\sqrt{2})$ et $(a-b\sqrt{2})^n = (a_n-b_n\sqrt{2})$ avec a_n et b_n entiers. On observe que si $a^2-2b^2 = (a+b\sqrt{2})(a-b\sqrt{2}) = 1$ alors $(a+b\sqrt{2})^n(a-b\sqrt{2})^n = (a_n+b_n\sqrt{2})(a_n-b_n\sqrt{2}) = 1$. L'entier n étant quelconque dans $\mathbb{Z}$. Pour la dernière question on se ramènera à l'équation précédente. On observera les solutions sur le graphe ci-dessous.



22. a) La plus grande puissance de 2 divisant $2^n!$ est 2^{2^n-1} .
24. On trouve $\frac{\ln(n)}{\ln(2)}$. On a les majorants suivants $C_{\text{sup}}(k, \ell)$, $C_{\text{sup}}(k_i)$, la constante dépend de n , $C(k\ell)$, $Cn^2\ln^2(n)$, et Ck^3 .

II. Polynômes

1. b) Pour montrer qu'il existe une unique famille finie d'entiers λ_i tels que $P = \sum_i \lambda_i E_i$ on évaluera le polynôme P en des entiers bien choisis.
- c) $\Delta(E_n) = E_{n-1}$.
2. Pour définir F_n , considérer le sous-anneau engendré par $E_1, E_p, \dots, E_{p^n}$.
4. a) La factorisation est, en polynômes de degré 1, et en polynômes de degré 2 à racines complexes conjuguées.
- b) Utiliser la décomposition en facteurs irréductibles et l'exercice 11 ci-dessus.

5. a) Utiliser les critères du cours.
b) Étudier le comportement de la fonction $V(x)$ quand x décrit le segment. En particulier étudier le comportement lors du franchissement d'un des zéros de la suite de polynômes.
7. Pour montrer que les polynômes n'ont que des racines simples, on réduira modulo 2 ou 3, et on utilisera le discriminant.
8. Se ramener au cas de monômes.
9. Compter le nombre d'éléments de la forme ux^2 et $w - vy^2$ quand x et y décrivent $\mathbb{Z}/p\mathbb{Z}$. Montrer que chacun des ensembles a $\frac{p+1}{2}$ éléments. Constater que les deux ensembles sont d'intersection non vide.
10. Raisonner par l'absurde, et écrire $P = QR$. Les valeurs prises par Q et R en les m_j sont nécessairement 1 ou -1 . Si le degré de Q est différent de celui de R montrer que Q ou R prend la même valeur (1 ou -1) un nombre de fois supérieur à son degré. Si le degré de Q est égal à celui de R et si la condition précédente n'est pas réalisée en conclure que Q et R sont opposés. Conclure.
11. Écrire le polynôme cherché H sous la forme $H = \sum_i a_{i,j} X^i Y^{n-i}$. Calculer les coefficients de proche en proche à l'aide des relations.
12. On procédera comme plus haut, en prenant soin de vérifier que les coefficients binomiaux utilisés ne sont pas nuls modulo p .
15. Exprimer les polynômes symétriques suivants à l'aide des sommes de Newton :
 - $N_a N_b - N_{a+b}$,
 - faire apparaître $N_a N_b N_c$ ainsi que des produits de la forme $N_{a+b} N_c$.
16. Dans le déterminant, soustraire les $n-1$ premières colonnes aux $n-1$ dernières. On trouve $n^n q^{n-1} + (-1)^{n-1} (n-1)^{n-1} p^n$.
17. On se place dans l'anneau des polynômes $k[X, T_1, \dots, T_n] \cong k[X_1, \dots, X_n][T]$. On considère le polynôme $\prod_i (X - T_i)$. On calcule le discriminant de ce polynôme. Montrer que c'est un polynôme symétrique de $k[T_1, \dots, T_n]$. Montrer qu'il est divisible par $T_i - T_j$, $i \neq j$. Puis par $(T_i - T_j)^2$, pour ce dernier point, on utilisera la symétrie. Comparer les degrés pour conclure qu'au signe près le produit $\prod_{i \neq j} (T_i - T_j)$ est égal au discriminant. Enfin remplacer les T_i par les λ_i .
18. Appliquer l'exercice précédent.
19. Dans le premier cas on trouve l'idéal engendré par $(X - Y)$, dans le second, celui engendré par $X^2 + Y^2 - 1$. Pour le quatrième, on se ramène à un analogue du second en adjoignant $X - Z$. Pour le troisième, on cherchera une équation différentielle du second ordre. Puis on cherchera à déterminer si la fonction peut satisfaire une équation du premier ordre.
20. Écrire la décomposition en éléments simples de la fraction rationnelle $\frac{P'}{P}$.

III. Séries formelles

1. Soit $k[[X]]$, montrer que le seul idéal maximal est (X) . Les autres idéaux sont de la forme (X^n) .

2. Pour donner un sens au produit infini :

$$\prod_{n \geq 1} \frac{1}{1 - T^n},$$

on observe que les coefficients de

$$\prod_{1 \leq i \leq n} \frac{1}{1 - T^i},$$

et

$$\prod_{1 \leq i \leq n+1} \frac{1}{1 - T^i},$$

coïncident jusqu'en degré n .

En effectuant le développement on montre que le coefficient du terme de degré n est $p(n)$.

3. Montrer que le coefficient du terme de degré k dans le développement de

$$\frac{T^n}{\prod_{i=1, \dots, n} (1 - T^i)}$$

est le nombre de partitions de k , en au plus n entiers.

4. Procéder comme en 1, en modifiant ce qu'il faut.

5. a) Écrire le développement.

b) Calculer la dérivée comme à l'ordinaire.

c) et d) On reconnaît les fonctions symétriques en développant le produit $\prod_i (1 - X_i T)$. Puis, on identifie à l'expression à l'aide des N_h , et on multiplie des deux côtés par le dénominateur. Enfin, on identifie les coefficients de T^n .

Chapitre 4

Extensions des corps

Applications

Ce chapitre est consacré à la théorie des corps. Sauf dans la section 8 tous les corps considérés seront supposés commutatifs. On commencera par définir la notion de caractéristique d'un corps, on en donnera quelques conséquences, par exemple la formule de Newton en caractéristique p . Après ces généralités on passera à l'étude des éléments algébriques sur un corps et du corps de rupture d'un polynôme. On ne fera pas une étude détaillée du corps de décomposition et l'existence et l'unicité de la clôture algébrique sera admise. On passera après à la description des corps finis. L'étude des racines primitives de l'unité et des polynômes cyclotomiques sera faite, avec en perspective, le théorème de Wedderburn et un premier aperçu sur les extensions cyclotomiques.

On reviendra dans les deux dernières sections sur les corps finis en décrivant l'algorithme de factorisation des polynômes sur un corps fini dû à Berlekamp. Enfin on donnera une courte introduction à la théorie des codes correcteurs d'erreurs, elle sera basée sur l'étude d'un code BCH . Le matériel décrit dans ces sections n'est pas plus difficile que celui des sections précédentes, mais dans la mesure où elles mettent en jeu de longs enchaînements d'arguments, leur lecture n'est pas conseillée dans une première approche. Par ailleurs leur côté introductif peut aussi rendre leur lecture plus difficile. Leur présence est cependant nécessaire dans ce livre dans la mesure où elles abordent des sujets riches en applications. Pour ces questions on renvoie à [LN] et [MS] dont ont été tirés certains énoncés et exercices.

1. Définitions, caractéristique d'un corps, applications

On rappelle la définition d'un corps :

Définition 1

On appelle corps, un anneau dans lequel tout élément non nul, c'est-à-dire distinct de l'élément neutre de la loi additive, a un inverse pour la loi multiplicative.

L'élément neutre de la multiplication sera comme d'habitude noté 1 pour tous les corps considérés. On aura soin en lisant le texte de bien vérifier à quel corps (ou éventuellement anneau) appartient cet élément. La multiplication sera toujours supposée commutative, sauf dans la section 6 où on démontrera le théorème de Wedderburn.

Voici trois exemples fondamentaux de corps :

- le corps $\mathbb{Q}$ des nombres rationnels dont la construction est donnée dans les rappels,
- le corps des fractions rationnelles $\mathbb{Q}(X)$,
- enfin si p est un nombre premier l'anneau quotient $\mathbb{Z}/p\mathbb{Z}$ qui est un corps et que l'on note $\mathbb{F}_p$.

Définition 2

Étant donné un corps k on appelle homomorphisme caractéristique, l'homomorphisme d'anneaux, $c : \mathbb{Z} \rightarrow k$, défini par :

- $c(m) = \underbrace{1 + \cdots + 1}_{m \text{ fois}} = m.1$ pour $m \geq 0$,
- $c(m) = -c(-m)$, pour $m < 0$.

On rappelle que 1 désigne l'élément neutre de la loi multiplicative du corps.

Considérons I le noyau de cet homomorphisme, c'est un idéal. Comme le quotient de $\mathbb{Z}$ par I est un sous-anneau d'un corps, il est intègre. Ceci implique que I est, soit trivial, c'est-à-dire réduit à $\{0\}$, soit premier.

- S'il est premier, il est maximal et constitué par les multiples d'un nombre premier p puisque $\mathbb{Z}$ est principal.
- S'il est trivial le corps est dit de *caractéristique zéro*, si l'idéal est constitué par les multiples d'un nombre premier p le corps est dit de caractéristique p . Le corps $\mathbb{Q}$ est de caractéristique 0, le corps $\mathbb{F}_p$ est de caractéristique p .

Observons maintenant qu'un corps k contient toujours un plus petit sous-corps, c'est l'intersection de tous les sous-corps contenus dans k .

Proposition 1. Si le corps k est de caractéristique 0 il contient comme plus petit sous-corps un sous-corps isomorphe à $\mathbb{Q}$. Si le corps k est de caractéristique p il contient comme plus petit sous-corps un sous-corps isomorphe à $\mathbb{F}_p$. Dans chaque cas ce sous-corps est appelé le corps premier.

Démonstration. Dans le second cas il suffit d'observer que l'image de l'homomorphisme caractéristique est isomorphe à $\mathbb{F}_p$. Et que par ailleurs cette image est contenue, par construction, dans tout sous-corps de k , en effet un sous-corps contient l'unité et donc tous les éléments dans l'image de l'homomorphisme caractéristique. Pour ce qui est de la première affirmation on constate que l'image de l'homomorphisme est isomorphe à $\mathbb{Z}$, et que le plus petit sous-corps de k contenant cette image est isomorphe à $\mathbb{Q}$. ■

Le résultat suivant est juste une reformulation de la définition de la caractéristique, son importance fait qu'il mérite d'être dégagé.

Proposition 2. Soit k un corps de caractéristique p . Pour tout $x \in k$ on a

$$\underbrace{x + \cdots + x}_{p \text{ fois}} = p.x = 0.$$

Démonstration. Il suffit d'observer que

$$\underbrace{x + \cdots + x}_{p \text{ fois}} = p.x = (p.1)x = 0. \quad \blacksquare$$

Dans la suite on notera $n.x$ simplement nx . Voici une autre propriété capitale de la caractéristique p :

Théorème 1. Soient a et b deux éléments d'un corps k de caractéristique p . Alors pour tout n on a :

$$(a + b)^{p^n} = a^{p^n} + b^{p^n}.$$

Démonstration. Rappelons que k est supposé commutatif. La démonstration se fait par récurrence sur n .

Pour le cas $n = 1$ on commence par écrire la formule de Newton :

$$(a + b)^p = \sum_{i=0}^p C_p^i a^i b^{p-i},$$

donc on a :

$$(a + b)^p = a^p + \sum_{i=1}^{p-1} C_p^i a^i b^{p-i} + b^p,$$

le coefficient binomial C_p^i est égal à $\frac{p!}{i!(p-i)!}$. Pour $0 < i < p$ le nombre premier p apparaît au numérateur et n'apparaît pas au dénominateur. On a donc $i!C_p^i = pq$,

avec q entier. Comme p ne divise pas $i!$ il divise C_p^i . Les coefficients binômiaux C_p^i sont donc divisibles par p et donc nuls modulo p pour les valeurs de i considérées. Ceci implique donc que $(a+b)^p = a^p + b^p$.

Passons au cas général et raisonnons par récurrence. Supposons le résultat établi pour $n-1$. On a alors :

$$(a+b)^{p^n} = ((a+b)^{p^{n-1}})^p = (a^{p^{n-1}} + b^{p^{n-1}})^p = a^{p^n} + b^{p^n},$$

la première égalité est déduite de l'hypothèse de récurrence, la seconde l'est du cas $n=1$. ■

Donnons un exemple de calcul modulo 2. Si on veut calculer $(a+b)^9$, où a et b sont éléments d'un corps de caractéristique 2 on écrit :

$$(a+b)^9 = (a+b)(a+b)^8 = (a+b)(a^8 + b^8) = a^9 + a^8b + ab^8 + b^9.$$

Voici deux applications très classiques de ce qui précède. Le petit théorème de Fermat (théorème 2) et le Théorème de Wilson (théorème 3).

Théorème 2. Soit p un nombre premier, et soit n un entier. Alors n^p est congru à n modulo p .

Démonstration. Il suffit de montrer que les classes résiduelles modulo p de n^p et n sont égales.

- Si n est divisible par p les deux classes sont nulles modulo p .
- Sinon, comme $\mathbb{F}_p$ est un corps, il suffit de montrer que pour tout élément non nul $x \in \mathbb{F}_p$ on a $x^{p-1} = 1$. Autrement dit, il suffit de montrer que l'ordre de tout élément dans le groupe multiplicatif $\mathbb{F}_p^*$ est un diviseur de $p-1$. Mais ceci résulte de ce que l'ordre de ce groupe est $p-1$. ■

Théorème 3. Le produit $(p-1)!$ est congru à -1 modulo p .

Démonstration. On supposera que $p \geq 3$. Classons les éléments du groupe $\mathbb{F}_p^*$ selon qu'ils sont égaux à leur inverse ou non. Un élément x égal à son inverse est tel que $x = x^{-1}$, donc que $x^2 = 1$. Un élément est égal à son inverse si et seulement si il est zéro du polynôme à coefficients dans $\mathbb{F}_p$ $X^2 - 1$. Ce polynôme a pour zéro 1 et -1 . Les autres éléments de $\mathbb{F}_p^*$ se groupent par paires d'éléments distincts (x_i, y_i) , $1 \leq i \leq \frac{p-1}{2}$, inverses l'un de l'autre : $x_i y_i = 1$. On a donc, en identifiant un entier à sa classe résiduelle modulo p :

$$(p-1)! = 1 \times (-1) \times \prod_{1 \leq i \leq \frac{p-1}{2}} x_i y_i = -1. \quad \blacksquare$$

Le résultat suivant sera utilisé plus loin :

Proposition 3. Soit k un corps de caractéristique p . Soit P un polynôme à coefficients dans k , alors la dérivée formelle P' de P est nulle si et seulement si il existe un polynôme Q à coefficients dans k tel que $P(X) = Q(X^p)$.

Démonstration. Soit $P = \sum_{i \geq 0} a_i X^i$, alors $P' = \sum_{i > 0} i a_i X^{i-1}$. Pour que P' soit nul il faut et suffit que tous les coefficients soient nuls. Si p divise i le coefficient $i a_i$ est nul. Sinon il est non nul dès que a_i est non nul car la classe de congruence de i est non nulle. Donc pour que P' soit nul il faut et suffit que les coefficients a_i , avec p ne divisant pas i , soient nuls. Donc $P = \sum_{i \geq 0} b_{pi} X^{pi}$ et $P = Q(X^p)$ avec $Q = \sum_i b_{pi} X^i$. Le résultat suit. ■

On notera pour conclure cette partie qu'il existe des corps de caractéristique p ayant un nombre infini d'éléments, par exemple le corps $\mathbb{F}_p(X)$ des fractions rationnelles à coefficients dans $\mathbb{F}_p$.

2. Éléments algébriques dans un corps et extensions algébriques des corps

Définition 1

Soit K un corps et k un sous-corps. On dit que K est une *extension* de k .

Par exemple $\mathbb{C}$ est une extension de $\mathbb{R}$. Voilà une remarque élémentaire mais fondamentale :

Proposition 1. Soit K un corps et k un sous-corps. Le corps K est un espace vectoriel sur le corps k . Dans ces conditions on appelle k le corps de base.

Démonstration. Rappelons qu'un espace vectoriel sur un corps k est un ensemble K muni d'une loi de groupe abélien et d'une loi externe, c'est-à-dire d'une application $k \times K$ dans K satisfaisant à certaines propriétés (chapitre 0). La loi de groupe abélien sur K est évidemment l'addition de K . Pour ce qui est de la loi externe on choisit l'application définie comme étant la composée de l'inclusion de $k \times K$ dans $K \times K$ suivie de la multiplication de $K \times K$ dans K . Elle vérifie les propriétés requises. Chacune des propriétés à vérifier est impliquée par l'associativité de la multiplication de K et la distributivité de la multiplication de K par rapport l'addition de K . En effet ce sont des cas particuliers de ces dernières propriétés. ■

Définition et notation

On notera $[K : k]$ la dimension de K en tant que k espace vectoriel. On appellera cette dimension le *degré de l'extension* sur le corps de base k .

Proposition 2. Soit k un corps fini de caractéristique p . Son cardinal est de la forme p^n .

Démonstration. Le corps k contient un sous-corps isomorphe à $\mathbb{F}_p$, et que l'on identifiera à $\mathbb{F}_p$. C'est donc d'après la proposition précédente un espace vectoriel sur $\mathbb{F}_p$.

Commençons par démontrer que comme k est fini il est de dimension finie sur $\mathbb{F}_p$. Il suffit d'observer qu'un espace vectoriel qui n'est pas de dimension finie contient des systèmes libres de cardinal arbitrairement grand, donc ne peut avoir un nombre fini d'éléments.

Soit donc n le cardinal d'une base de k sur $\mathbb{F}_p$, et soit $\{e_1, \dots, e_n\}$ une base. Tout élément x de k a une écriture unique sous la forme : $\sum_{i=1, \dots, n} \lambda_i e_i$ où $\lambda_i \in \mathbb{F}_p$. On a donc p choix pour la première coordonnée, p choix pour la seconde, et ainsi de suite. On a à choisir successivement n fois entre p éléments. Ces choix se multiplient. On a donc en fin de compte p^n choix, soit p^n éléments. ■

Le corollaire suivant est une généralisation du petit théorème de Fermat.

Corollaire. Soit k un corps fini à p^n éléments. Tout élément de k est solution de l'équation $X^{p^n} - X = 0$. Dans le corps k ce polynôme est donc produit de facteurs de degré 1 :

$$X^{p^n} - X = \prod_{\alpha \in k} (X - \alpha).$$

Démonstration. Soit k^* le groupe multiplicatif des éléments non nuls de k . L'ordre de tout élément divise $p^n - 1$ donc pour tout $x \in k$ non nul on a $x^{p^n-1} = 1$. Mais les solutions de l'équation $X^{p^n} - X = 0$ sont, d'un côté la solution nulle, d'un autre côté les solutions de $X^{p^n-1} = 1$.

En fin de compte, tout élément de k étant soit nul soit inversible est solution de $X^{p^n} - X = 0$.

Pour ce qui est de la formule $X^{p^n} - X = \prod_{\alpha \in k} (X - \alpha)$, on vient de montrer que le polynôme de droite divise le polynôme de gauche. Ils ont par ailleurs même degré, et le coefficient du terme de degré dominant est le même. La formule suit. ■

Éléments algébriques et transcendants

On va maintenant distinguer les éléments dans les extensions, selon qu'ils sont ou non solutions d'une équation polynomiale non triviale. On commence par définir la notion d'élément algébrique dans une extension :

Définition 2

Soit K une extension de k et soit x un élément de K . On dit que x est un *élément algébrique* sur k s'il existe un polynôme non nul $P \in k[X]$ dont x est racine.

On a :

Proposition 3. *Un élément $x \in K$ est algébrique sur k si et seulement si il existe un sous-corps L de K contenant k , de dimension finie sur k , dont x est élément.*

Avant de passer à la démonstration donnons des exemples : les éléments i , $\sqrt{2}$, les racines n -ièmes de l'unité sont algébriques sur $\mathbb{Q}$. On montre facilement que $a = \sqrt{3} + \sqrt{5}$ est algébrique sur $\mathbb{Q}$. En effet on écrit $(a - \sqrt{3})^2 = 5$. Soit $a^2 - 5 = 2\sqrt{3}a$, puis on élève au carré cette dernière équation. On obtient alors $a^4 - 22a^2 + 25 = 0$. Donc a est racine de $X^4 - 22X^2 + 25$. Ce cas est particulier, en général il est difficile de montrer directement que la somme de deux nombres algébriques est algébrique. Ceci est démontré en général plus loin.

Démonstration de la proposition 3. Supposons la condition de la définition satisfaite. Considérons alors l'homomorphisme

$$\begin{aligned}\phi : k[X] &\rightarrow K \\ \sum_i a_i X^i &\mapsto \sum_i a_i x^i.\end{aligned}$$

Si $R = \sum_i a_i X^i$, on écrira $\phi(R) = R(x)$. C'est un homomorphisme d'anneaux (proposition 2, section 5, chapitre 3), l'image de cet homomorphisme sera souvent notée $k[x]$, ce sont les polynômes en x , ou plus rigoureusement les valeurs prises par les polynômes à coefficients dans k en x .

L'anneau $k[X]$ étant principal le noyau de ϕ est constitué par les multiples d'un polynôme f , tout polynôme non nul de plus bas degré annulant x convient. Ce polynôme est uniquement défini si on lui impose d'être unitaire. Ce polynôme f est non nul car par hypothèse il existe un polynôme non nul annulant x . L'image de l'homomorphisme ϕ est isomorphe à l'anneau quotient $k[X]/(f)$. Elle est intègre car c'est un sous-anneau de K . L'idéal (f) est donc premier, et donc maximal car $k[X]$ est un anneau principal. On en conclut que le polynôme f est irréductible et que $k[X]/(f)$ est un corps. Ce corps est de dimension finie sur k , en notant d le degré du polynôme f on a :

$$[k[X]/(f) : k] = \dim_k k[X]/(f) = d.$$

■

Cette formule démontre que la première condition de la proposition implique la seconde partie de la proposition.

Avant de démontrer cette formule introduisons une définition :

Définition 3

Le *polynôme minimal* de x dans $k[X]$ est, par définition, le polynôme unitaire de plus petit degré annulant x .

Le polynôme minimal de x dans $k[X]$ est irréductible dans $k[X]$.

Revenons à la démonstration de la formule.

Soit $f = a_0 + a_1X + \dots + X^d$ le polynôme minimal de x . On va montrer que l'anneau quotient a même dimension sur k que l'espace vectoriel des polynômes de degré strictement inférieur à d .

Si on a un élément quelconque dans l'image de ϕ , il est de la forme $R(x)$, où R est un polynôme. Effectuons la division euclidienne du polynôme R par f on obtient $R = qf + T$, avec $0 \leq |T| < |f|$, d'où $R(x) = T(x)$. Ceci montre que $\{1, x, \dots, x^{d-1}\}$ est un système générateur de $k[X]/(f)$.

Inversement montrons que le système est libre. Si on a $b_0 + b_1x + \dots + b_hx^h = 0$, avec $b_h \neq 0$ et $h < d$, cela veut dire que x est zéro du polynôme non nul $b_0 + b_1X + \dots + b_hX^h$. Ce qui est impossible parce que $h < d$ et qu'il n'existe pas de polynôme non nul de degré $h < d$ annulant x .

Le système décrit est donc libre et est une base de $k[X]/(f)$.

Démontrons que la seconde condition de la définition implique la première. Soit m la dimension de L comme espace vectoriel sur k . L'ensemble $\{x^n \mid m \geq n \geq 0\}$ est lié car il comporte $m + 1$ éléments. On a donc au moins une relation de la forme $\sum_{i=1}^{i=m} a_i x^i = 0$ avec les a_i dans k et non tous nuls. Ce qui donne le résultat.

La définition implique le fait suivant qui est très utile. Soit K une extension de k , soit L un sous-corps de K contenant k , et soit $x \in K$ un élément algébrique sur k . Alors x est algébrique sur L . En effet s'il est zéro d'un polynôme non nul à coefficients dans k il l'est du même, considéré comme polynôme à coefficients dans L .

Introduisons quelques notations qui seront commodes dans la suite. La première l'a déjà été plus haut. Soit $x \in K$, on note $k[x]$ le sous-ensemble de K constitué par les éléments de la forme $P(x)$ où P est un polynôme à coefficients dans k . Il est facile de voir que c'est un sous-anneau car la somme et le produit de deux tels éléments sont encore du même type. De même on note $k(x)$ le plus petit sous-corps de K contenant k et x , c'est-à-dire le corps des fractions de $k[x]$. On a montré que :

Théorème 1. Si l'élément x est algébrique sur k alors $k[x]$ est égal à $k(x)$ et isomorphe à $k[X]/(P)$ où P est le polynôme minimal de x .

Soit x un élément qui n'est pas algébrique. Alors, l'homomorphisme qui à P associe $P(x)$ est injectif. On introduit la définition suivante :

Définition 4

On dit que $x \in K$ est un *élément transcendant* sur k s'il n'est zéro d'aucun polynôme non nul à coefficients dans k .

Si l'élément x est transcendant sur k alors $k[x]$ est isomorphe à $k[X]$ et $k(x)$ est isomorphe à $k(X)$.

Si on considère le corps des fractions rationnelles $k(X)$ la variable X est transcendante sur k . Les réels π et e sont transcendants sur $\mathbb{Q}$. Dans les exercices on donnera un exemple de nombre transcendant.

Le lemme suivant est très commode pour démontrer qu'un nombre est algébrique dans une grande variété de situations :

Lemme 1 (lemme de la base télescopique). Soit E une extension de k , et F une extension de E . On a la formule suivante :

$$[F : k] = [F : E][E : k].$$

Démonstration. On procède de la manière suivante. Considérons une base de E en tant qu'espace vectoriel sur k , soit $\{x_1, \dots, x_n\}$. Soit également $\{y_1, \dots, y_m\}$ une base de F en tant qu'espace vectoriel sur E . Nous allons montrer que l'ensemble $\{x_i y_j, 1 \leq i \leq n, 1 \leq j \leq m\}$ est une base de F sur k .

Il faut d'abord montrer que tout élément x de F a une écriture unique de la forme :

$$\sum_{(i,j)} a_{i,j} x_i y_j, \quad \text{avec } a_{i,j} \in k.$$

Pour montrer l'existence d'une décomposition on observe que x s'écrit $x = \sum_i b_i y_i$, où $b_i \in E$, les y_i constituant une base de F sur E . Par ailleurs pour tout i on a $b_i = \sum_j a_{i,j} x_j$, avec $a_{i,j} \in k$, les x_j constituant une base de E sur k . On a donc $x = \sum_{i,j} a_{i,j} x_j y_i$. Le résultat suit.

Pour ce qui est de l'unicité il suffit de montrer que si on a une relation

$$\sum_{(i,j)} a_{i,j} x_i y_j = 0, \quad \text{où } a_{i,j} \in k,$$

alors $a_{i,j} = 0$ pour tous (i, j) . Mais la relation s'écrit aussi

$$\sum_j \left(\sum_i a_{i,j} x_i \right) y_j = 0,$$

avec $\sum_i a_{i,j} x_i \in E$. Comme l'ensemble des y_j forme une base de F sur E on en déduit que $\sum_i a_{i,j} x_i$ est nul pour tous j .

Fixons j ; comme l'ensemble des x_i forme une base de E sur k , la nullité du coefficient $\sum_i a_{i,j} x_i \in E$ entraîne la nullité de $a_{i,j}$ pour tout i . D'où le résultat. ■

Voici une application sur laquelle on reviendra plus loin. Supposons donné un corps fini de caractéristique p , soit k , ayant p^n éléments et k' un sous-corps. Alors k' a p^d éléments pour un certain entier d . Montrons que l'entier d divise n . Les corps k et k' contiennent le corps premier $\mathbb{F}_p$. Ils sont de dimensions respectives n et d sur $\mathbb{F}_p$. Le lemme de la base télescopique montre alors que :

$$n = [k : k']d.$$

Théorème 2. Soit E une extension de k . L'ensemble des éléments de E algébriques sur k est un sous-corps de E .

Démonstration. Il suffit de montrer que si x et y sont algébriques sur k il en est de même pour $x + y$, xy , $-x$, et y^{-1} . On va montrer que tous ces éléments sont contenus dans une extension de k , contenue dans E , et de dimension finie sur k . C'est d'abord le cas pour x qui est par hypothèse algébrique sur k . Soit L , $k \subset L \subset E$ une telle extension. L'élément y est algébrique sur k , donc algébrique sur L .

Il existe donc une extension de L , contenue dans E , et de dimension finie sur L qui contient y . Mais alors l'extension M de k est de dimension finie sur k à cause du Lemme de la base télescopique (lemme 1, section 2) et contient tous les éléments considérés. Le résultat suit (définition 2, section 2). ■

Terminons par une définition.

Définition 5

On dit qu'une extension $k \subset K$ est une *extension algébrique* si tout élément $x \in K$ est algébrique sur k .

Il résulte de la définition 2 qu'une extension $k \subset K$, telle que K est de dimension finie sur k est algébrique.

3. Corps de rupture d'un polynôme

Soit $P \in k[X]$ un polynôme irréductible.

Théorème 1. Il existe une extension K de k unique à isomorphisme près telle que P ait au moins une racine α dans K et que le plus petit sous-corps de K contenant k et α soit K lui-même. Plus précisément si on a deux extensions K et K' de k avec des éléments α et α' vérifiant les hypothèses précédentes il existe un isomorphisme ϕ de K sur K' qui est égal à l'identité sur k et tel que $\phi(\alpha) = \alpha'$.

Démonstration. Pour ce qui est de l'existence on observe que l'anneau quotient $k[X]/(P)$ est un corps puisque P étant irréductible l'idéal qu'il engendre est maximal. La classe de X , notons la x , dans l'anneau quotient, est une racine de P .

Pour ce qui est de l'unicité il suffit de montrer que tout corps K satisfaisant aux conditions imposées est isomorphe à $k[X]/(P)$. Considérons l'application suivante :

$$k[X] \longrightarrow K, \quad R = \sum_i a_i X^i \longmapsto \sum_i a_i \alpha^i = R(\alpha).$$

Son noyau est engendré par le polynôme minimal de α qui n'est autre que P . Le quotient $k[X]/(P)$ s'injecte donc dans K , $k[X]/(P)$ est donc isomorphe à un sous-corps de K via un homomorphisme envoyant x sur α . Comme le plus petit sous-corps de K contenant α et k est K lui-même on en déduit que $k[X]/(P)$ est isomorphe à K . ■

Définition

Par définition ce corps est le *corps de rupture* du polynôme P .

Par exemple $\mathbb{C}$ est le corps de rupture du polynôme $X^2 + 1$ à coefficients dans $\mathbb{R}$.

Ce qui suit est une part de la définition du corps de décomposition d'un polynôme. En fait nous nous contentons de démontrer que pour tout polynôme à coefficients dans un corps k on peut trouver un corps K dans lequel P s'écrit comme produit de facteurs du premier degré. Nous ne discuterons pas l'unicité d'une telle extension.

Théorème 2. Soit k un corps et soit $P \in k[X]$. Il existe une extension K de k dans laquelle P est produit de facteurs du premier degré. (i.e. scindé).

Démonstration. La démonstration se fait par récurrence sur le degré du polynôme P . Le départ de la récurrence est immédiat. Soit d le degré du polynôme considéré. Supposons le théorème démontré pour tous les polynômes de degré strictement inférieur à d , à coefficients dans un corps quelconque. Soit alors f un facteur irréductible de P . Considérons le corps de rupture L de f , soit $k[X]/(f)$, notons α la classe de X dans L . Le polynôme P , considéré comme polynôme à coefficients dans L s'écrit $(X - \alpha)Q$. On peut donc appliquer l'hypothèse de récurrence au polynôme $Q \in L[X]$ et construire une extension M de L dans laquelle Q est produit de facteurs du premier degré. Mais M est une extension de k dans laquelle P est produit de facteurs du premier degré. ■

Voici un application de ce qui précède :

Proposition 1. Soit $k \subset K$ une extension de degré finie, et soit f un polynôme irréductible à coefficients dans k qui a au moins une racine dans K . Alors le degré de f divise le degré de l'extension.

Démonstration. En effet si α désigne la racine de f le degré du corps $k(\alpha)$ est celui de f . Par ailleurs $[k(\alpha) : k]$ divise $[K : k]$ par le théorème de la base télescopique. D'où le résultat. ■

4. Corps algébriquement clos, clôture algébrique

L'objet de cette section est de donner les définitions et des exemples de corps algébriquement clos et de clôtures algébriques. On ne démontrera pas en général l'existence de la clôture algébrique. Celle-ci dépend du lemme de Zorn et nous avons préféré insister sur les cas (très importants) où l'on peut justement s'en passer.

Définition 1

Un corps K est dit algébriquement clos si tout polynôme à coefficients dans K a au moins une racine dans K .

Voici des formulations équivalentes :

Proposition 1. On dit qu'un corps K est un corps algébriquement clos, si et seulement si il satisfait à l'une des propriétés équivalentes suivantes :

- tout polynôme à coefficient dans K est produit de facteurs de degré 1,
- tout polynôme irréductible est de degré 1,
- toute extension algébrique de K est égale à K .

Démonstration. La première condition de la proposition implique celle de la définition. Montrons que la définition l'implique. On le fait par récurrence sur le degré du polynôme $P \in K[X]$. Le départ de la récurrence est évident. On suppose le résultat établi pour les polynômes de degré inférieurs ou égaux à $n-1$. soit P un polynôme de degré n , P a au moins une racine dans K , et est donc produit d'un polynôme de degré 1 et d'un polynôme de degré $n-1$ auquel on peut appliquer l'hypothèse de récurrence, P est produit de facteurs de degré 1. D'un autre côté la première condition est équivalente au fait que tout polynôme irréductible est de degré 1. Montrons que la troisième implique la seconde. On raisonne par l'absurde, si P est irréductible de degré strictement supérieur à 1, alors $K[X]/P$ est une extension algébrique non-triviale de K (de degré strictement supérieur à 1). Donc il y a contradiction. Inversement si la seconde a lieu le polynôme minimal d'un élément dans une extension algébrique quelconque de K est de degré 1 et est donc dans K . ■

Le résultat suivant est fondamental :

Théorème 1 (Laplace-D'Alembert-Gauss). *Le corps $\mathbb{C}$ est algébriquement clos.*

La démonstration de ce résultat dépend des propriétés topologiques du corps des réels. Ce théorème est admis.

Un corps fini n'est jamais algébriquement clos. En effet soit k un corps fini, alors le polynôme à coefficients dans k

$$\prod_{\alpha \in k} (X - \alpha) + 1$$

n'a aucune racine dans k .

Tout corps est contenu dans un corps algébriquement clos. Introduisons une définition :

Définition 2

Soit k un corps, on appelle clôture algébrique de k une extension algébrique de k qui est algébriquement close.

Les clôtures algébriques existent, plus précisément on a le résultat suivant :

Théorème 2. *Tout corps k a une clôture algébrique K . Elle est unique à isomorphisme près en ce sens que si K' désigne une autre clôture algébrique il existe un isomorphisme $\phi : K \rightarrow K'$ qui est égal à l'identité quand on le restreint à k . La clôture algébrique de k est en général noté $\bar{k}$.*

Nous ne démontrerons pas ce théorème. Comme on l'a dit sa démonstration est une application du Lemme de Zorn. Mais nous allons montrer l'existence dans des cas particuliers sans utiliser le Lemme de Zorn.

Théorème 3. *Soit $\bar{\mathbb{Q}}$ l'ensemble des éléments de $\mathbb{C}$ qui sont algébriques sur $\mathbb{Q}$. alors $\bar{\mathbb{Q}}$ est une clôture algébrique de $\mathbb{Q}$.*

Démonstration. Le théorème 2 de la section 2 montre que c'est un sous-corps, en effet x et y sont algébriques sur $\mathbb{Q}$ il en est de même de $x - y$ et xy . Montrons qu'il est algébriquement clos. Soit $P(X) = \sum_{i=0, \dots, n} a_i X^i$ un polynôme irréductible à coefficients dans $\bar{\mathbb{Q}}$. Soit α un nombre complexe racine de ce polynôme. On va montrer que cet élément est dans $\bar{\mathbb{Q}}$. Définissons par récurrence un corps $k_i \subset \bar{\mathbb{Q}}$ par $k_0 = \mathbb{Q}(a_0)$, $k_i = k_{i-1}(a_i)$. Le corps k_n est de dimension finie sur $\mathbb{Q}$. En effet pour le démontrer il suffit, grâce au lemme de la base télescopique (section 2), de démontrer que pour tout i k_i est de dimension finie sur k_{i-1} . Ceci provient du fait que a_i est algébrique sur $\mathbb{Q}$ et donc sur k_i .

L'élément α est algébrique sur k_n , car son polynôme minimal est P qui est à coefficients dans k_n . Le corps $k_n(\alpha)$ est donc de dimension finie sur k_n . Donc par application du lemme de la base télescopique (section 2) le corps $k_n(\alpha)$ est de dimension finie sur $\mathbb{Q}$ donc α est algébrique sur $\mathbb{Q}$. Donc tout polynôme non constant à coefficients dans $\mathbb{Q}$ a au moins une racine dans $\overline{\mathbb{Q}}$. ■

Le même théorème reste vrai en substituant à $\mathbb{Q}$ un sous-corps $\mathbb{K}$ quelconque de $\mathbb{C}$:

Théorème 4. *Soit $\overline{\mathbb{K}}$ l'ensemble des éléments de $\mathbb{C}$ qui sont algébriques sur $\mathbb{K}$. alors $\overline{\mathbb{K}}$ est une clôture algébrique de $\mathbb{K}$.*

La démonstration est identique. Dans les exercices on décrira la clôture algébrique des corps finis.

5. Structure et classification des corps finis

Commençons par observer qu'un corps fini est nécessairement de caractéristique non nulle, faute de quoi il contiendrait $\mathbb{Q}$ qui n'est pas fini ! Soit p cette caractéristique. Les corps finis sont des objets mathématiques intéressants pour eux-mêmes et qui ont des applications concrètes : les codes correcteurs d'erreur. Le but de cette section est de les décrire tous, et d'établir quelques unes de leurs propriétés.

Voici le résultat fondamental de cette section. Il sera précisé par les énoncés ultérieurs :

Théorème 1. *Soit p un nombre premier, et soit $n \geq 1$ un entier. Alors :*

- soient $\mathbb{F}$ et $\mathbb{F}'$ deux corps à p^n éléments, alors $\mathbb{F}$ est isomorphe à $\mathbb{F}'$,
- il existe (au moins) un corps à p^n éléments.

Il s'agit là d'un théorème d'existence et d'unicité. Nous donnerons deux démonstrations de l'existence. L'une dépendra du théorème 2 de la section 3, l'autre de l'existence d'un polynôme irréductible de degré n .

Commençons par appliquer les théorèmes généraux à notre situation.

Proposition 1. *Soit P un polynôme irréductible de degré n sur $\mathbb{F}_p$. Alors le corps de rupture de P , soit $L = \mathbb{F}_p[X]/(P)$, est constitué par l'ensemble des solutions de l'équation $X^{p^n} - X = 0$. C'est un corps à p^n éléments. Enfin le polynôme P divise le polynôme $X^{p^n} - X$.*

Démonstration. La première partie est une reformulation immédiate du corollaire 2 de la section 2. La seconde partie se démontre comme suit. Soit x la classe de X dans L , c'est une racine de P , le polynôme minimal de $x \in L$ sur $\mathbb{F}_p$ est P . En effet il est irréductible et divisible par le polynôme minimal sur $\mathbb{F}_p$ de x .

Par ailleurs on vient de voir que x annule le polynôme $X^{p^n} - X$. Donc par définition, le polynôme minimal de l'élément x , soit P divise $X^{p^n} - X$.

Enfin il est de degré n sur $\mathbb{F}_p$. ■

À titre d'exemple prenons le polynôme $X^2 + X + 1 \in \mathbb{F}_2[X]$, il est irréductible car il n'a pas de racines dans $\mathbb{F}_2$. L'anneau quotient $\mathbb{F}_2[X]/(X^2 + X + 1)$ est donc un corps à 4 éléments. De même on peut considérer le polynôme $X^3 + X + 1$, il est aussi irréductible pour la même raison, il n'a pas de racines dans $\mathbb{F}_2$. L'anneau quotient $\mathbb{F}_2[X]/(X^3 + X + 1)$ est donc un corps à 8 éléments.

On peut aussi prendre le polynôme $X^3 + 2X + 1 \in \mathbb{F}_3[X]$. On vérifie qu'il est irréductible pour la même raison : il n'a pas de racines dans $\mathbb{F}_3$. L'anneau quotient $\mathbb{F}_3[X]/(X^3 + 2X + 1)$ est donc un corps à 27 éléments.

Existence des corps finis

Passons maintenant à la question de l'existence d'un corps à p^n éléments.

Comme on ne sait pas s'il existe un polynôme irréductible sur $\mathbb{F}_p$ de degré n pour un n quelconque, on ne peut utiliser cet argument pour prouver l'existence d'un corps à p^n éléments. Ainsi qu'on l'a dit plus haut on va donner deux démonstrations indépendantes.

Dans la première on démontre l'existence d'un corps à p^n éléments à l'aide du théorème 2 de la section 3 :

Proposition 2. Soit P un polynôme quelconque à coefficients dans $\mathbb{F}_p$; il existe une extension k de $\mathbb{F}_p$ dans laquelle P est produit de facteurs du premier degré. Soit P le polynôme $X^{p^n} - X$, et soit K une telle extension. Alors :

- l'ensemble des solutions dans K de l'équation $X^{p^n} - X = 0$ est un sous-corps L à p^n éléments,
- tout polynôme irréductible de degré n à coefficients dans $\mathbb{F}_p$ est produit de facteurs du premier degré comme polynôme à coefficients dans L .

Démonstration. L'existence de l'extension est une application du théorème 2 de la section 3. Considérons la seconde assertion, par hypothèse le polynôme $X^{p^n} - X$ est, en tant que polynôme à coefficients dans K , produit de facteurs du premier degré.

L'ensemble L est par définition l'ensemble des solutions de l'équation $X^{p^n} - X = 0$, c'est un sous-corps de K .

En effet, si a et b sont deux solutions on vérifie que $a + b$, ab , $-a$, et a^{-1} sont également solutions. Faisons le dans le cas de $a + b$, on a :

$$(a + b)^{p^n} = a^{p^n} + b^{p^n} = a + b.$$

C'est bien un sous-corps et puisque les racines de $X^{p^n} - X$ sont deux à deux distinctes, car la dérivée formelle de $X^{p^n} - X$ est -1 , il a p^n éléments.

Considérons la dernière assertion. Soit P un polynôme à coefficients dans $\mathbb{F}_p$, irréductible et de degré n . Le théorème 1 implique qu'il divise $X^{p^n} - X$. Le polynôme $X^{p^n} - X$ a toutes ses racines dans L , il en est donc de même pour P . ■

Avant de passer à la seconde approche on va énoncer des résultats intermédiaires.

Proposition 3. *Soit P un polynôme irréductible à coefficients dans $\mathbb{F}_p$ de degré d . Alors P divise le polynôme $X^{p^n} - X = 0$ si et seulement si d divise n .*

Démonstration. On commence par montrer que $X^{p^d} - X$ divise $X^{p^n} - X$ si et seulement si d divise n .

En fait on montre un peu plus, à savoir que le pgcd de $X^{p^m} - X$ et $X^{p^n} - X$ est $X^{p^h} - X$, où h est le pgcd de m et n . Ceci implique le résultat annoncé.

Posons $m = an + r$, avec $0 \leq r < n$. Alors

$$p^m - 1 = (p^n - 1)(p^{(a-1)n} + \dots + 1)p^r + p^r - 1,$$

posons $\lambda = p^{(a-1)n} + \dots + 1$

$$X^{p^m-1} - 1 = X^{p^r-1}(X^{(p^n-1)\lambda p^r} - 1) + X^{p^r-1} - 1,$$

soit

$$X^{p^m-1} - 1 = X^{p^r-1}(X^{p^{n-1}} - 1)((X^{p^{n-1}})^{\lambda p^r-1} + \dots + 1) + X^{p^r-1} - 1.$$

Il résulte de cette formule que le pgcd de $X^{p^m} - X$ et $X^{p^n} - X$ est le même que celui de $X^{p^n} - X$ et $X^{p^r} - X$. Le résultat suit par itération comme dans l'algorithme d'Euclide.

Alors si d divise n par application de la proposition 1, P divise $X^{p^d} - X$ et donc P divise $X^{p^n} - X$.

Considérons la réciproque, soit donc d le degré du polynôme considéré. Le polynôme P divise donc $X^{p^n} - X$ et $X^{p^d} - X$, donc leur pgcd $X^{p^r} - X$. On conclut alors en faisant un raisonnement par récurrence. Supposons la propriété démontrée pour tous les entiers inférieurs à n . Si $d = n$ il n'y a rien à démontrer. Si $d < n$, alors $r < n$ et par récurrence $d \mid r$, donc $d \mid n$. ■

La réciproque résulte aussi de l'argument suivant où on utilise le théorème 2 de la section 3. Soit un polynôme irréductible P divisant $X^{p^n} - X$. Il a une racine dans le corps $\mathbb{F}$ à p^n éléments, il y est même produit de facteurs de degré 1. Donc $\mathbb{F}$ contient un corps de rupture L pour P . Le cardinal de $\mathbb{F}$ soit p^n est donc une puissance du cardinal de L qui est $p^{|P|}$, on en déduit que $|P|$ divise n .

Dans la suite on notera $\mathcal{J}_d$ l'ensemble des polynômes unitaires irréductibles à coefficients dans $\mathbb{F}_p$ de degré d .

Théorème 2. *On a la formule suivante :*

$$X^{p^n} - X = \prod_{d|n} \left(\prod_{f \in \mathcal{J}_d} f \right).$$

Démonstration. Décomposons $X^{p^n} - X$ en facteurs irréductibles dans $\mathbb{F}_p[X]$. D'après le théorème 3 précédent apparaissent tous les polynômes irréductibles unitaires de degré d divisant n . Ils n'apparaissent qu'une fois car la dérivée formelle de $X^{p^n} - X$ est -1 . ■

En comparant les degrés des deux membres on obtient aussitôt :

Corollaire. Notons $\psi_p(n) = \# \mathcal{J}_n$. On a alors :

$$p^n = \sum_{d|n} d \psi_p(d).$$

Lemme. La fonction $\psi_p(n)$ est non nulle pour tout n .

Démonstration. Il suffit de montrer que $n\psi_p(n)$ est non nul pour tout n . La relation ci-dessus montre que $n\psi_p(n) \leq p^n$. Donc on a

$$n\psi_p(n) \geq p^n - \sum_{d|n, d \neq n} p^d,$$

l'entier n a au plus $\frac{n}{2}$ diviseurs : c'est un sous-ensemble de $\{1, \dots, [\frac{n}{2}]\}$. Pour les valeurs d appartenant à cet intervalle la fonction p^d est inférieure à $p^{\frac{n}{2}}$. Il suffit donc de montrer que $p^n - (\frac{n}{2})p^{\frac{n}{2}} > 0$. Ce qu'on vérifie facilement. ■

Le résultat précédent donne une autre démonstration de l'existence d'un corps fini à p^n éléments. En effet on vient de montrer qu'il existe au moins un polynôme irréductible de degré n .

Unicité des corps finis

Voici en corollaire de ce qui précède un énoncé plus précis sur l'unicité à isomorphisme près des corps finis :

Théorème 3. Soient L et L' deux corps finis, de caractéristique p , ayant le même nombre d'éléments p^n , alors L et L' sont isomorphes. De plus, si f est un polynôme irréductible de degré n à coefficients dans $\mathbb{F}_p$, alors ils sont isomorphes au corps $\mathbb{F}_p[X]/(f)$.

Autrement dit, étant donnée une puissance d'un nombre premier p , soit p^n , il y a un corps et un seul, à isomorphisme près, ayant p^n éléments.

Le corps fini à p^n éléments sera noté $\mathbb{F}_{p^n}$, respectivement si la puissance de p est écrite sous la forme q le corps fini à q éléments sera noté $\mathbb{F}_q$. La terminologie « le corps fini à p^n éléments » est bien entendue abusive et doit être comprise comme il est précisé plus haut.

Démonstration. Soit un polynôme f irréductible de degré n , à coefficients dans $\mathbb{F}_p$, f divise le polynôme $X^{p^n} - X$ et a donc (corollaire, section 2) une racine α dans L . Donc l'homomorphisme d'anneaux de $\mathbb{F}_p[X]$ dans L qui à P associe $P(\alpha)$ a pour noyau l'idéal engendré par f . On a donc un homomorphisme de corps de $\mathbb{F}_p[X]/(f)$ dans L . Les cardinaux de ces deux corps sont égaux, c'est donc un isomorphisme. On fait le même travail avec L' et on en conclut que L et L' étant tous les deux isomorphes à un même corps sont isomorphes. ■

Nous concluons par l'étude des automorphismes du corps fini $\mathbb{F}_{p^n}$. On a :

Proposition 4. L'application F de $\mathbb{F}_{p^n}$ dans lui-même qui à l'élément x associe x^p est un automorphisme de $\mathbb{F}_{p^n}$. L'ensemble des automorphismes de $\mathbb{F}_{p^n}$ est un groupe cyclique d'ordre n dont F est un générateur.

Démonstration. Le fait que F est un homomorphisme est une conséquence de la formule de Newton en caractéristique p . C'est un homomorphisme injectif. En effet le seul élément tel que $F(\alpha) = 1$ est l'unité car le polynôme $X^p + 1$ est égal à $(X + 1)^p$. Donc F est surjectif car c'est une application injective d'un ensemble fini dans lui-même. C'est donc un automorphisme. L'ensemble des automorphismes est un sous-groupe du groupe des permutations de $\mathbb{F}_{p^n}$ pour la composition des applications. Notons maintenant que l'ordre de F divise n , c'est une reformulation de la généralisation du théorème de Fermat car on a $F^n(x) = x^{p^n}$. L'ordre de F est exactement n . En effet on ne peut avoir pour tout $x \in \mathbb{F}_{p^n}$ l'égalité $F^d(x) = x$, c'est-à-dire $x^{p^d} = x$ que si $d = n$. Car l'équation $X^{p^d} - X = 0$ a au plus p^d solutions et il y a p^n éléments dans $\mathbb{F}_{p^n}$.

Pour achever la démonstration on montre que l'ensemble des automorphismes de $\mathbb{F}_{p^n}$ a au plus n éléments. Comme il y en a déjà n le résultat suit. On procède

comme suit. Soit P un polynôme irréductible de degré n , soit α une racine de P dans $\mathbb{F}_{p^n}$, donc P est polynôme minimal de α , soit enfin σ un automorphisme de $\mathbb{F}_{p^n}$. Le polynôme minimal de l'élément $\sigma(\alpha)$ est également P . Il suffit de montrer que $\sigma(\alpha)$ est racine de P . Or si

$$P = \sum_i a_i X^i,$$

avec $a_i \in \mathbb{F}_p$, on a

$$\sum_i a_i \alpha^i = 0.$$

Donc on a

$$\sigma(\sum_i a_i \alpha^i) = 0,$$

donc

$$\sum_i \sigma(a_i) \sigma(\alpha)^i = 0.$$

Mais on a $\sigma(u) = u$ pour tout $u \in \mathbb{F}_p$, car $\sigma(1) = 1$, et il en est de même pour tous les éléments de $\mathbb{F}_p$ qui sont de la forme $\underbrace{1 + \dots + 1}_{m \text{ fois}}$. En conclusion on a

$$\sum_i a_i \sigma(\alpha)^i = 0.$$

soit $P(\sigma(\alpha)) = 0$. Étant donné un automorphisme σ , l'élément $\sigma(\alpha)$ est donc racine de l'équation $P = 0$. Il ne peut donc prendre au plus que n valeurs, les racines de l'équation. On vient donc de construire une application du groupe des automorphismes vers un ensemble à n éléments. On va montrer que cette application est injective.

Pour cela on observe qu'un automorphisme σ est déterminé par sa valeur sur α . En effet les éléments de $1, \alpha, \dots, \alpha^{n-1}$ constituent une base de l'espace vectoriel $\mathbb{F}_{p^n}$ sur $\mathbb{F}_p$, donc si on connaît la valeur de σ en α on la connaît sur toute puissance de α . Mais un élément $x \in \mathbb{F}_{p^n}$ s'écrit $x = a_0 + a_1 \alpha + \dots + a_{n-1} \alpha^{n-1}$, $a_i \in \mathbb{F}_p$. Et donc on a $\sigma(x) = a_0 + a_1 \sigma(\alpha) + \dots + a_{n-1} \sigma(\alpha)^{n-1}$. Si on a donc σ et σ' tels que $\sigma(\alpha) = \sigma'(\alpha)$ on a $\sigma(x) = \sigma'(x)$ pour tout x . L'application est donc injective et le groupe des automorphismes de $\mathbb{F}_{p^n}$ a au plus n éléments. ■

6. Applications arithmétiques

On va dans cette section donner quelques applications arithmétiques de ce qui précède.

Proposition 1. Soit p un nombre premier impair. Un élément x dans $\mathbb{F}_p^*$ est un carré si et seulement si $x^{\frac{p-1}{2}} = 1$.

Démonstration. L'application de $\mathbb{F}_p^*$ dans lui-même qui à x associe x^2 est un homomorphisme de groupes. Son noyau est constitué par les éléments 1 et -1 , car il est constitué par les solutions de l'équation $X^2 - 1 = 0$. Son image consiste en l'ensemble des éléments de $\mathbb{F}_p^*$ qui sont des carrés, elle a donc $\frac{p-1}{2}$ éléments. Cette image est contenue dans l'ensemble des solutions de l'équation $X^{\frac{p-1}{2}} - 1 = 0$. En effet, un élément de la forme y^2 est tel que $(y^2)^{\frac{p-1}{2}} = y^{p-1} = 1$. Mais l'équation $X^{\frac{p-1}{2}} - 1 = 0$ a au plus $\frac{p-1}{2}$ solutions, puisqu'il y a $\frac{p-1}{2}$ carrés elle en a exactement $\frac{p-1}{2}$, et donc les carrés sont l'ensemble des solutions de cette équation. ■

En particulier on en conclut que :

Corollaire. Soit p un nombre premier différent de 2. L'élément -1 est un carré modulo p , si et seulement si p est congru à 1 modulo 4.

Démonstration. Un nombre premier impair est soit de la forme $4k - 1$, soit de la forme $4k + 1$. Il suffit alors d'appliquer le critère, $(-1)^{\frac{p-1}{2}}$ est égal à 1 modulo p si et seulement si p est de la forme $4k + 1$. ■

Voici une conséquence :

Théorème. Il existe une infinité de nombres premiers de la forme $4k + 1$.

Démonstration. On raisonne par l'absurde. Supposons qu'il y ait un nombre fini de nombres premiers de la forme considérée, soient $p_1, \dots, p_n$. Considérons alors l'entier $N = 4(p_1 \cdots p_n)^2 + 1$. Soit p un nombre premier qui divise cet entier, p est nécessairement impair et distinct de $p_1, \dots, p_n$. Si on réduit modulo p on constate que -1 est égal au carré de $2p_1 \cdots p_n$, donc p est congru à 1 modulo 4. ■

Ce résultat est un cas particulier du théorème de Dirichlet qui affirme que si a et b sont deux entiers premiers entre eux il existe une infinité de nombres premiers de la forme $a + kb$. Ce théorème se démontre par des méthodes entièrement différentes.

Donnons rapidement à titre d'exemple à quelle condition sur p (p premier impair) 2 est un carré modulo p . Soit K une extension de $\mathbb{F}_p$ qui contient une racine de l'équation $X^4 + 1 = 0$, soit ζ une telle racine. Considérons alors $x = \zeta + \zeta^{-1}$. On a $x^2 = 2 + \zeta^2 + \zeta^{-2} = 2 + \zeta^{-2}(\zeta^4 + 1) = 2$. On doit déterminer à quelle condition l'élément x est dans le corps $\mathbb{F}_p$. Ceci a lieu si et seulement si $x^p = x$. Or un nombre premier impair est de la forme $8k + 1$, $8k - 1$, $8k + 3$ ou $8k - 3$. On a dans le premier cas $x^p = \zeta^p + \zeta^{-p} = \zeta^{8k} \zeta + \zeta^{-8k} \zeta^{-1} = (-1)^{2k} \zeta + (-1)^{-2k} \zeta^{-1} = x$. On vérifie de manière analogue que la même équation a lieu si $p = 8k - 1$. Par contre on vérifie qu'il n'en est pas ainsi dans les deux autres cas.

7. Racines de l'unité, polynômes cyclotomiques

Dans cette section nous allons étudier les racines de l'unité dans un corps. On introduira les *polynômes cyclotomiques*, leur irréductibilité.

Définition 1

Soit ζ un nombre complexe, on dit que c'est une *racine n -ième de l'unité* s'il existe un entier n tel que $\zeta^n = 1$. On note μ_n l'ensemble des racines n -ièmes de l'unité.

Cette définition s'étend évidemment au cas d'un corps quelconque.

Proposition 1. *L'ensemble μ_n est un groupe pour la multiplication. Il a n éléments et est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.*

Démonstration. Pour ce qui est de la première assertion on observe que si $\zeta^n = 1$ et $\zeta^m = 1$ on a $\zeta^n \zeta^{m-n} = 1$, donc μ_n est bien un groupe pour la multiplication. Pour ce qui est de la seconde assertion l'ensemble des racines n -ièmes de l'unité est l'ensemble des solutions complexes du polynôme $X^n - 1$. Comme ce polynôme n'a pas de zéros communs avec son polynôme dérivé nX^{n-1} l'équation $X^n - 1 = 0$ a n zéros dans $\mathbb{C}$. L'élément $e^{\frac{2\pi i}{n}}$ est dans μ_n et est d'ordre n . On en conclut que le groupe est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. ■

On observera que si d divise n , μ_d est un sous-groupe de μ_n .

Définition 2

On appelle *racine primitive n -ième de l'unité* un élément de μ_n qui est d'ordre n , c'est-à-dire un générateur de μ_n . On note μ_n^* l'ensemble des racines primitives n -ièmes de l'unité, cet ensemble est de cardinal $\varphi(n)$, où on rappelle que φ est la fonction d'Euler.

Définition 3

On appelle *polynôme cyclotomique* Φ_n le polynôme

$$\prod_{\zeta \in \mu_n^*} (X - \zeta),$$

il est de degré $\varphi(n)$.

Proposition 2. *Le polynôme cyclotomique Φ_n est à coefficients dans $\mathbb{Z}$.*

Démonstration. On commence par observer que l'on a la formule suivante :

$$X^n - 1 = \prod_{d|n} \Phi_d.$$

Les ensembles μ_d^* , pour $d|n$, constituent une partition de $\{1, \dots, n\}$. En effet, soit une racine n -ième de l'unité d'ordre d dans le groupe μ_n , cet ordre divise n . Par définition elle appartient à μ_d^* , si et seulement si son ordre est d . On a donc :

$$X^n - 1 = \prod_{d|n} \left(\prod_{\zeta \in \mu_d^*} (X - \zeta) \right),$$

la formule annoncée en résulte.

La formule utilisée dans la démonstration mérite d'être énoncée à part :

Proposition 3. *On a pour tout entier n :*

$$X^n - 1 = \prod_{d|n} \Phi_d.$$

On remarquera que si on calcule les degrés des polynômes de chaque côté on retrouve la formule :

$$n = \sum_{d|n} \varphi(d).$$

Terminons maintenant la démonstration de la proposition 2. On procède par récurrence sur n . On suppose donc le résultat établi pour les polynômes cyclotomiques Φ_d , avec $d < n$. On observe aussi que tous les polynômes cyclotomiques sont unitaires. On a donc une relation : $X^n - 1 = \Phi_n \times R$, où $R \in \mathbb{Z}[X]$ et R est unitaire. On peut alors (théorème 2, section 5, chapitre 3) faire la division dans $\mathbb{Z}[X]$, le résultat suit. ■

Donnons quelques exemples. Soit p un nombre premier, on a :

$$\Phi_p(X) = X^{p-1} + X^{p-2} + \dots + 1,$$

car on a $X^p - 1 = \Phi_p(X)(X - 1)$. Ce polynôme est irréductible comme polynôme à coefficients dans $\mathbb{Z}$ par application du critère d'Eisenstein effectuée en substituant $X + 1$ à X dans le polynôme (proposition 3, section 7, chapitre 3).

On vérifiera en exercice que :

$$\Phi_4(X) = X^2 + 1,$$

et

$$\Phi_6(X) = X^2 - X + 1$$

ainsi que

$$\Phi_{12}(X) = X^4 - X^2 + 1.$$

On vérifie aisément qu'ils sont irréductibles. On renvoie aux exercices pour d'autres calculs.

La démonstration du théorème suivant est sensiblement plus difficile. Elle est faite dans les exercices.

Théorème 1. *Le polynôme cyclotomique Φ_n est irréductible dans $\mathbb{Z}[X]$.*

En voici une application :

Corollaire. *Soit ζ une racine primitive n -ième de l'unité. Le corps $\mathbb{Q}(\zeta)$ est de degré $\varphi(n)$ sur $\mathbb{Q}$.*

Cette extension est appelée l'extension cyclotomique de degré n .

Pour terminer cette section nous allons faire une brève incursion du côté des racines de l'unité dans un corps fini. Observons d'abord que tout élément non nul dans un corps fini k est une racine de l'unité. En effet on sait que si k est un corps fini à q éléments pour tout $x \in k^*$ on a $x^{q-1} = 1$.

On peut définir comme il a été fait pour $\mathbb{C}$ des racines primitives. Nous n'allons pas faire une étude détaillée. Contentons nous de mentionner (exercice) que si l'on réduit modulo p un polynôme cyclotomique sa réduction n'est pas en général irréductible. Nous donnerons aussi une de leur propriété comme corollaire du théorème qui suit.

Théorème 2. *Le groupe multiplicatif d'un corps fini est cyclique.*

Démonstration. Soit $\mathbb{F}$ le corps fini, soit q l'ordre de $\mathbb{F}$, et soit $\mathbb{F}^*$ son groupe multiplicatif. Donc $q - 1$ est le cardinal de $\mathbb{F}^*$. Tout élément $x \in \mathbb{F}^*$ satisfait à $x^{q-1} = 1$ car son ordre divise $q - 1$. Il faut montrer qu'il y a un élément d'ordre $q - 1$. Soit d un diviseur de $q - 1$, les éléments d'ordre d sont solutions de l'équation $X^d - 1 = 0$. Soit cette équation a une solution qui est un élément d'ordre d . Dans ce cas l'ensemble des solutions de l'équation constitue un sous-groupe multiplicatif qui est isomorphe à $\mathbb{Z}/d\mathbb{Z}$ et il y a autant d'éléments d'ordre d que de générateurs, soit $\varphi(d)$, où φ est la fonction d'Euler. Soit il n'y a pas d'éléments d'ordre d .

Soit maintenant C_l l'ensemble des éléments d'ordre l . Les sous-ensembles C_l , pour l divisant n , constituent une partition de $\mathbb{F}^*$. On a donc

$$q - 1 = \sum_{l|q-1} \# C_l.$$

On vient de montrer que le cardinal de C_l est 0 ou $\varphi(l)$ (proposition 2, section 3, chapitre 3). Par ailleurs on a la formule classique

$$q - 1 = \sum_{d|q-1} \varphi(d).$$

En comparant les deux formules, et puisque $\#C_d \leq \varphi(d)$ on en déduit que $\#C_d = \varphi(d)$, donc que C_{q-1} est non vide et qu'il existe un élément d'ordre $q - 1$.

Pour conclure faisons l'observation suivante. Soit $q = p^n$, et soit x un générateur de $\mathbb{F}^*$. Le polynôme minimal de x sur $\mathbb{F}_p$ est de degré n . En effet, soit k le degré de ce polynôme minimal, le corps $\mathbb{F}_p(x)$ contient $p^k = p^n$ éléments, par hypothèse sur x , donc $k = n$. ■

Voici un exemple, considérons le corps $\mathbb{F}_{16}$. Nous pouvons le représenter comme étant le quotient $\mathbb{F}_2[X]/(X^4 + X + 1)$. Notons α la classe de X . Le groupe $(\mathbb{F}_2[X]/(X^4 + X + 1))^*$ a 15 éléments, il nous suffit donc de trouver dans ce groupe un élément, distinct de 0 et 1, qui ne soit pas d'ordre 5 ou 3, il sera forcément d'ordre 15. Le corps $\mathbb{F}_2[X]/(X^4 + X + 1)$ admet pour base sur $\mathbb{F}_2$ les éléments $1, \alpha, \alpha^2, \alpha^3$. L'élément α n'est pas d'ordre 3 et comme $\alpha^5 = \alpha^2 + \alpha$ il n'est pas d'ordre 5. Il convient donc.

Le lecteur vérifiera que le théorème reste vrai si on remplace k^* par un sous-groupe fini d'un corps commutatif quelconque.

Notons enfin :

Corollaire. *Soit k un corps fini à q éléments. Alors k contient une racine primitive n -ième de 1 si et seulement si n divise $q - 1$.*

Démonstration. En effet, il suffit de se demander quand le groupe multiplicatif k^* contient un élément d'ordre n . ■

8. Théorème de Wedderburn

L'objet de cette section est de démontrer le théorème suivant :

Théorème (Wedderburn). *Un corps fini est commutatif et donc isomorphe à $\mathbb{F}_{p^n}$ pour un nombre premier p et un entier n bien déterminés.*

Démonstration. On commence par examiner ce qui reste vrai dans les résultats démontrés dans ce chapitre si on enlève l'hypothèse de commutativité faite au début. Spécifiquement on vérifie sans peine qu'un corps k non-commutatif est un espace vectoriel sur un sous-corps quelconque k' et que par conséquent le cardinal de k est une puissance de celui de k' .

Soit alors k un corps fini et soit $Z(k)$ le sous-ensemble de k défini par

$$Z(k) = \{x \in k \mid xy = yx, \forall y \in k\}.$$

C'est un sous-corps commutatif de k . Soit q le cardinal de $Z(k)$.

Le corps k est donc un espace vectoriel sur $Z(k)$ de dimension finie n . Il a donc q^n éléments pour un certain entier n . On veut en fait montrer que $n = 1$.

Soit maintenant un élément $x \in k$ non nul et soit

$$k_x = \{y \in k \mid xy = yx\}$$

l'ensemble des éléments de k qui commutent avec x . C'est un sous-corps de k . Il contient par définition $Z(k)$ car un élément qui commute avec tous les éléments du corps commute avec un élément donné. Le corps k_x est donc un espace vectoriel sur $Z(k)$ de dimension finie n_x . Le cardinal de k_x est donc de la forme q^{n_x} . L'entier n_x divise n par le même type d'argument que celui employé plus haut : le lemme de la base télescopique s'applique.

Précisons, le corps k est un espace vectoriel sur k_x , soit d_x sa dimension sur k_x . Le nombre d'éléments de k s'écrit donc d'une part q^n et d'autre part $(q^{n_x})^{d_x}$. On a donc en écrivant l'égalité de ces expressions $n = n_x d_x$. Notons que n_x est distinct de n .

Pour éviter les espaces vectoriels sur un corps non commutatif on peut aussi raisonner comme suit. On observe à l'aide du théorème de Lagrange que $q^{n_x} - 1$ divise $q^n - 1$, puis on conclut (proposition 3, section 5).

On va maintenant écrire l'équation des classes pour le groupe multiplicatif k^* . Le centre de k^* est égal à $Z(k)^*$. Le centralisateur d'un élément $x \in k^*$ est égal à k_x^* . Leur nombre d'éléments sont respectivement $q - 1$ et $q^{n_x} - 1$. On a donc :

$$q^n - 1 = q - 1 + \sum_x \frac{q^n - 1}{q^{n_x} - 1},$$

où la somme dans le terme de droite est prise sur un système de représentants des classes de conjugaison de k^* n'appartenant pas au centre. Considérons la formule de la proposition 3 de la section 3, et évaluons la en q . On en déduit

$$q^n - 1 = \prod_{d \mid n} \Phi_d(q).$$

Les termes de droite sont des entiers.

L'entier $\Phi_n(q)$ divise $q^n - 1$. Il divise chaque terme dans la somme du membre de droite de l'équation. En effet chacun de ces termes est un produit de la forme $\prod_{d \mid n, d \neq n_x} \Phi_d(q)$ et $n_x \neq n$ par hypothèse. Donc $\Phi_n(q)$ apparaît dans ce produit d'entiers.

L'entier $\Phi_n(q)$ divise donc $q - 1$. Ceci n'est possible que si $n = 1$. En effet pour une racine ζ de l'unité distincte de 1 on a $|q - \zeta| > |q - 1| \geq 1$. On en déduit que si $n > 1$ on a $|\Phi_n(q)| > |q - 1|$ si $n > 1$, ce qui est impossible puisque l'entier $\Phi_n(q)$ divise $q - 1$. Donc on a $n = 1$ ce qui veut dire que $k = Z(k)$ et donc que k est commutatif. ■

9. L'algorithme de Berlekamp

Dans cette section, nous présentons un algorithme, dit de Berlekamp, de décomposition des polynômes en produit de facteurs irréductibles. Il est bien clair que le problème de factoriser un polynôme sur un corps fini est fini, comme celui de factoriser un entier. Il suffit en effet par récurrence de donner la liste de tous les polynômes irréductibles de degré inférieur ou égal. Mais nous allons présenter un algorithme plus rapide, pouvant être programmé efficacement sous certaines hypothèses. Par ailleurs, cet algorithme est conceptuellement intéressant. Il se présente en fait comme une succession d'applications des techniques déjà étudiées.

Pour cette section et la suivante, on trouvera une masse énorme d'informations supplémentaires dans [De], [MS], et tout spécialement à [LN] dont on s'est fortement inspiré.

Soit $k = \mathbb{F}_q$ un corps fini de caractéristique p , et soit P un polynôme à coefficients dans k . Soit

$$P = P_1^{\alpha_1} \dots P_\ell^{\alpha_\ell}$$

sa décomposition en produit de facteurs irréductibles.

L'algorithme qui va être décrit permet de décomposer un polynôme en produit de puissances de polynômes irréductibles. Mais il ne permet pas de décomposer une puissance d'un polynôme irréductible. Cette dernière opération peut se faire à l'aide de l'algorithme d'Euclide. Cependant il est plus économique de partir d'emblée d'un polynôme dont les facteurs irréductibles sont simples. C'est-à-dire au cas où dans la décomposition ci-dessus tous les α_i sont égaux à 1.

Ceci est obtenu en remplaçant le problème initial par celui de la factorisation d'une famille de polynômes. Deux cas peuvent se présenter :

- Soit le polynôme P est une puissance p -ième. C'est-à-dire qu'il est de la forme $Q(X^p)$ et la factorisation de P se ramène à celle de Q qui est de degré $|P|/p$. On est ramené à factoriser Q .
- Soit que le polynôme P n'est pas une puissance p -ième, auquel cas le polynôme dérivé P' est non nul. On calcule alors le pgcd D_0 de P et P' par l'algorithme d'Euclide. Tous les facteurs irréductibles du polynôme $Q_0 = P/D_0$ sont simples. L'algorithme décrit ci-dessous s'appliquera donc à Q_0 .

Il reste alors à factoriser D_0 et on recommence ce qui a été fait plus haut pour P . On détermine d'abord si D_0 est ou n'est pas une puissance p -ième. Si D_0 est une puissance p -ième, c'est-à-dire de la forme $R(X^p)$ on analyse R . Sinon on calcule le pgcd D_1 du polynôme D_0 et de son polynôme dérivé D'_0 . Tous les facteurs irréductibles de la décomposition de $Q_1 = D_0/D_1$ sont simples et on lui applique l'algorithme ci-dessous. Il reste alors à factoriser D_1 . Comme le degré des polynômes considérés décroît strictement le processus s'achève après un nombre fini d'itérations.

On doit donc décomposer un polynôme dont tous les facteurs irréductibles sont simples. Ainsi qu'on l'a dit, l'algorithme donne en fait la décomposition d'un polynôme quelconque en puissances de polynômes irréductibles.

Voici le résultat essentiel :

Proposition 1. Soit H un polynôme à coefficients dans k tel que $H^q - H$ est divisible par P . Alors on a :

$$P = \prod_{\lambda \in k} \text{pgcd}(P, H - \lambda).$$

Démonstration. Le membre de droite divise P . En effet les termes du produit sont premiers entre eux deux à deux et divisent tous P , donc leur produit divise P . Deux termes distincts sont premiers entre eux car ils sont diviseurs, disons de $H - \lambda$ et $H - \lambda'$ avec $\lambda \neq \lambda'$. Et comme ces deux derniers polynômes sont premiers entre eux, deux diviseurs de ces polynômes le sont aussi.

Montrons maintenant que le polynôme P divise le produit. Observons que $H^q - H$ s'écrit (en substituant H à X dans $X^q - X = \prod_{\lambda} (X - \lambda)$)

$$\prod_{\lambda \in k} (H - \lambda),$$

mais on sait que P divise $H^q - H$. Il divise donc

$$\text{pgcd}(P, H^q - H) = \text{pgcd}\left(P, \prod_{\lambda \in k} (H - \lambda)\right) = \prod_{\lambda \in k} \text{pgcd}(P, H - \lambda).$$

La dernière égalité résulte de ce que les polynômes $H - \lambda$ sont premiers entre eux deux à deux (on a en tous cas une relation de divisibilité, qui nous suffit). ■

Définition 1

On dit qu'un polynôme H est un *polynôme réducteur* pour le polynôme P si la factorisation associée est non-triviale.

Cette proposition a lieu sans supposer que le polynôme P ait tous ses facteurs irréductibles simples. À partir du moment où l'on a pu trouver un polynôme H

tel que P divise $H^q - H$ et que $|H| < |P|$ elle permet d'écrire P comme produit de polynômes premiers entre eux deux à deux. Pour que ce produit comporte au moins deux termes il faut et suffit que le polynôme H ne soit pas une constante. Les termes du produit ne sont pas nécessairement irréductibles. Mais s'il y a au moins deux termes dans le produit, on a ramené la factorisation du polynôme initial à celle de polynômes de plus bas degré.

La prochaine proposition montre que l'on peut trouver un polynôme réducteur de degré non nul si et seulement si le polynôme P n'est pas puissance d'un polynôme irréductible. Avant de l'énoncer il nous faut expliquer comment déterminer des polynômes H réducteurs pour P .

On réduit ce problème à la résolution d'un système d'équations linéaires. Soit N le degré de P , et pour un entier j fixé soit

$$\sum_{0 \leq i \leq N-1} a_{i,j} X^j$$

le reste de la division de X^{jq} par P . Considérons la matrice (N, N) $A(P) = (a_{i,j})$ (on indexe lignes et colonnes de 0 à $N-1$). Pour que le polynôme $H = \sum_{0 \leq j \leq N-1} \alpha_j X^j$ à coefficients dans k soit tel que $H^q - H$ soit divisible par P , il faut et suffit que :

$$A(P) \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{N-1} \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{N-1} \end{pmatrix}.$$

On s'est donc bien ramené à la résolution d'un système linéaire. On a :

Proposition 2. *Le rang de ce système est égal à $N - \ell$, où N est le degré de P et ℓ est le nombre de facteurs irréductibles apparaissant dans la décomposition de P .*

Ce système a toujours comme solution les polynômes constants. Mais ceux-ci donnent des décompositions triviales de P . Comme corollaire de ce résultat on a :

Corollaire. *Si le système linéaire précédent n'admet pour solutions que les polynômes constants, P est la puissance d'un polynôme irréductible.*

Ceci justifie la démarche suivie plus haut. Passons à la démonstration de la proposition 2.

Démonstration de la proposition 2. On utilise le lemme chinois pour construire des polynômes réducteurs pour P . Si $P = P_1^{\alpha_1} \cdots P_\ell^{\alpha_\ell}$ est la décomposition de P en facteurs irréductibles et si $(a_1, \dots, a_\ell)$ est un élément de k^ℓ le lemme chinois implique qu'il existe un polynôme H tel que H est congru à a_i modulo $P_i^{\alpha_i}$ pour tout $1 \leq i \leq \ell$. Si de plus nous imposons que le degré de H soit strictement inférieur à celui de P alors le polynôme est uniquement déterminé (exercice!).

Réciproquement, soit H un polynôme réducteur pour P avec $|H| < |P|$. La décomposition

$$H^q - H = \prod_{\lambda \in k} (H - \lambda)$$

est en facteurs premiers entre eux deux à deux. Donc, à i fixé, le facteur $P_i^{\alpha_i}$ divise l'un des termes du produit, et un seul. Il existe donc un élément $(a_1, \dots, a_\ell) \in k^\ell$ tel que H est congru à a_i modulo $P_i^{\alpha_i}$ pour tout i . ■

On vient d'établir une correspondance bi-univoque entre les éléments de k^ℓ et les polynômes H tels que $H^q - H$ soit divisible par P et $|H| < |P|$. Cette correspondance est une application linéaire entre l'espace des solutions du système linéaire considéré plus haut et k^ℓ .

La dimension du sous-espace des solutions est donc égal à ℓ . Reformulons le résultat, en gardant les notations précédentes :

Proposition 3. *Le nombre de facteurs indécomposables divisant le polynôme P est égal à $N - \varrho$, où ϱ est le rang de la matrice $B(P) = A(P) - I_N$.*

Voici deux exemples d'application sur le corps $\mathbb{F}_2$.

On considère le polynôme

$$P = X^9 + X^8 + X^6 + X^5 + X^3 + X + 1.$$

Sa dérivée P' est $X^8 + X^4 + X^2 + 1$. Calculons le pgcd de P et P' . Le reste R_1 de la division de P par P' est $X^6 + X^4 + X^2$, celui de la division de P' par R_1 est $R_2 = X^4 + 1$, celui de R_1 par R_2 est $R_3 = X^2$, enfin celui de R_2 par R_3 est 1. Le polynôme n'admet donc pas de facteurs multiples dans sa décomposition et est produit de facteurs indécomposables deux à deux distincts.

La matrice $A(P)$ considérée s'écrit

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

la matrice $B(P) = A(P) - I_9$ s'écrit :

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

son déterminant est évidemment nul. Le rang de cette matrice est 8. Le polynôme considéré est donc irréductible.

Voici le second exemple. Soit $P = X^{11} + X^8 + X^5 + X^4 + 1$, sa dérivée est $X^{10} + X^4$. Par l'algorithme d'Euclide on montre que le pgcd de P et P' est $(X^2 + X + 1)$.

On va donc appliquer l'algorithme au polynôme :

$$Q = P/(X^2 + X + 1) = X^9 + X^8 + X^3 + X + 1$$

qui n'a que des facteurs simples dans sa décomposition.

La matrice $B(Q)$ associée à Q est :

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Elle est de rang 6, les polynômes 1 , $X^6 + X^5 + X^2$ et $X^7 + X^6 + X^4$ constituent une base du noyau et sont réducteurs pour Q . On a donc :

$$P = \text{pgcd}(P, X^6 + X^5 + X^2) \times \text{pgcd}(P, X^7 + X^6 + X^4 + 1),$$

ce qui donne, après application de l'algorithme d'Euclide :

$$Q = (X^5 + X + 1)(X^4 + X^3 + 1).$$

On vérifie facilement que $X^4 + X^3 + 1$ est irréductible. Pour ce qui est de $R = X^5 + X + 1$, on peut écrire la décomposition directement ou recommencer l'algorithme. La matrice $B(R)$ associée est :

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Cette matrice est de rang 3. Les polynômes suivants forment une base du noyau 1 , $X^4 + X^3 + X$. On obtient donc :

$$X^5 + X + 1 = \dots$$

$$\dots = \text{pgcd}(X^5 + X + 1, X^4 + X^3 + X) \times \text{pgcd}(X^5 + X + 1, X^4 + X^3 + X + 1),$$

soit

$$X^5 + X + 1 = (X^2 + X + 1)(X^3 + X^2 + 1).$$

Finalement

$$P = (X^2 + X + 1)^2 (X^3 + X^2 + 1)(X^4 + X^3 + 1).$$

10. Les codes correcteurs d'erreurs

La théorie des *codes correcteurs d'erreurs* fournit une application de la théorie des corps finis aux télécommunications. Le problème est de savoir comment, à partir de données qui sont partiellement erronées, reconstituer les données exactes. Ce problème se pose lors de la transmission de données, celles-ci peuvent être altérées

à la suite d'erreurs dans la transmission. Pour résoudre ce problème l'idée est de transmettre l'information de telle manière que l'on puisse à partir de données plus ou moins altérée rétablir l'information originelle. Les corps finis fournissent des moyens, effectivement mis en oeuvre, de le faire.

Encore une fois cette section est une introduction à ces techniques, on renvoie à [De], [MS], et en particulier à [LN] ainsi qu'on l'a dit plus haut, pour une étude plus générale. Encore une fois précisons que la difficulté tient à ce que l'on a une longue chaîne d'arguments classiques.

Soit donc dans toute la suite $\mathbb{K}$ un corps fini à q éléments de caractéristique p . Sur $\mathbb{K}^n$ on introduit la fonction suivante, dite distance de Hamming :

Définition 1

Soient $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ des éléments de $\mathbb{K}^n$. La fonction de Hamming $h(x, y)$ entre x et y est le nombre d'indices i , avec $1 \leq i \leq n$, tels que $x_i \neq y_i$. Le poids de Hamming $w(x)$ est par définition $h(x, 0)$.

La proposition suivante justifie que l'on ait avant la définition appelée cette fonction *distance de Hamming* :

Proposition 1. *La fonction de Hamming est une distance sur $\mathbb{K}^n$.*

On emploiera donc la terminologie « distance de Hamming » pour cette fonction.

La démonstration est triviale. Seule l'inégalité triangulaire mérite un commentaire. Si k coordonnées de x et y diffèrent et si ℓ coordonnées de y et z diffèrent, les vecteurs x et z ont au plus $k + \ell$ coordonnées distinctes.

Supposons que l'on ait à transmettre des données, et que ces données soient exprimées comme une suite de h éléments de $\mathbb{K}$. On appellera cette suite l'information initiale. Lors du processus de transmission des erreurs peuvent se produire.

La question qui se pose est de reconstituer l'information initiale à partir des données reçues. À cette fin on commence par adjoindre des données supplémentaires à l'information initiale, on peut y ajouter $n - h$ éléments de $\mathbb{K}$, déterminés par la suite initiale. On peut par exemple répéter le message deux ou trois fois ! De façon plus générale on peut déterminer une suite, plus longue, de n éléments de $\mathbb{K}$ à partir de la suite initiale. Celle-ci n'apparaît pas nécessairement explicitement dans la suite finale qui doit satisfaire à certaines relations. On appellera cette suite les données transmises.

Comme elles sont déterminées par l'information initiale, elles doivent satisfaire à certaines relations. Dans l'exemple donné, il doit y avoir répétition des suites de symboles. S'il y a erreur durant la transmission les données reçues diffèrent des données transmises. Ceci se détecte en observant si les données reçues satisfont ou non aux relations mentionnées plus haut.

C'est ici qu'intervient la distance de Hamming. On considère la différence entre les données transmises et celles qui ont été reçues. Supposons que le nombre d'erreurs soit petit, disons inférieur à un entier fixé t . Les données transmises se trouvent à distance (pour la distance de Hamming) moins que t des données reçues. Si dans une boule de rayon t autour des données reçues il y a un seul élément qui satisfait aux relations imposées on a reconstitué les données transmises.

Dans le processus décrit on a donc à la source un espace $\mathbb{K}^h$ correspondant à l'information initiale, à l'arrivée un espace $\mathbb{K}^n$ correspondant aux données transmises et reçues, une application f de $\mathbb{K}^h$ dans $\mathbb{K}^n$ correspondant au processus de transmission théorique (sans altération).

Définition

Le sous-ensemble $C \subset \mathbb{K}^n$ image de f est appelé le code.

Enfin on a une autre application g de $\mathbb{K}^h$ dans $\mathbb{K}^n$ correspondant au processus de transmission réel, dans lequel il y a altération. La question est de savoir quand on peut déterminer f à partir de g .

Définition 2

L'ensemble C est par définition le *code*, un élément de C est un mot. L'entier h est la *dimension du code*, l'entier n est la *longueur du code*.

Définition 3

Un code est dit linéaire si l'ensemble C est un sous-espace vectoriel de $\mathbb{K}^n$, on le supposera de dimension h . Une matrice (h, n) à coefficients dans $\mathbb{K}$ dont les vecteurs lignes engendrent le code sera appelée une *matrice génératrice* du code, on la notera G . Une matrice $(n - h, n)$ à coefficients dans $\mathbb{K}$ dont les vecteurs lignes constituent un système d'équations pour le code sera appelée une *matrice de parité* du code, on la notera H .

Supposons que le processus de transmission réel g fasse au plus t erreurs. Soit x un élément de $\mathbb{K}^h$. S'il n'y a qu'un point du code dans la boule de centre $g(x)$ et de rayon t , ce point est $f(x)$. Ce sont nécessairement les données transmises. On dit alors que le code C peut corriger jusqu'à t erreurs.

Précisons comment ceci peut être assuré.

Définition 4

On appelle *distance minimale* du code C la quantité :

$$\delta_C = \min_{x, y \in C, x \neq y} h(x, y).$$

Dans le cas d'un *code linéaire* le code est un sous-espace vectoriel de dimension h de $\mathbb{K}^n$. Dans ce cas la distance minimale est le poids minimal d'un élément non nul. Pour le montrer il suffit de faire une translation : pour $x, y \in C$ on a $h(x, y) = w(x - y)$.

Les exemples suivants montrent que la distance minimale dépend du choix que l'on fait.

Si on prend pour sous-espace $\mathbb{K}^h \subset \mathbb{K}^n$, le sous-ensemble des vecteurs dont les $n - h$ dernières coordonnées sont nulles la distance minimale est 1. Prenons maintenant $n = 2h$ et choisissons pour sous-espace le sous-espace diagonal : celui formé par les vecteurs

$$(x_1, \dots, x_n, x_1, \dots, x_n).$$

Dans ce cas, si on a deux vecteurs distincts dans le code, ils ont nécessairement deux composantes distinctes, donc leur distance est au moins 2. La distance minimale de ce code est donc 2. De façon plus générale, on peut choisir $n = ah$ et définir le code comme le sous-espace diagonal dans $\mathbb{K}^{ah}$. La distance minimale de ce code est a .

Proposition 2. *Si la distance minimum δ_C d'un code C est inférieure ou égale à $2t + 1$ le code peut corriger jusqu'à t erreurs.*

Démonstration. La condition dit que dans une boule fermée de rayon t il y a au plus un élément du code. Donc si on a transmis des données, et qu'au plus t erreurs ont été commises, l'inégalité triangulaire montre que dans la boule qui a pour centre les données reçues et de rayon t , il y a un élément du code et un seul. Ce sont les données transmises. ■

Il est possible de donner pour la distance minimale d'un code linéaire une borne supérieure, ceci sera fait en exercice.

Le résultat suivant caractérise la distance minimale à partir de la matrice de parité.

Proposition 3. *Un code linéaire de matrice de parité H est de distance minimale δ supérieure ou égale à $d + 1$ si et seulement si tout système de d colonnes de la matrice de parité H est linéairement indépendant.*

Démonstration. Puisque le code est linéaire il suffit d'évaluer le poids minimal d'un vecteur non nul du code. Dire que la distance minimale δ est supérieure ou égale à $d + 1$ équivaut donc à dire qu'il n'existe pas de vecteur non nul dans le code qui appartienne à la boule de centre 0 et de rayon d . Soit qu'il n'existe pas de vecteurs non nuls dans le code ayant au plus d coordonnées non nulles.

Les vecteurs lignes de la matrice de parité sont les coefficients des formes linéaires définissant le code. Dire qu'il y a un élément non nul du code qui a au plus d

coordonnées non nulles équivaut donc à dire que l'on peut trouver une relation linéaire non triviale entre d colonnes de la matrice de parité. Les coefficients de cette relation linéaire sont les coordonnées du vecteur considéré. Donc dire que la distance minimale est supérieure ou égale à $d + 1$ équivaut à dire que tout système de d colonnes de la matrice de parité est libre. ■

On va dans la suite décrire deux types de code, et ce dans un des cas jusqu'au décodage. Cela permettra d'illustrer des techniques diverses d'algèbre, aucune n'est compliquée en elle-même, c'est leur succession qui rend le problème délicat.

Définition 5

On dira qu'un code C de longueur n est cyclique si et seulement si

$$(a_1, \dots, a_n) \in C \implies (a_n, a_1, \dots, a_{n-1}) \in C.$$

On va donner une interprétation multiplicative des codes cycliques. Considérons l'anneau quotient $\mathbb{K}[X]/(X^n - 1)$, l'application linéaire

$$\begin{aligned} \psi : \mathbb{K}^n &\longrightarrow \mathbb{K}[X]/(X^n - 1), \\ (a_1, \dots, a_n) &\longmapsto a_1 + a_2X + \dots + a_nX^{n-1}, \end{aligned}$$

est bijective.

Proposition 4. Soit C un code dans $\mathbb{K}^n$, C est cyclique si et seulement si $\psi(C)$ est un idéal.

Démonstration. En effet, si la classe d'un polynôme P est dans le code, vu comme idéal de $\mathbb{K}[X]/(X^n - 1)$, il en est de même pour la classe XP . Ce qui veut exactement dire que si $(a_1, \dots, a_n) \in C$ il en est de même pour $(a_n, a_1, \dots, a_{n-1})$. Inversement, cette dernière condition, interprétant cette suite comme la classe d'un polynôme P , signifie que si P est dans le code, il en est de même pour XP et plus généralement pour RP , où R est un polynôme quelconque. Le résultat suit. ■

Lemme 1. Soit P un polynôme quelconque à coefficients dans k . Tout idéal dans $\mathbb{K}[X]/(P)$ est principal, c'est-à-dire admet un générateur.

Démonstration. Soit $\mathcal{I}$ un idéal dans $\mathbb{K}[X]$, et soit π l'application canonique de $\mathbb{K}[X]$ vers $\mathbb{K}[X]/\mathcal{I}$. L'idéal $\pi^{-1}(\mathcal{I})$ de $\mathbb{K}[X]$ est principal, il admet un générateur, soit R . Alors, $\pi(R)$ est un générateur pour $\mathcal{I}$. ■

Soit C un code cyclique, que l'on identifie à l'idéal de l'anneau quotient $\mathbb{K}[X]/(X^n - 1)$ auquel il correspond. On appellera *polynôme générateur* de C un polynôme G de plus bas degré dont la classe modulo $X^n - 1$ engendre C . Le polynôme générateur est bien défini à multiplication par un scalaire non nul près.

Lemme 2. Le polynôme G divise $X^n - 1$.

Démonstration. En effet, par définition de G , la classe du pgcd de G et $X^n - 1$ est dans l'idéal de l'anneau quotient de

$$\mathbb{K}[X]/(X^n - 1)$$

engendré par G . Ce pgcd doit donc être du degré de G , donc égal à G à un scalaire près.

Notons que la dimension du code h est égale $n - |G|$. ■

Définition 6

On appellera *polynôme de parité* du code le polynôme :

$$\frac{X^n - 1}{G(X)}.$$

La façon la plus classique pour coder l'information initiale dans ce cas précis est la suivante. On interprète les données initiales comme un polynôme à coefficients dans $\mathbb{K}$, soit D , les données transmises sont la classe du produit GD dans l'anneau quotient.

Étude d'un code BCH

Une façon commode de déterminer un idéal dans $\mathbb{K}[X]$ est de le construire comme ensemble des polynômes qui ont pour racines certains éléments d'une extension algébrique $\mathbb{L}$ de $\mathbb{K}$. Puis on peut considérer l'image dans l'anneau quotient $\mathbb{K}[X]/(X^n - 1)$, comme l'application est surjective c'est un idéal. Si les éléments considérés dans $\mathbb{L}$ sont des racines n -ièmes de 1, la valeur prise par un élément d'une classe modulo $X^n - 1$ ne dépend que de la classe et non de l'élément. On peut donc directement définir un idéal dans $\mathbb{K}[X]/(X^n - 1)$ comme étant l'ensemble des classes prenant la valeur 0 sur certaines racines n -ièmes de 1. Choisissons donc une racine primitive n -ième de 1, soit μ dans un corps $\mathbb{L}$ de caractéristique p , on suppose évidemment que n est premier à la caractéristique p . Pour que $\mathbb{L}$ contienne des racines primitives n -ièmes de l'unité il faut et il suffit que n divise $q - 1$. Considérons l'idéal des classes de polynômes s'annulant sur $\mu^b, \dots, \mu^{b+d-2}$.

On suppose que ces racines sont deux à deux distinctes, c'est-à-dire que l'on suppose que l'ordre n de μ dans le groupe multiplicatif $\mathbb{L}^*$ est supérieur ou égal à d .

On peut évidemment aussi considérer les polynômes nuls sur ces racines, ce sont ceux divisibles par le ppcm des polynômes minimaux des μ^i , $b \leq i \leq b + d - 2$, qui est le polynôme générateur du code. Un code déterminé par un idéal de ce type est appelé un *code BCH* (pour Bose, Chauduri, Hocquenghem). À cause de la proposition suivante on dit qu'il est de distance prescrite d .

Proposition 5. Un code BCH est de distance minimale au moins d . Si $d = 2t + 1$ il peut donc corriger jusqu'à t erreurs.

Démonstration. En effet, on doit vérifier qu'il n'y a pas, dans le code, de vecteurs non nuls ayant moins de d coordonnées non nulles. Identifions un vecteur à un polynôme non nul $P = a_1 + a_2X + \dots + a_{n-1}X^{n-1}$. On doit vérifier que les équations $P(\mu^i) = 0$, avec $b \leq i \leq b + d - 2$ ont lieu si le polynôme a au moins d coefficients non nuls. On a $d - 1$ équations à n inconnues et la matrice du système linéaire est la suivante :

$$\begin{pmatrix} 1 & \mu^b & \dots & \mu^{b(n-1)} \\ 1 & \mu^{b+1} & \dots & \mu^{(b+1)(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mu^{b+d-2} & \dots & \mu^{(b+d-2)(n-1)} \end{pmatrix}.$$

Si on sélectionne $d - 1$ colonnes quelconques, on reconnaît, à un facteur multiplicatif près, un déterminant de Van der Monde qui est non nul. Si bien que $d - 1$ colonnes quelconques de cette matrice sont linéairement indépendantes. Donc un élément non nul dans le code doit avoir au moins d coordonnées non nulles. Donc si P n'a que $d - 1$ coefficients éventuellement non nuls, il est nul. ■

Algorithme de décodage

Venons en maintenant au processus de décodage. Supposons que l'on ait le code BCH de longueur n et de distance minimale au moins égale à $d = 2t + 1$ défini plus haut. Supposons que le processus de transmission des données fasse au plus t erreurs, si bien que le code peut corriger ces erreurs. Soit $R \in \mathbb{K}[X]/(X^n - 1)$ les données reçues, cherchons à déterminer les données transmises. Considérons la différence, que l'on appellera le polynôme d'erreur, $E = R - T$ entre les données transmises T et les données reçues R . On va procéder en deux étapes, d'abord on va déterminer en quels emplacements se sont produites les erreurs, puis on les déterminera.

Les valeurs $R(\mu^i)$ pour $b \leq i \leq b + d - 2$ sont connues, par construction on a $E(\mu^i) = R(\mu^i)$ pour $b \leq i \leq b + d - 2$.

Les coefficients de E sont donc solutions d'un système linéaire de $d - 1$ équations et n inconnues (les coefficients du polynôme E). Il y a au moins une solution par hypothèse, mais en général il y en a plusieurs. Pour déterminer la bonne on procède comme suit. Établissons d'abord un lemme.

Considérons un entier $u \leq t$ un entier, et un polynôme :

$$F = \sum_{i=0, \dots, u-1} f_i X^{a'_i},$$

où les exposants sont a'_j deux à deux distincts et strictement inférieurs à n . Les coefficients $f_i \in \mathbb{K}$ ne sont pas nécessairement non nuls. Posons $y_j = \mu^{a'_j}$ et posons :

$$\prod_{j=1, \dots, u} (X - y_j) = \sum_{\ell=0, \dots, u} \alpha_{u-\ell} X^\ell.$$

Si, dans ce polynôme, on substitue la valeur y_j à la variable X , on obtient :

$$\sum_{\ell=0, \dots, u} \alpha_{u-\ell} y_j^\ell = 0.$$

Multiplions cette équation par $f_j y_j^k$, pour $b \leq k \leq b+d-2$, et sommons sur j , nous obtenons, par définition de F , la relation suivante :

$$\alpha_u F(\mu^k) + \alpha_{u-1} F(\mu^{k+1}) + \dots + \alpha_1 F(\mu^{k+u-1}) + F(\mu^{k+u}) = 0.$$

On obtient donc un système de $d-1$ équations linéaires en les u inconnues α_j .

Lemme 3. *Le système linéaire précédent admet une solution unique seulement si les coefficients de f_j de F sont tous non nuls.*

Démonstration. Nous ne discutons pas ici l'existence mais seulement l'unicité de la solution. Pour étudier cette question, sélectionnons les u premières équations parmi les $d-1$ équations données. On obtient comme matrice du système :

$$\begin{pmatrix} F(\mu^b) & F(\mu^{b+1}) & \dots & F(\mu^{b+u-1}) \\ F(\mu^{b+1}) & F(\mu^{b+2}) & \dots & F(\mu^{b+u}) \\ \vdots & \vdots & & \vdots \\ F(\mu^{b+u-1}) & F(\mu^{b+u}) & \dots & F(\mu^{b+2u-1}) \end{pmatrix}.$$

Or cette matrice est égale au produit suivant VDV^t où on a

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ y_1 & y_2 & \dots & y_u \\ \vdots & \vdots & & \vdots \\ y_1^{u-1} & y_2^{u-1} & \dots & y_u^{u-1} \end{pmatrix}, \quad \text{et} \quad D = \begin{pmatrix} f_1 y_1^{u-1} & 0 & \dots & 0 \\ 0 & f_2 y_2^{u-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & f_u y_u^{u-1} \end{pmatrix}.$$

Donc le déterminant de la matrice du système est égal à :

$$f_1 \dots f_u (y_1 \dots y_u)^{u-1} \prod_{0 \leq i < j \leq u} (y_i - y_j)^2.$$

Comme les racines y_i sont supposées deux à deux distinctes ce déterminant est non nul si et seulement si tous les coefficients f_j sont non nuls. ■

Appliquons ces résultats au polynôme d'erreur E , supposons que r erreurs aient été commises, avec $r \leq t$, et soit enfin $u \leq t$. Considérons le déterminant :

$$D_u = \begin{vmatrix} E(\mu^b) & E(\mu^{b+1}) & \dots & E(\mu^{b+u-1}) \\ E(\mu^{b+1}) & E(\mu^{b+2}) & \dots & E(\mu^{b+u}) \\ \vdots & \vdots & & \vdots \\ E(\mu^{b+u-1}) & E(\mu^{b+u}) & \dots & E(\mu^{b+2u-1}) \end{vmatrix}.$$

On a :

Corollaire. *Le déterminant D_u , $1 \leq u \leq t$, est nul dès que $u > r$, il est non nul pour $u = r$.*

Pour déterminer le nombre d'erreurs commises on doit donc calculer ces déterminants et calculer la plus grande valeur inférieure ou égale à t pour laquelle il est non nul.

On calcule alors le polynôme d'erreurs comme suit. Posons

$$E = \sum_{i=0, \dots, r} e_i X^{a_i},$$

avec $r \leq t$, $e_i \in \mathbb{K}^*$ et les entiers a_i deux à deux distincts et strictement inférieurs à n . Les entiers a_i correspondent aux emplacements où se sont produites les erreurs. Posons $x_i = \mu^{a_i}$ et considérons le polynôme à coefficients dans $\mathbb{L}$:

$$\prod_{i=1, \dots, e} (X - x_i) = \sum_{j=0, \dots, e} \tau_{e-j} X^j.$$

Si, dans ce polynôme on substitue à la variable X la valeur x_i , puis si on somme sur i on obtient comme précédemment un système linéaire :

$$\tau_r E(\mu^k) + \tau_{r-1} E(\mu^{k+1}) + \dots + \tau_1 E(\mu^{k+r-1}) + E(\mu^{k+r}) = 0,$$

pour $k = b, \dots, b+r-1$. On peut donc calculer la valeur des τ_i qui sont, au signe près, les fonctions symétriques élémentaires des x_i . Précisément $(-1)^i \tau_i$ est la i -ième fonction symétrique élémentaire des racines x_i , si bien que l'on obtient ces racines en résolvant l'équation :

$$X^r + \tau_1 X^{r-1} + \dots + \tau_r = 0.$$

Notons que l'on peut toujours résoudre une telle équation, quitte à tester tous les éléments du corps fini $\mathbb{L}$!

Enfin, on calcule les coefficients e_i en résolvant le système linéaire :

$$\sum_{1 \leq i \leq r} e_i \mu^{j(a_i)} = R(\mu^j).$$

Donnons un exemple, prenons pour $\mathbb{K}$ le corps $\mathbb{F}_2$. Choisissons pour μ une racine primitive de 1, qui a pour polynôme minimal $X^5 + X^2 + 1$. Puis choisissons $b = 1$ et $d = 5$. Notons que, μ^i pour $1 \leq i \leq 5$ sont bien deux à deux distinctes. Voici les polynômes minimaux de μ^2, μ^3, μ^4 , et μ^5 .

Celui de μ , c'est-à-dire $X^5 + X^2 + 1$, l'est aussi pour μ^2 et μ^4 . Celui de μ^3 est $X^5 + X^4 + X^3 + X^2 + 1$, pour le vérifier on élève au cube l'équation $\mu^6 + \mu^3 = \mu$. Celui de μ^5 est $X^5 + X^4 + X^2 + X + 1$, pour le vérifier on écrit que $\mu^5 + 1 = \mu^2$ et on élève l'équation à la puissance 5. Il faut bien entendu vérifier que tous ces polynômes sont irréductibles.

Le polynôme générateur est donc de degré 15, et est égal au produit

$$(X^5 + X^2 + 1)(X^5 + X^4 + X^3 + X^2 + 1)(X^5 + X^4 + X^2 + X + 1).$$

Il est égal à :

$$1 + X + X^2 + X^3 + X^5 + X^7 + X^8 + X^9 + X^{10} + X^{11} + X^{15}.$$

Considérons le code *BCH* associé de dimension 16, il est de longueur 31, il peut corriger jusqu'à deux erreurs, supposons donc qu'au plus deux erreurs se soient produites. Considérons le cas où les données reçues sont le polynôme :

$$1 + X + X^3 + X^7 + X^9 + X^{10} + X^{12} + X^{13} + X^{17} + X^{18} + X^{19}.$$

Le polynôme d'erreur E prend les mêmes valeurs que ce polynôme en μ^i pour $1 \leq i \leq 5$, ce sont dans l'ordre : $\mu^4, \mu^3 + \mu^2 + 1, \mu^3 + \mu^2 + 1, \mu^3 + \mu^2 + \mu, \mu^4 + \mu^3 + \mu + 1$.

Le déterminant

$$\begin{vmatrix} \mu^4 & \mu^3 + \mu^2 + 1 \\ \mu^3 + \mu^2 + 1 & \mu^3 + \mu^2 + \mu \end{vmatrix},$$

est nul. Donc seulement une erreur s'est produite. Le polynôme E est donc de la forme X^k , comme il prend en μ la valeur μ^4 il est égal à X^4 . Le polynôme transmis était donc :

$$1 + X + X^3 + X^4 + X^7 + X^9 + X^{10} + X^{12} + X^{13} + X^{17} + X^{18} + X^{19}.$$

Pour obtenir les données initiales il faut diviser par le polynôme générateur. On obtient :

$$1 + X^2 + X^3 + X^4.$$

Supposons maintenant que les données reçues soient :

$$1 + X + X^2 + X^4 + X^5 + X^7 + X^8 + X^9 + X^{10} + X^{11} + X^{15} + X^{20}.$$

Le polynôme d'erreur E prend les mêmes valeurs que ce polynôme en μ^i pour $1 \leq i \leq 5$, ce sont dans l'ordre $\mu^2, \mu^4, \mu^4 + \mu + 1, \mu^3 + \mu^2 + 1, \mu^3 + \mu + 1$.

Le déterminant

$$\begin{vmatrix} \mu^2 & \mu^4 \\ \mu^4 & \mu^4 + \mu + 1 \end{vmatrix},$$

est non nul et égal à $\mu^3 + \mu + 1$. Donc seulement deux erreurs se sont produites. Le polynôme E est donc somme de deux termes, c'est-à-dire de la forme $X^{a_1} + X^{a_2}$. Les quantités $\tau_0 = \mu^{a_1+a_2}$ $\tau_1 = \mu^{a_1} + \mu^{a_2}$ sont solutions du système :

$$\tau_1 \mu^2 + \tau_0 \mu^4 = \mu^4 + \mu + 1$$

$$\tau_1 \mu^4 + \tau_0 (\mu^4 + \mu + 1) = \mu^3 + \mu^2 + 1$$

On obtient

$$\tau_0 = \mu^3 + \mu^2 + \mu$$

$$\tau_1 = \mu^2$$

Donc μ^{a_1} et μ^{a_2} sont solutions de l'équation du second degré

$$X^2 + \mu^2 X + \mu^3 + \mu^2 + \mu = 0.$$

Les zéros sont μ^3 et μ^{20} . Les données transmises étaient donc, comme dans le cas précédent :

$$1 + X + X^3 + X^4 + X^7 + X^9 + X^{10} + X^{12} + X^{13} + X^{17} + X^{18} + X^{19}.$$

On notera dans ce cas que le déterminant :

$$\begin{vmatrix} \mu^2 & \mu^4 & \mu^4 + \mu + 1 \\ \mu^4 & \mu^4 + \mu + 1 & \mu^3 + \mu^2 + 1 \\ \mu^4 + \mu + 1 & \mu^3 + \mu^2 + 1 & \mu^3 + \mu + 1 \end{vmatrix},$$

est nul.

Codes de Goppa

Dans ce paragraphe on étudie une autre classe de codes, les codes de Goppa. Par comparaison au cas précédent les codes de Goppa sont construits de manière à avoir une distance minimale supérieure à une valeur prescrite à l'avance. Nous donnerons dans ce cas seulement les définitions, le calcul d'une matrice de parité et des propriétés de base du code. Encore une fois, un des intérêts de cette étude, outre les applications, est la mise en oeuvre de techniques de base d'algèbre linéaire ainsi que de la théorie de corps finis. Nous commençons par introduire quelques notations et un lemme.

Soient $\mathbb{F}_q$ un corps fini, m un entier positif non-nul, $G \in \mathbb{F}_{q^m}[X]$, et soit enfin $L = \{\alpha_1, \dots, \alpha_n\}$ une famille d'éléments de $\mathbb{F}_{q^m}$ qui ne sont pas des racines du polynôme G . Le polynôme $X - \alpha_i$ est donc par hypothèse premier avec G .

On commence par observer que la classe de $X - \alpha_i$ est un élément inversible dans l'anneau quotient $\mathbb{F}_{q^m}[X]/(G)$. C'est une conséquence de l'identité de Bezout. Soient en effet U et V deux polynômes tels que $U(X - \alpha_i) + VG = 1$. En réduisant modulo G on obtient $U(X - \alpha_i) \equiv 1$ modulo G .

Définition 7

Le code de Goppa Γ associé à G et L est le sous-espace de $\mathbb{F}_q^n$ constitué par les vecteurs $(a_1, \dots, a_n)$ tels que l'élément de l'anneau quotient $\mathbb{F}_{q^m}[X]/(G)$:

$$\sum_{i=1, \dots, n} \frac{a_i}{X - \alpha_i}$$

soit nul.

Si le polynôme G est irréductible le code est dit irréductible.

Proposition 6. *Le code Γ est un code linéaire de longueur n , de dimension k telle que $k \geq |G|$. En fin la distance minimale d de Γ est telle que $d \geq |G| + 1$.*

La première propriété est conséquence immédiate de la définition. La seconde et la troisième doivent être démontrées. Pour ce qui est de la seconde on considère une base de $\mathbb{F}_{q^m}$ comme espace vectoriel sur $\mathbb{F}_q$, $\mathbb{F}_{q^m}$ est de dimension m sur $\mathbb{F}_q$; soit donc $\{v_1, \dots, v_m\}$ une telle base. Posons $r = |G|$, l'anneau quotient $\mathbb{F}_{q^m}[X]/(G)$ est lui de dimension r comme espace vectoriel sur $\mathbb{F}_{q^m}$, de base $\{1, X, \dots, X^{r-1}\}$. L'équation de définition du code Γ se réécrit comme suit :

$$\sum_{i=1, \dots, n} a_i \left(\sum_{j=0, \dots, r-1} u_{i,j} X^j \right) = 0$$

où $(X - \alpha_i)^{-1} = \sum_{j=0, \dots, r-1} u_{i,j} X^j$ dans $\mathbb{F}_{q^m}[X]/(G)$. Ce qui peut encore se réécrire

$$\sum_{i=1, \dots, n} a_i \left(\sum_{j=0, \dots, r-1} \sum_{k=1, \dots, m} \nu_{i,j,k} v_k X^j \right) = 0$$

où $\sum_{k=1, \dots, m} \nu_{i,j,k} v_k$, $\nu_{i,j,k} \in \mathbb{F}_q$, est la décomposition de l'élément $u_{i,j}$ de $\mathbb{F}_{q^m}$ dans la base des v_k .

La nullité de l'élément

$$\sum_{i=1, \dots, n} a_i \left(\sum_{j=0, \dots, r-1} u_{i,j} X^j \right)$$

se traduit par la nullité des coefficients de X^j , $0 \geq j \geq r-1$, soit

$$\sum_{i=1, \dots, n} a_i u_{i,j} = 0, \quad 0 \geq j \geq r-1.$$

Puis on écrit que ces éléments de $\mathbb{F}_{q^m}$ sont nuls, c'est-à-dire que leurs coefficients dans la base des v_k le sont. Soit $\sum_{i=1, \dots, n} a_i \nu_{i,j,k} = 0$, à j et k fixés. On a donc mr équations à coefficients dans $\mathbb{F}_q$. Le code est donc de dimension au moins $n - mr$.

Pour ce qui est du second point il faut montrer qu'un point non nul du code a au moins $r+1$ coordonnées non nulles.

C'est un exercice que de vérifier que l'élément $(X - \alpha_i)^{-1}$ est, dans l'anneau quotient $\mathbb{F}_q^m[X]/(G)$, égal à

$$-\frac{G - G(\alpha_i)}{X - \alpha_i} G(\alpha_i)^{-1}.$$

Il en résulte que $a = (a_1, \dots, a_n) \in \Gamma$ si et seulement si l'élément de l'anneau quotient :

$$\sum_{i=1, \dots, n} a_i \frac{G - G(\alpha_i)}{X - \alpha_i} G(\alpha_i)^{-1}$$

est nul. Cet élément étant en tant que polynôme de degré inférieur ou égal à $r-1$ cette condition d'annulation peut être prise indifféremment dans l'anneau de polynômes.

Posons $G = g_r X^r + \dots + g_0$, on a :

$$\frac{G - G(\alpha_i)}{X - \alpha_i} = g_r(X^{r-1} + \alpha_i X^{r-2} + \dots + \alpha_i^{r-1}) + g_{r-1}(X^{r-2} + \dots + \alpha_i^{r-2}) + \dots + g_2(X + \alpha_i) + g_1.$$

On écrit alors que les coefficients de $X^{r-1}, \dots, 1$ dans l'élément ci-dessus sont nuls et on obtient la matrice de parité dont le terme $p_{i,j}$ est égal à

$$(g_{r+1-i} + \alpha_j g_{r+2-i} + \dots + \alpha_j^{i-1}) G(\alpha_j)^{-1}.$$

soit la matrice :

$$\begin{pmatrix} g_r G(\alpha_1)^{-1} & g_r G(\alpha_2)^{-1} & g_r G(\alpha_n)^{-1} \\ (g_{r-1} + \alpha_1 g_r) G(\alpha_1)^{-1} & (g_{r-1} + \alpha_1 g_r) G(\alpha_2)^{-1} & (g_{r-1} + \alpha_1 g_r) G(\alpha_n)^{-1} \\ \dots & \dots & \dots \\ (g_1 + \alpha_1 g_2 + \dots + \alpha_1^{r-1} g_r) G(\alpha_1)^{-1} & (g_1 + \alpha_1 g_2 + \dots + \alpha_1^{r-1} g_r) G(\alpha_2)^{-1} & \dots (g_1 + \alpha_1 g_2 + \dots + \alpha_1^{r-1} g_r) G(\alpha_n)^{-1} \end{pmatrix}$$

La terminologie matrice de parité est dans ce cas abusive, car chacune des équations obtenues à partir d'une ligne (un coefficient du polynôme) est une équation à coefficients dans $\mathbb{F}_q^m$ et non dans $\mathbb{F}_q$. En décomposant comme il a été dit plus haut on obtient pour chaque ligne m équations linéaires à coefficients dans $\mathbb{F}_q$.

La matrice qui précède est égale au produit (dans l'ordre) des deux matrices suivantes :

$$\begin{pmatrix} g_r & 0 & \dots & 0 \\ g_{r-1} & g_r & \dots & 0 \\ \dots & \dots & \dots & \dots \\ g_1 & g_2 & g_3 & \dots & g_r \end{pmatrix} \begin{pmatrix} G(\alpha_1)^{-1} & G(\alpha_2)^{-1} & \dots & G(\alpha_n)^{-1} \\ \alpha_1 G(\alpha_1)^{-1} & \alpha_1 G(\alpha_2)^{-1} & \dots & \alpha_1 G(\alpha_n)^{-1} \\ \dots & \dots & \dots & \dots \\ \alpha_1^{r-1} G(\alpha_1)^{-1} & \alpha_1^{r-1} G(\alpha_2)^{-1} & \dots & \alpha_1^{r-1} G(\alpha_n)^{-1} \end{pmatrix}$$

Comme la première matrice est inversible ($g_r \neq 0$) la seconde matrice est encore, par définition, une matrice de parité. Le calcul du déterminant de Van der Monde montre alors que r colonnes quelconques sont linéairement indépendantes (les α_i étant supposés deux à deux distincts).

Le calcul de la distance minimum en résulte.

On va préciser ce résultat dans le cas où $q=2$. Soit un élément du code $\mathbf{a} = (a_1, \dots, a_n)$. Supposons que exactement ℓ scalaires a_i soient non-nuls : $a_{\phi(1)}, \dots, a_{\phi(\ell)}$. Définissons un polynôme $f_{\mathbf{a}}(X)$ par $f_{\mathbf{a}} = \prod_{i=1, \dots, \ell} (X - \alpha_{\phi(i)})$.

Alors la dérivée logarithmique donne :

$$R_a = \sum_{i=1, \dots, \ell} \frac{1}{X - \alpha_{\phi(i)}} = \frac{f'_a}{f_a}.$$

Les polynômes G et f_a sont premiers entre eux, par conséquent $R_a = 0$ si et seulement si G divise f'_a . Notons $\bar{G}$ le polynôme, carré parfait, de plus bas degré, qui est divisible par G , il est bien défini à un carré scalaire non-nul près. Alors,

Lemme 4. *Le polynôme G divise f'_a si et seulement si $\bar{G}$ divise f'_a .*

Comme on travaille en caractéristique 2, et sur le corps $\mathbb{F}_2$ on a :

Lemme 5. *Un polynôme dérivé ne contient que des puissances paires et est donc un carré parfait.*

Il en résulte que si une puissance impaire H^{2k+1} d'un polynôme irréductible H sur $\mathbb{F}_2$ divise le polynôme f'_a , la puissance H^{2k+2} le divise aussi. Le lemme suit si on part de l'hypothèse que G divise f'_a . L'autre sens de l'équivalence est clair.

On en conclut que l'élément $a \in \Gamma$ si et seulement si $\bar{G}$ divise f'_a . Si a est un élément non nul, comme $|f'_a| \geq |\bar{G}|$, la distance minimale de Γ est supérieure ou égale à $|\bar{G}| + 1$. En particulier, si G n'a pas de racines doubles, elle est supérieure ou égale à $2|G| + 1$.

Exercices

I. Éléments algébriques, polynômes cyclotomiques

1. Calculer les polynômes minimaux des éléments suivants : $\sqrt{2} + \sqrt{7}$, $i + \sqrt{5}$, $i + \sqrt{3} + \sqrt{7}$, $i + 2^{\frac{1}{5}}$.
2. Soit μ une racine primitive n -ième de l'unité. Soit k un entier. Calculer la somme

$$\sum_{i=0, \dots, n-1} \mu^{ki}.$$

Discuter en fonction de k .

3. Soit $\alpha \in \mathbb{R} - \mathbb{Q}$ une racine d'un polynôme irréductible $P \in \mathbb{Z}[X]$ de degré $n > 1$. Montrer qu'il existe une constante C telle que pour tout $\frac{p}{q} \in \mathbb{Q}$ on ait

$$|\alpha - \frac{p}{q}| \geq \frac{C}{q^n}.$$

On distinguera suivant le cas où $\frac{p}{q}$ est proche, en un sens à préciser, de α ou non, on évaluera le polynôme P en la valeur $\frac{p}{q}$, et on appliquera le théorème des accroissements finis.

4. Nombres de Liouville.

En utilisant l'exercice précédent on montrera que le réel

$$\sum_n \frac{1}{10^{n!}}$$

est transcendant.

5. Calculer $\Phi_n(1)$ et $\Phi_n(-1)$.

6. Calculer les polynômes cyclotomiques Φ_{12} , Φ_{24} , Φ_{15} , Φ_9 , Φ_{27} , Φ_{36} .

7. Soit p un nombre premier, et soit $h > 1$. Montrer que $\Phi_{p^h}(X) = \Phi_{p^{h-1}}(X^p)$.

8. Soit h un entier impair. Montrer que $\Phi_{2h}(X) = \Phi_h(-X)$.

9. Indépendance linéaire des caractères.

Soit $\mathbb{K}$ un corps. Soient $c_1, \dots, c_n$, n homomorphismes deux à deux distincts de $\mathbb{Z}$ dans le groupe multiplicatif $\mathbb{K}^*$, de tels homomorphismes sont aussi appelés des caractères. Supposons que pour certains éléments $a_i \in K$ on ait une relation de la forme :

$$\sum a_i c_i(m) = 0,$$

pour tout entier $m \in \mathbb{Z}$. Montrer alors que les a_i sont tous nuls.

On pourra soit faire une récurrence sur n , soit utiliser un déterminant de Van der Monde.

10. Symbole de Legendre.

Soit p un nombre premier impair. Définissons le symbole $\left(\frac{x}{p}\right)$, pour $x \in \mathbb{F}_p^*$ par $\left(\frac{x}{p}\right) = 1$ si $x \in (\mathbb{F}_p^*)^2$, par $\left(\frac{x}{p}\right) = -1$ sinon.

a) ** Montrer que l'application de $\mathbb{F}_p^*$ dans $(1, -1)$ est un homomorphisme, montrer que :

$$\sum_{x \in \mathbb{F}_p^*} \left(\frac{x}{p}\right) = 0.$$

b) Soit μ une racine primitive p -ième de l'unité dans $\mathbb{C}$. Posons :

$$s = \sum_{x \in \mathbb{F}_p^*} \left(\frac{x}{p}\right) \mu^x.$$

Montrer que :

$$s^2 = \sum_{x, y \in \mathbb{F}_p^*} \left(\frac{y}{p}\right) \mu^{x(y+1)}.$$

Puis écrire que :

$$s^2 = \sum_{x, y \in \mathbb{F}_p^*, y \neq -1} \left(\frac{y}{p}\right) \mu^{x(y+1)} + \sum_{x \in \mathbb{F}_p^*} \left(\frac{-1}{p}\right).$$

En conclure que $s^2 = \left(\frac{-1}{p}\right)p$.

c) Soit d un entier. Montrer que le corps $\mathbb{Q}(\sqrt{d})$ est un sous-corps d'un corps cyclotomique.

11. Soit p un nombre premier, considérons la réduction modulo p du polynôme cyclotomique Φ_n . Montrer que cette réduction n'est pas, en général, irréductible.

12. Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$.

Soit Φ_n le polynôme cyclotomique, et soit p un nombre premier ne divisant pas n . Soit μ une racine primitive n -ième de l'unité.

a) Soit $P_1 \cdots P_k$ la décomposition de Φ_n en facteurs irréductibles dans $\mathbb{Z}[X]$. Supposons que μ soit racine de P_1 et que μ^p soit racine de P_2 . Montrer que μ est une racine de $P_2(X^p)$ et que donc P_1 divise $P_2(X^p)$. En déduire que la réduction modulo p des polynômes P_1 et P_2 ont un facteur commun.

b) En utilisant les résultats du chapitre 3 montrer que la réduction modulo p du polynôme cyclotomique Φ_n n'a que des facteurs simples dans sa décomposition. En déduire que si μ est racine de P_1 , μ^p est aussi racine de P_1 . On utilisera la question précédente.

c) En conclure que si μ est racine de P_1 toutes les racines primitives n -ièmes le sont et que Φ_n est irréductible.

13. a) (Construction à la règle et au compas) L'objet de cet exercice est de donner un aperçu sur le problème des constructions géométriques à la règle et au compas. On s'en tiendra aux aspects algébriques, renvoyant à [A] et [Pe] pour les liens avec la géométrie.

Soit K un sous-corps de $\mathbb{R}$. On dira qu'un réel a , algébrique sur $\mathbb{Q}$, est constructible si il existe une suite de corps F_i tels que :

$$\mathbb{Q} \subset F_1 \subset F_2 \subset \cdots \subset \mathbb{Q}(a) \subset \mathbb{R},$$

pour tout i le corps F_i est obtenu à partir de F_{i-1} par adjonction de racines carrées d'éléments positifs de F_{i-1} .

Montrer que l'ensemble des éléments constructibles est un sous-corps de $\mathbb{R}$.

b) Montrer que le degré d'un élément constructible est une puissance de 2.

c) Montrer que $\cos(\frac{\pi}{9})$ n'est pas constructible.

d) Montrer que $2^{\frac{1}{3}}$ n'est pas constructible.

Ces deux dernières questions montrent l'impossibilité de la trissection de l'angle et de la duplication du cube à la règle et au compas.

e) Montrer que $\cos(\frac{2\pi}{5})$ est constructible.

14. En utilisant la formule de Newton en caractéristique 2, donner une condition nécessaire et suffisante pour que le coefficient binomial C_n^k soit divisible par 2.

II. Corps finis

1. Démontrer que les polynômes suivants, à coefficients dans $\mathbb{F}_2$, sont irréductibles :

► $P_1 = X^7 + X + 1,$

► $P_2 = X^6 + X^3 + 1.$

Trouver un générateur du groupe multiplicatif du corps $\mathbb{F}_2[X]/(P_i)$, $i = 1, 2$.

2. Démontrer que les polynômes suivants à coefficients dans $\mathbb{F}_3$ sont irréductibles :

► $P_1 = X^3 + 2X + 1,$

► $P_2 = X^3 + X^2 + 2,$

► $P_3 = X^4 + X^2 + 2.$

Trouver un générateur du groupe multiplicatif du corps $\mathbb{F}_3[X]/(P_i)$, $i = 1, 2, 3$.

3. Factoriser les polynômes suivants à coefficients dans $\mathbb{F}_2$:

► $X^{11} + X^7 + X^6 + X + 1,$

► $X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1,$

► $X^{10} + X^7 + X^6 + X + 1,$

► $X^{14} + X^8 + X^6 + 1,$

► $X^{15} + X^7 + X^6 + X^4 + X + 1.$

4. Montrer en faisant des réductions modulo un nombre premier p bien choisi que les polynômes suivants à coefficients entiers sont irréductibles :

► $X^8 + 2X^7 + 3X^6 + X^5 + 4X^4 + 3X + 1,$

► $X^8 + X^7 + 2X^6 + 3X^4 + 3X^3 + 1,$

► $X^7 + 3X^6 + 4X^4 + X^2 + 2X + 2,$

► $X^6 + X^3 + X^2 + 1.$

5. a) Déterminer tous les automorphismes de $\mathbb{F}_{p^{nm}}$ laissant fixe le sous-corps $\mathbb{F}_{p^n}$.
- b) Soit $k = \mathbb{F}_{p^n}$ et soit P un polynôme irréductible à coefficients dans k de degré m . Montrer que le corps de rupture de P sur k est isomorphe à $\mathbb{F}_{p^{nm}}$.
- c) Donner une formule analogue à celle du cours pour calculer le nombre de polynômes unitaires, de degré m , irréductibles, à coefficients dans $\mathbb{F}_{p^n}$.
6. Soit p un nombre premier supérieur ou égal à 3. Soient u, v, w des éléments de $\mathbb{F}_p$ non tous nuls. Montrer que l'équation, en les inconnues x, y , $ux^2 + vy^2 = w$ a au moins une solution non nulle. On comptera le nombre d'éléments de la forme ux^2 , et le nombre d'éléments $vy^2 - w$, x et y décrivant $\mathbb{F}_p$.
7. Soit $P \in k[X]$, où $K = \mathbb{F}_q$ est un corps fini à q éléments. Supposons P irréductible. On appelle exposant de P le plus petit entier e tel que P divise $X^e - 1$. Soit k le degré de P .
- a) Montrer que P est scindé en facteurs de degré 1 dans $\mathbb{F}_{q^k}$ et que e est l'ordre d'une racine quelconque de P dans le groupe multiplicatif $\mathbb{F}_{q^k}^*$.
- b) Montrer que e divise $q^k - 1$.
- c) Montrer qu'un polynôme irréductible P d'exposant e divise $X^d - 1$ si et seulement si e divise d .
8. Calculer les exposants des polynômes suivants, dont on vérifiera qu'ils sont irréductibles, à coefficients dans $\mathbb{F}_2$
- $X^8 + X^4 + X^3 + X^2 + 1$,
 - $X^6 + X^3 + 1$,
 - $X^7 + X^4 + 1$.
9. a) ** Soit $k = \mathbb{F}_q$ un corps fini de caractéristique p . Soit n un entier premier à p . On suppose que k contient toutes les racines n -ièmes de l'unité, c'est-à-dire que le polynôme $X^n - 1$ a n racines dans k . Comme dans le cas de $\mathbb{C}$ on appelle racine primitive n -ième de l'unité une racine dont l'ordre dans le groupe multiplicatif k^* est n . On notera $\mu_n^*(k)$ l'ensemble des racines primitives n -ièmes de k . Soit $\Phi_{n,k} = \prod_{\zeta \in \mu_n^*(k)} (X - \zeta) \in k[X]$. Démontrer les propriétés correspondant à celles des polynômes cyclotomiques Φ_n . En particulier montrer que $\Phi_{n,k} \in \mathbb{F}_p[X]$.
- b) Montrer que la réduction modulo p de Φ_n est égale à $\Phi_{n,k}$.
- c) Montrer que la réduction modulo p du polynôme cyclotomique Φ_n se factorise en un produit de $\frac{\varphi(n)}{d}$ polynômes irréductibles à coefficients dans k , où d est l'ordre de q (le cardinal de k) dans le groupe $(\mathbb{Z}/n\mathbb{Z})^*$.
10. Soit P un polynôme irréductible de degré n , à coefficients dans $\mathbb{F}_p$. Soit d un diviseur de n , montrer que P factorise en $\frac{n}{d}$ polynômes irréductibles de même degré dans $\mathbb{F}_{p^d}[X]$.

11. ** On reprend la terminologie de l'exercice 7. Soit $k = \mathbb{F}_p$, p premier, calculer le nombre de polynômes irréductibles de degré n , unitaires et d'exposant e . On commencera par étudier la décomposition de la réduction modulo p du polynôme cyclotomique $\Phi_e(X)$.
12. Soit V un espace vectoriel de dimension fini d sur un corps fini k . Soit $(e_1, \dots, e_d)$ une base de V . Soit $v \in V$, et soit $(a_1, \dots, a_d)$ les coordonnées de v dans cette base.
- a) L'élément v étant fixé, montrer qu'il existe un polynôme à coefficients dans k en d variables qui prend la valeur 1 en $(a_1, \dots, a_d)$, et la valeur 0 partout ailleurs.
 - b) Soit de façon plus générale f une application de V dans k . Montrer qu'il existe un polynôme P en d variables tel que, pour tout $v \in V$, on ait $f(v) = P(a_1, \dots, a_d)$.
 - c) Le polynôme précédent est-il unique ?
13. Soit k un corps fini à q éléments de caractéristique p . Soit $a \in k$. Montrer que le polynôme $X^p - X - a$ est irréductible sur k si et seulement si il n'a pas de racines dans k .

14. Fonction d'Euler.

Soit k un corps fini à q éléments, de caractéristique p . Étant donné $P \in k[X]$ on définit $\Phi_k(P)$ comme étant le nombre de polynômes de degré inférieur à celui de P et premiers avec P .

a) Calculer

- $\Phi_k(X^p - 1)$ si $k = \mathbb{F}_p$,
- $\Phi_k(PQ)$ si P et Q sont premiers entre eux.

b) Soit $P_1^{\alpha_1} \cdots P_k^{\alpha_k}$ la décomposition d'un polynôme P de degré n en facteurs irréductibles. Montrer que

$$\Phi_k(P) = q^n \times \prod_i (1 - q^{-n_i}),$$

où n_i est le degré de P_i .

15. Théorème de König-Rados.

Soit k un corps fini à q éléments. Soit P un polynôme à coefficients dans k , $P = a_0 + a_1X + \cdots + a_{q-2}X^{q-2}$. Montrer que le nombre des racines dans k de l'équation $P(X) = 0$ est égal à $q - 1 - r$, où r est le rang de la matrice :

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{q-2} \\ a_1 & a_2 & \cdots & a_0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{q-2} & a_1 & \cdots & a_{q-2} \end{pmatrix}$$

16. (Clôture algébrique des corps finis)

Soit p un nombre premier, et soit k_n le corps $\mathbb{F}_{p^{n!}}$.

- a) Montrer que qu'il y a un sous-corps de k_n isomorphe à k_{n-1} et un seul. On notera i_{n-1} l'inclusion de k_{n-1} dans k_n .
- b) On pose $e_n = k_n - i_{n-1}(k_{n-1})$. Posons $l_n = \mathbb{F}_p \cup (\cup_{i=2, \dots, n} e_i)$. Montrer que l_n a une structure de corps, et qu'il est isomorphe à k_n .
- c) Montrer que $l = \mathbb{F}_p \cup (\cup_{2 \geq i} e_i)$ est un corps et est une clôture algébrique de $\mathbb{F}_p$.

III. Codes linéaires

1. On appelle code de Hamming tout code de longueur $n = 2^m - 1$, $m \geq 2$ et de dimension $2^m - m - 1$, dont les colonnes de la matrice de parité sont les coefficients de la décomposition en base 2 des entiers $1, \dots, 2^m - 1$.

Démontrer que les codes de Hamming corrigent jusqu'à 1 erreur, c'est-à-dire que leur distance minimale est 3.

2. Borne de Plotkin. Soit $\mathbb{K}$ un corps fini à q éléments, et soit C un code de longueur n et de dimension h . Soit δ_C la distance minimale du code. Démontrer que

$$\delta_C \leq \frac{nq^{h-1}(q-1)}{q^h - 1}.$$

3. Borne de Gilbert-Varshamov. Soit $\mathbb{K}$ un corps fini à q éléments. Démontrer qu'il existe un code linéaire de longueur n et de dimension h de distance minimale supérieure ou égale à un entier d donné dès que :

$$q^{n-h} > \sum_{i=0}^{d-2} C_{n-1}^i (q-1)^i.$$

4. Étude d'un code BCH.

Soit μ une racine du polynôme $X^5 + X^3 + 1$ dans une extension du corps $\mathbb{F}_2$. Montrer que son ordre est 31.

- a) On considère le code BCH de longueur 31 déterminé par les racines $\mu, \mu^2, \dots, \mu^9$. Quelle est sa distance minimale ? Calculer son polynôme générateur et son polynôme de parité.
- b) On considère le code BCH déterminé par les racines $\mu, \mu^2, \dots, \mu^{11}$. Quelle est sa longueur, sa distance minimale ? Calculer son polynôme générateur et son polynôme de parité.
- c) Dans le cas de la question 4.a) on considère le message suivant qui a été reçu : $X^7 + X^{10} + X^{11} + X^{13} + X^{14} + X^{16} + X^{18}$. Quelles étaient les données initiales ? Même question avec $1 + X^2 + X^7 + X^{10} + X^{11} + X^{13} + X^{14} + X^{16} + X^{18} + X^{20}$.

5. Étude d'un code cyclique.

Soit μ une racine du polynôme $X^6 + X^5 + X^4 + X^2 + 1$ dans une extension du corps $\mathbb{F}_2$. Montrer que son ordre est 21.

a) On considère le code cyclique de longueur 10 déterminé par les racines

$$\mu, \mu^2, \mu^3, \mu^4, \mu^6, \mu^7, \mu^8, \mu^{11}, \mu^{12}, \mu^{14}, \mu^{16}.$$

Quelle est sa distance minimale ?

Calculer son polynôme générateur et son polynôme de parité.

b) Calculer les données initiales à partir des données reçues suivantes :

$$X + X^3 + X^8 + X^{10} + X^{11}, \quad X + X^9 + X^{12}.$$

6. (Étude d'un code de Goppa) On considère le polynôme $G = X^2 + X + 1$ et le corps $L = \mathbb{F}_8$, α désigne un générateur du groupe multiplicatif dont le polynôme minimal est $X^3 + X + 1$.

a) Montrer que G est irréductible sur L .

b) Calculer la matrice de parité du code de Goppa de longueur 8 sur $\mathbb{F}_2$ déterminé par G et l'ensemble des éléments de L . L'exprimer sur le corps $\mathbb{F}_2$. Calculer sa dimension et sa distance minimum.

c) Montrer que ce code est cyclique.

d) Faire le même travail avec $G = X^3 + X + 1$ et $L = \mathbb{F}_{16}$.

Quelques réponses ou indications

I. Éléments algébriques, Polynômes cyclotomiques

1. On commence par calculer le degré du polynôme. On trouve 4, 4, 8 et 10. Pour calculer ces degrés on utilise le lemme de la base télescopique. Puis, par exemple dans le premier cas, on pose $\alpha = \sqrt{2} + \sqrt{7}$, soit $\alpha - \sqrt{2} = \sqrt{7}$. Donc $(\alpha - \sqrt{2})^2 = 7$, et $\alpha^2 - 5 = 2\alpha\sqrt{2}$, et on élève cette dernière relation au carré.

2. Utiliser la formule donnant la somme des n premiers termes d'une progression géométrique.

3. On considérera la différence

$$P\left(\frac{p}{q}\right) - P(\alpha).$$

En appliquant le théorème des accroissements finis on obtient

$$P\left(\frac{p}{q}\right) - P(\alpha) = \left(\frac{p}{q} - \alpha\right)P'(\theta).$$

Si $\frac{p}{q}$ est assez proche, et différent de α , $P\left(\frac{p}{q}\right)$ est non nul. Par ailleurs $P\left(\frac{p}{q}\right)$ est non nul et peut s'écrire sous la forme $\frac{c}{q^n}$ où c est un entier non nul. On a donc

$$\frac{c}{q^n} = \left(\frac{p}{q} - \alpha\right)P'(\theta).$$

On obtiendra l'inégalité sur un voisinage de α en majorant $|P'(\theta)|$, pour ce faire on montrera que $P'(\alpha)$ est non nul.

4. Nombre de Liouville.

Montrer que le réel $\sum_n \frac{1}{10^{n!}}$ ne satisfait pas au critère précédent. On majorera le reste de la série. En conclure qu'il est transcendant.

5. Dans le premier cas on utilisera la formule

$$\frac{X^n - 1}{X - 1} = \prod_{d|n, d \neq 1} \Phi_d(X),$$

et un raisonnement par récurrence. On distinguera également suivant que n est puissance d'un nombre premier ou non. Si p est un nombre premier on trouve $\Phi_{p^k}(1) = p$, et on trouve la valeur 1 si n est divisible par au moins deux nombres premiers distincts.

On adaptera pour le second cas. En particulier on montrera que $\Phi_2(-1) = 0$, $\Phi_{2^h}(-1) = 1$ si $h > 1$, $\Phi_k(-1) = 1$, k impair, et que $\Phi_{2p^h}(-1) = 1$ si p est premier impair.

6,7. Pour les polynômes cyclotomiques Φ_9 , Φ_{27} , le résultat résulte de l'exercice 7. Cet exercice se démontre à partir de la formule :

$$X^n - 1 = \prod_{d|n} \Phi_d(X),$$

appliquée dans le cas $n = p^h$. On observera que $\Phi_{p^{h-1}}(X)$ est facteur de $X^{p^h} - 1$.

8. Appliquer la définition, c'est-à-dire comparer l'ensemble des racines primitives $2k+1$ -ièmes à l'ensemble des racines primitives $2(2k+1)$ -ièmes.

9. D'abord, disons qu'il y a une relation linéaire entre les homomorphismes c_i si, pour tout $m \in \mathbb{Z}$, on a :

$$\sum a_i c_i(m) = 0.$$

Si les coefficients a_i sont non tous nuls, on dira que la relation linéaire est non-triviale. Il faut montrer qu'il n'y a pas de relations linéaires non-triviales.

Si on procède par récurrence sur le nombre d'homomorphismes, on suppose le résultat démontré pour $n-1$ homomorphismes. Puis on raisonne par l'absurde, on suppose avoir une relation linéaire non-triviale entre n homomorphismes. On fixe alors un entier h , et soit m un entier quelconque. On considère alors les relations obtenues en évaluant la relation linéaire en h , et en $m+h$. En multipliant la première par un élément adéquat et en la soustrayant à la seconde on obtiendra une relation linéaire non-triviale entre $n-1$ homomorphismes et donc une contradiction.

Sinon en utilisant un déterminant de Van der Monde bien choisi on montrera que les coefficients d'une relation linéaire sont solutions d'un système linéaire homogène de rang maximal, donc sont nuls.

10. a) Il y a $p-1$ éléments qui sont des carrés dans $\mathbb{F}_p^*$ et $p-1$ éléments qui ne le sont pas.

b) Calculer le carré :

$$s^2 = \sum_{x, z \in \mathbb{F}_p^*} \left(\frac{xz}{p}\right) \mu^{x+z}.$$

Puis poser $z = xy^{-1}$. Puis montrer que le premier terme de la somme :

$$s^2 = \sum_{x, y \in \mathbb{F}_p^*, y \neq -1} \left(\frac{y}{p}\right) \mu^{x(y+1)} + \sum_{x \in \mathbb{F}_p^*} \left(\frac{-1}{p}\right),$$

est nul. Utiliser l'exercice 2.

11. Chercher des cas modulo 2, 3, ...

12. a) b) Utiliser entre autres le critère utilisant la dérivation pour tester l'existence de racines multiples.

c) Montrer qu'en élevant à des puissances de nombres premiers bien choisis on prouve que toutes les racines primitives sont racines de P_1 . On pensera aux générateurs du groupe multiplicatif de $\mathbb{Z}/n\mathbb{Z}$.

13. a) b) Utiliser le théorème de la base télescopique.

c) d) Écrire les polynômes minimaux des éléments concernés. On trouve $8X^3 - 6X - 1$ et $X^3 - 2$. On vérifiera qu'ils sont irréductibles sur $\mathbb{Q}$.

e) Montrer que l'on se ramène à l'équation $X^4 + X^3 + X^2 + X + 1 = 0$ qui se ramène à une équation du second degré par le changement de variables $Y = X + \frac{1}{X}$.

14. Montrer que le coefficient est premier à 2 si et seulement toutes les puissances de 2 qui apparaissent dans la décomposition en base 2 de k apparaissent aussi dans celle de n . On écrira la formule de Newton comme suit, si $k = 2^{a_1} + \dots + 2^{a_h}$:

$$(X + Y)^k = \prod_i (X + Y)^{2^{a_i}} = \prod_i (X^{2^{a_i}} + Y^{2^{a_i}}).$$

II. Corps finis

1. Dans ce cas, on vérifie l'irréductibilité en vérifiant que les polynômes en question ne sont pas divisibles par des polynômes irréductibles de degré 1, 2, et 3. Il suffit de connaître la liste de ces polynômes : X , $X + 1$, $X^2 + X + 1$, $X^3 + X + 1$ et $X^3 + X^2 + 1$.

Pour ce qui est de trouver un générateur, dans le premier cas on doit trouver un générateur d'un groupe cyclique d'ordre 63, il suffit donc de trouver un élément qui ne soit ni d'ordre 7 ni d'ordre 9. Dans le second on doit trouver un générateur d'un groupe d'ordre 127, 127 est premier donc tout élément différent de l'élément neutre est générateur.

2. L'exercice est analogue au précédent mais en caractéristique 3.

4. On travaillera avec $p = 2$ et $p = 3$.

5. a) Montrer que sont les automorphismes de la forme $(F^n)^\lambda$.

b) Utiliser les résultats généraux sur les corps finis.

c) Posons $q = p^n$. Si $\psi_q(m)$ est le nombre de polynômes unitaires, de degré m , irréductibles, à coefficients dans $\mathbb{F}_q$ montrer que :

$$\sum_{d|m} d\psi_q(m) = q^m.$$

6. On compte le nombre d'éléments de la forme ux^2 , et le nombre d'éléments $vy^2 - w$, x et y décrivant $\mathbb{F}_p$. Il y en a $\frac{p+1}{2}$. Et on montre que l'intersection des deux ensembles est non-vide.

7. a) b) Appliquer la définition et les propriétés de l'ordre d'un élément dans un groupe.

c) On rappelle que $X^e - 1$ divise $X^f - 1$ si et seulement si e divise f .

8. 255, 9 et 127.

9. a) et b) On commencera par montrer que l'on a la relation $X^n - 1 = \prod_{d|n} \Phi_{d,k}$.
- c) On montrera que le degré du polynôme minimal dans k d'une racine primitive n -ième de 1 est l'entier d défini dans l'énoncé.
10. Soit H un facteur irréductible apparaissant dans la décomposition de P comme polynôme à coefficient dans $\mathbb{F}_{p^d}$. Montrer que $\frac{n}{d}$ divise le degré de H en observant que le corps de rupture de H contient $\mathbb{F}_{p^n}$.
Puis appliquer un résultat analogue à la proposition 3 de la section 5.
11. On considère la réduction du polynôme cyclotomique Φ_e . On remarquera que e est premier à p . On utilisera l'exercice 9.
Le nombre de polynômes irréductibles unitaires de degré m et d'exposant e à coefficients dans $\mathbb{F}_q$ est $\frac{\varphi_p(e)}{m}$ si m est l'ordre de q modulo e , 2 si $m = e = 2$, 0 sinon.
12. a)
- $$\prod_{i=1, \dots, d} \left(\prod_{\alpha \neq a_i} (X_i - \alpha) \right) + 1.$$
- 1 en $(a_1, \dots, a_d)$
- c) Non, si k a q éléments on peut le remplacer par sa puissance q -ième.
13. Observer que si α est racine d'un facteur irréductible $\alpha + 1, \alpha + 2, \dots$ l'est aussi.
15. Théorème de König-Rados. Considérer le produit de la matrice introduite par une matrice de Van der Monde construite à partir de q éléments distincts de $\mathbb{K}$.

III. Codes linéaires

- Deux colonnes quelconques de la matrice de parité sont linéairement indépendantes.
- Majorer la somme des poids de tous les éléments du code, puis diviser par le nombre d'éléments non nuls du code. Enfin la distance minimale est aussi le poids minimum non nul d'un élément. Pour majorer la somme des poids on tiendra compte des éléments du code qui sont situés dans les hyperplans de coordonnées et leurs intersections.
- On construit une matrice de parité dans laquelle $d - 1$ colonnes sont toujours linéairement indépendantes. Pour ce faire, on commence par choisir une colonne quelconque, pour la seconde on élimine toutes les colonnes linéairement dépendantes. On itère le processus en supposant avoir choisi k colonnes satisfaisant à la propriété. Pour en trouver $k + 1$ colonnes on doit éliminer toutes les combinaisons linéaires de $d - 2$ colonnes parmi les k précédentes.
- Voici la liste des polynômes irréductibles de degré 5 sur $\mathbb{F}_2$; $X^5 + X^2 + 1, X^5 + X^4 + X^2 + X + 1, X^5 + X^3 + 1, X^5 + X^4 + X^3 + X^2 + 1, X^5 + X^3 + X^2 + X + 1, X^5 + X^4 + X^3 + X + 14$. Le second est le polynôme minimal de μ^{11} , le troisième de μ, μ^2, μ^4, μ^8 , le quatrième de μ^7 , le cinquième de μ^3, μ^6 , enfin le dernier de μ^5, μ^9, μ^{10} . Ceci permet de calculer le polynôme générateur.

On trouve 4 erreurs dans le premier cas 1 dans le second. Les données transmises étaient $1 + X^2 + X^3 + X^7 + X^{10} + X^{11} + X^{13} + X^{14} + X^{16} + X^{18} + X^{20}$.

5. On vérifie que le polynôme $X^6 + X^5 + X^4 + X^2 + 1$ divise $X^{21} + 1$. Les polynômes minimaux de μ, μ^3, μ^7 sont $X^6 + X^5 + X^4 + X^2 + 1, X^3 + X + 1, X^2 + X + 1$.
Le polynôme générateur est $1 + X^2 + X^7 + X^8 + X^{11}$
5. b) Il y a 4 erreurs dans le premier cas, 2 dans le second. Les données transmises étaient $X + X^3 + X^8 + X^9 + X^{12}$ dans les deux cas.
6. a) b) On trouve d'abord la matrice :

$$\begin{pmatrix} 1 & 1 & \alpha^2 & \alpha^2 + \alpha & \alpha^2 & \alpha & \alpha & \alpha^2 + \alpha \\ 0 & 1 & \alpha + 1 & \alpha^2 + 1 & \alpha^2 + \alpha + 1 & \alpha^2 + \alpha + 1 & \alpha^2 + 1 & \alpha + 1 \end{pmatrix}$$

soit :

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Le code est de dimension 2, sa distance minimale est 5 (appliquer la proposition 3 section 10 et calculer le rang de la matrice).

Dans le second exemple on choisit un générateur α de polynôme minimal $X^5 + X^2 + 1$, et on trouve comme dimension 17 et distance minimale est 7, on renvoie à [MS] pour plus de détails sur ces exemples.

Chapitre 5

Réductions des endomorphismes Structure du groupe linéaire

Ce chapitre est consacré en premier lieu à la réduction des endomorphismes. Le cas fondamental est celui où le corps de base est le corps des complexes $\mathbb{C}$. Une bonne part des arguments s'étend sans problème suivant le cas, soit à un corps quelconque, soit à un corps algébriquement clos quelconque. Nous nous placerons donc dans ces cas. Le lecteur pourra se restreindre au cas de $\mathbb{R}$ ou $\mathbb{C}$ dans une première approche. La première section sera consacrée à l'étude du polynôme minimal, du théorème des noyaux, puis à la diagonalisation. La seconde section est consacrée à la démonstration du théorème de Cayley-Hamilton, décrit les matrices compagnons et donne sans démonstration les propriétés des invariants de similitude. La troisième section expose la triangularisation et la réduction de Jordan. Dans la quatrième on donne des applications topologiques des précédents résultats.

En second lieu on étudiera la structure du groupe linéaire. On décrira en particulier un ensemble de générateurs. On déterminera le centre, et les sous-groupes distingués.

1. Polynôme minimal d'un endomorphisme

Soit k un corps quelconque et soit E un espace vectoriel de dimension finie sur k . Soit $f \in \mathcal{L}(E)$ une application linéaire de E dans lui-même. Définissons une application

$$\text{ev} : k[X] \longrightarrow \mathcal{L}(E)$$

par :

$$P = \sum_i a_i X^i \longmapsto \sum_i a_i f^i.$$

Dans la suite $\sum_i a_i f^i$ sera noté $P(f)$. Dans cette formule f^0 désigne l'identité Id . Concrètement, dans le polynôme, à la variable X on substitue l'application linéaire f . On a (proposition 2, section 5, chapitre 3) :

Proposition 1. *L'application ev est un homomorphisme d'anneaux.*

Démonstration. Refaisons brièvement la démonstration.

Si $P = \sum_i a_i X^i$ et $Q = \sum_j b_j X^j$ on a :

$$\text{ev}(P + Q) = \sum_i (a_i + b_i) f^i = \sum_i a_i f^i + \sum_i b_i f^i = \text{ev}(P) + \text{ev}(Q).$$

De même on calcule :

$$\text{ev}(PQ) = \sum_{\ell+k=i} a_\ell b_k f^i = \left(\sum_\ell a_\ell f^\ell \right) \left(\sum_k b_k f^k \right) = \text{ev}(P) \text{ev}(Q).$$

Ce qui montre que ev est un homomorphisme d'anneaux. ■

Notons que pour tous P et Q dans $k[X]$ on a $P(f)Q(f) = Q(f)P(f)$.

Le noyau de l'application ev est donc un idéal. Comme l'anneau $k[X]$ est principal cet idéal est engendré par un polynôme qui est uniquement défini si lui on impose d'être unitaire, (i.e. que le coefficient de son terme de degré dominant soit égal à 1) :

Définition 1

Ce polynôme est appelé par définition le polynôme minimal de l'application linéaire f .

Le polynôme minimal de f est donc le polynôme unitaire non nul de degré minimal annulant f .

Tout ce que l'on vient de dire se transcrit aussitôt au cas des matrices carrées. Soit A une matrice carrée. Considérons l'application qui à un polynôme $P \in k[X]$, $P = \sum_i a_i X^i$ associe la matrice $\sum_i a_i A^i$. C'est un homomorphisme d'anneaux de

$k[X]$ dans l'anneau des matrices carrées $M_n(k)$. Le noyau de cet homomorphisme est un idéal principal qui admet pour générateur un polynôme que l'on appellera polynôme minimal de la matrice A .

Revenons au cas d'un espace vectoriel E et d'un endomorphisme f . Si on choisit une base de E et si A est la matrice associée à f dans cette base le polynôme minimal de f coïncide avec le polynôme minimal de la matrice A .

Soit f une application linéaire de E dans E et soit $P \in k[X]$. Supposons que P s'écrive comme un produit RS où R et S sont des polynômes à coefficients dans k premiers entre eux. Notons E_0 le noyau de l'application linéaire $R(f)$ et E_1 le noyau de l'application linéaire $S(f)$.

Proposition 2. Avec les notations précédentes on a :

- Les sous-espaces vectoriels E_0 et E_1 sont stables par f , i.e. $f(E_0)$ est contenu dans E_0 et $f(E_1)$ est contenu dans E_1 ,
- le noyau de $P(f)$, $\text{Ker}(P(f))$ est somme directe de E_0 et E_1 ,
- si P est le polynôme minimal de f , alors $E = E_0 \oplus E_1$. Le polynôme minimal de la restriction de f à E_0 est R , et le polynôme minimal de la restriction de f à E_1 est S .

Démonstration. Montrons par exemple que E_0 est stable par f . On doit montrer que si on a un élément $x \in E$ tel que $R(f)(x) = 0$ alors on a aussi $R(f)(f(x)) = 0$. Mais on a $R(f)(f(x)) = f(R(f))(x) = 0$, ce qui démontre le résultat.

Comme R et S sont premiers entre eux d'après l'identité de Bézout il existe des polynômes à coefficients dans k , soient U et V , tels que

$$UR + VS = 1.$$

En substituant f à la variable X on obtient

$$U(f)R(f) + V(f)S(f) = \text{Id}.$$

En particulier pour tout $x \in E$ on a

$$x = U(f)R(f)(x) + V(f)S(f)(x).$$

Par ailleurs $U(f)R(f)(x) \in E_1$ car on a :

$$S(f)(U(f)R(f))(x) = U(f)R(f)S(f)(x) = U(f)P(f)(x) = 0.$$

De même on a $V(f)S(f)(x) \in E_0$, ceci montre que E est somme de E_0 et E_1 .

Pour montrer que la somme est directe il faut démontrer que l'intersection de E_0 et E_1 est réduite à $\{0\}$. Soit donc $x \in E_0 \cap E_1$. Considérons l'identité

$$x = U(f)R(f)(x) + V(f)S(f)(x),$$

les deux termes de droite sont nuls. Ceci montre que $x = 0$ et achève cette partie de la démonstration.

Il reste à démontrer que le polynôme minimal de f restreint à E_0 (resp. E_1) est égal à R (resp. S). Soit R' (resp. S') le polynôme minimal de f restreint à E_0 (resp. E_1). Le polynôme R' divise le polynôme R car $R(f)$ est nul sur E_0 par hypothèse. De même le polynôme S' divise le polynôme S .

Comme d'un autre côté $R'(f)S'(f)$ est nul sur $E = E_0 \oplus E_1$, RS divise $R'S'$. Comme on a aussi $R'S'$ divise RS on en déduit que $R = R'$ et $S = S'$. ■

Ce résultat est connu sous le nom de *théorème des noyaux*. Il se généralise comme suit :

Théorème 1. Soit E un espace vectoriel de dimension finie sur un corps k , et soit $f \in \mathcal{L}_k(E)$ une application linéaire. Soit $P \in k[X]$ un polynôme. Supposons que $P = P_1 \cdots P_i \cdots P_h$, où les polynômes $P_i \in k[X]$ sont premiers entre eux deux à deux. Soit K_i le noyau de $P_i(f)$, alors :

- les sous-espaces vectoriels K_i sont stables par f ,
- le noyau de l'application linéaire $P(f)$, $\text{Ker } P(f)$ est somme directe des sous-espaces vectoriels K_i ,
- si P est le polynôme minimal de f l'espace vectoriel E est somme directe des K_i , et le polynôme minimal de la restriction de f à K_i est P_i .

Démonstration. Introduisons les polynômes $Q_i = \prod_{j \neq i} P_j$. Comme les polynômes P_i sont premiers entre eux deux à deux les polynômes Q_i sont premiers dans leur ensemble.

D'après l'identité de Bézout il existe des polynômes $U_1, \dots, U_h$ tels que

$$U_1 Q_1 + \cdots + U_h Q_h = 1.$$

En particulier si on substitue f à X on obtient :

$$\sum_i U_i(f) Q_i(f) = \text{Id},$$

donc en particulier pour tout $x \in E$ on a :

$$x = \sum_i U_i(f) Q_i(f)(x).$$

Mais on a :

$$P_i(f) U_i(f) Q_i(f)(x) = U_i(f) P_i(f) Q_i(f)(x) = Q_i(f) P(f)(x) = 0,$$

donc pour tout i on a :

$$U_i(f) Q_i(f)(x) \in \text{Ker}(P_i(f)) = K_i.$$

Ce qui implique que E est somme des sous espaces vectoriels K_i .

Il reste à montrer que la somme est directe. Conservant les notations précédentes notons π_i l'application linéaire $U_i(f) Q_i(f)$.

Lemme 1. On a $\pi_i \pi_j = 0$ si $i \neq j$ et $\pi_i^2 = \pi_i$.

Démonstration. Considérons la première assertion. On a

$$\pi_i \pi_j = U_i(f) Q_i(f) U_j(f) Q_j(f) = U_i(f) U_j(f) Q_i(f) Q_j(f).$$

Par construction des polynômes Q_i le polynôme P divise $Q_i Q_j$, si $i \neq j$, donc $Q_i(f) Q_j(f) = 0$, et donc $\pi_i \pi_j = 0$.

Pour la seconde relation on considère la formule $\text{Id} = \sum_j \pi_j$ et on la multiplie par π_i (i fixé) :

$$\pi_i = \left(\sum_j \pi_j \right) \pi_i = \sum_j \pi_i \pi_j = \pi_i^2. \quad \square$$

Montrons que le sous-espace vectoriel K_i est aussi l'image L_i de l'application linéaire π_i . On vient de voir que $L_i \subset K_i$, en effet par définition de π_i on a $P_i \circ \pi_i = 0$. Inversement si $x \in K_i$ on a $x = \pi_i(x)$, car pour tout $j \neq i$ on a $\pi_j \circ \pi_i(x) = 0$. Montrons que la somme est directe. Soit que si on a $0 = \sum_j v_j$ avec $v_j = \pi_j(w_j)$ avec $w_j \in E$, pour tout j on a $v_j = 0$. Si on applique à cette formule π_i pour un i donné, comme plus haut on obtient $0 = \pi_i(v_i) = \pi_i^2(w_i) = \pi_i(w_i) = v_i$. Le résultat suit.

La dernière partie du théorème se démontre comme plus haut. ■

Définition 2

Soit E un espace vectoriel de dimension finie sur un corps k , et soit f une application linéaire. On dit que $\lambda \in k$, est *valeur propre* de f s'il existe un élément non nul $v \in E$ tel que $f(v) = \lambda v$. On dit que v est un *vecteur propre* associé à la valeur propre λ si $f(v) = \lambda v$.

La somme de deux vecteurs propres associés à la valeur propre λ est un vecteur propre associé à la valeur propre λ . De même si v est vecteur propre associé à la valeur propre λ , il en est de même pour av pour tout $a \in k$. Notons que, avec cette définition, 0 est toujours vecteur propre.

Proposition 3. Soit E un espace vectoriel sur un corps k , et soit f une application linéaire. Soit $\lambda \in k$ une valeur propre de f . L'ensemble des vecteurs propres de f associés à la valeur propre λ est un sous-espace vectoriel de E appelé le sous-espace propre associé à la valeur propre λ . Il est noté E_λ .

Soient $(\lambda_1, \dots, \lambda_n)$ les valeurs propres de f , et soient $E_{\lambda_1}, \dots, E_{\lambda_n}$ le sous-espaces propres associés.

Lemme 2. Les sous-espaces propres sont en somme directe.

Démonstration. En effet, supposons avoir une relation du type

$$\sum_{1 \leq i \leq k} v_i = 0,$$

où v_i est un vecteur propre associé à la valeur propre λ_i de f . Les valeurs propres λ_i sont deux à deux distinctes. Il faut montrer que tous les v_i sont nuls. On va montrer que si on a une relation comportant k termes non nuls on peut obtenir une relation du même type comportant $k - 1$ termes non nuls. Par récurrence descendante on obtiendra une contradiction.

Chaque v_i est donc non nul par hypothèse. Écrivons la relation sous la forme $v_1 = -\sum_{2 \leq i \leq k} v_i$. Appliquons f , on obtient $\lambda_1 v_1 = -\sum_{2 \leq i \leq k} \lambda_i v_i$. En multipliant par λ_1 la première relation et en soustrayant on obtient la relation :

$$\sum_{2 \leq i \leq k} (\lambda_i - \lambda_1) v_i = 0.$$

Chacun des termes de cette relation est non nul. Par itération on obtient :

$$\prod_{1 \leq i \leq k-1} (\lambda_i - \lambda_k) v_k = 0,$$

donc $v_k = 0$ en contradiction avec les hypothèses. ■

Définition 3

Si la somme des sous-espaces propres est égale à l'espace E l'application linéaire f est dite diagonalisable.

Justifions cette terminologie. Soient E_{λ_i} les sous-espaces propres de f . Si on choisit une base de E qui est réunion de bases des E_{λ_i} la matrice de f dans cette base est une matrice diagonale.

Cette définition s'applique évidemment aussi à une matrice (n, n) identifiée à une application linéaire de k^n dans k^n .

Le théorème suivant donne une caractérisation des applications linéaires diagonalisables.

Théorème 2. Soit E un espace vectoriel de dimension finie sur un corps k , et soit f une application linéaire.

- Supposons que le polynôme minimal soit scindé sur k et que les racines soient de multiplicité 1. Alors E est somme directe des sous-espaces propres de f . Inversement si E est somme directe des sous-espaces propres de f le polynôme minimal de f a toutes ses racines dans k et elles sont de multiplicité 1.

- Pour qu'un endomorphisme soit diagonalisable il faut et suffit qu'il soit annulé par un polynôme scindé en produit de facteurs du premier degré et ayant toutes ses racines simples.

Démonstration. Commençons par la seconde partie. C'est une application immédiate du théorème des noyaux. Soient λ_i les racines du polynôme scindé en facteurs du premier degré, ayant toutes ses racines simples qui annule f . On peut alors lui appliquer le théorème des noyaux. L'espace vectoriel E est somme directe des sous-espaces $\text{Ker}(f - \lambda_i \text{Id})$. Si λ_i est une valeur propre le noyau $\text{Ker}(f - \lambda_i \text{Id})$ est le sous-espace propre E_{λ_i} . Il est nul si λ_i n'est pas valeur propre. Le résultat suit.

Réciproquement supposons f diagonalisable. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de f . Soit $v \in E$, par hypothèse v peut s'écrire sous la forme $v = \sum_i v_i$ où $v_i \in E_{\lambda_i}$. Montrons que l'endomorphisme $\prod_j (f - \lambda_j \text{Id})$ est nul. On l'applique à v , on a

$$\prod_j (f - \lambda_j \text{Id})(v_i) = \prod_{j \neq i} (f - \lambda_j \text{Id})(f - \lambda_i \text{Id})(v_i) = 0.$$

Le résultat suit.

La première partie du théorème suit aussitôt en prenant pour P le polynôme minimal. ■

Ce résultat se transcrit immédiatement aux matrices.

Soient f et g deux applications linéaires diagonalisables, le résultat suivant indique quand ces applications sont simultanément diagonalisables.

Proposition 4. Soit E un espace vectoriel sur un corps k . Soient f et g deux applications linéaires qui commutent : $fg = gf$. Supposons que f et g soient diagonalisables, alors elles sont diagonalisables dans une base commune.

Démonstration. Notons E_λ les sous-espaces propres de f . On va montrer qu'ils sont stables par g . Soit en effet $v \in E_\lambda$, on a alors $f(g(v)) = g(f(v)) = g(\lambda v) = \lambda g(v)$, soit $g(v) \in E_\lambda$.

Montrons maintenant que la restriction de g à E_λ est diagonalisable. Comme g est diagonalisable elle est annulée par un polynôme ayant toutes ses racines dans le corps de base, celles ci étant de multiplicité 1. Ce polynôme annule aussi la restriction de g à E_λ qui est donc diagonalisable.

On peut aussi facilement démontrer que E_λ est somme directe des $E_\lambda \cap F_\mu$, où les F_μ sont les sous-espaces propres de g . ■

Corollaire. Soit E un espace vectoriel de dimension finie sur un corps k , et soient f_i , $i \in I$ un ensemble d'applications linéaires qui commutent deux à deux : $f_i f_j = f_j f_i$, pour tout i et tout j . Supposons que les applications linéaires f_i soient diagonalisables, alors elles sont diagonalisables dans une base commune.

Démonstration. Supposons avoir démontré le résultat pour un ensemble de n applications linéaires commutant deux à deux. Démontrons maintenant le résultat pour un ensemble de $n + 1$ applications linéaires $f_1, \dots, f_{n+1}$ commutant deux à deux. L'espace vectoriel E est somme directe des sous-espaces propres de f_{n+1} . Ceux-ci sont stables par $f_1, \dots, f_n$, on peut donc appliquer l'hypothèse de récurrence à la famille $f_1, \dots, f_n$ sur chacun de ces sous-espaces propres. Notons que chaque f_i est diagonalisable sur un sous-espace propre de f_{n+1} à cause du même argument que plus haut. On trouve une base de chacun des sous-espaces propres de f_{n+1} où chaque f_i , $1 \leq i \leq n$ a une matrice diagonale. Le résultat suit car f_{n+1} a bien évidemment une matrice diagonale dans cette base.

Le résultat s'étend à une famille infinie d'applications linéaires pourvu que l'on suppose E de dimension finie (exercice). ■

2. Théorème de Cayley-Hamilton

La section précédente n'a pas donné de moyens de calculer les vecteurs propres. C'est ce que l'on va rappeler maintenant.

Introduisons une définition. On rappelle que le *polynôme caractéristique* d'une matrice de taille (n, n) , M , est le déterminant $\det(M - X I_n)$. C'est un élément de $k[X]$. Dans cette formule I_n est, comme d'habitude, la matrice unité, diagonale (n, n) avec des 1 sur la diagonale. Soit f une application linéaire et M la matrice de f dans une base B . On appelle polynôme caractéristique de f , le déterminant $\det(M - X I_n)$. Ce polynôme ne dépend pas de la base choisie. En effet si on effectue un changement de base de matrice de passage P le nouveau polynôme est

$$\det(P^{-1}MP - X I_n) = \det(P^{-1}) \det(M - X I_n) \det(P) = \det(M - X I_n).$$

On le notera donc $\det(f - X \text{Id}) \in k[X]$.

Définition 1

Soit E un espace vectoriel de dimension finie sur un corps k , et soit f une application linéaire. On appelle polynôme caractéristique de f le polynôme $\det(f - X \text{Id}) \in k[X]$.

Le degré du polynôme caractéristique est la dimension de l'espace E .

On a :

Proposition 1. Soit E un espace vectoriel sur le corps k , et soit f une application linéaire de E dans E . Un élément $\lambda \in k$ est valeur propre de f si et seulement si il est racine du polynôme caractéristique de f .

Démonstration. Un élément $\lambda \in k$ est valeur propre de f si et seulement si l'application linéaire $f - \lambda \text{Id}$ n'est pas injective. Autrement dit si et seulement si $\det(f - \lambda \text{Id})$ est nul. Donc si et seulement si λ est racine du polynôme caractéristique $\det(f - X \text{Id})$. ■

Comme tout polynôme non constant a au moins une racine dans un corps algébriquement clos on a :

Corollaire. Soit E un espace vectoriel de dimension finie, non nulle, sur un corps k algébriquement clos. Une application linéaire de E dans lui-même a toujours au moins une valeur propre.

Soit f une application linéaire d'un espace vectoriel sur un corps k , de dimension finie sur k , dans lui-même. Si elle a toutes ses valeurs propres dans le corps k et deux à deux distinctes, le polynôme caractéristique est égal au polynôme minimal. Cette propriété a lieu dès que les racines du polynôme caractéristique dans une clôture algébrique du corps de base sont deux à deux distinctes.

Les commentaires précédents peuvent être démontrés directement ou être vus comme une conséquence du *théorème de Cayley-Hamilton* qui suit :

Théorème 1. Soit $f \in \mathcal{L}(E)$. Le polynôme minimal de f divise le polynôme caractéristique de f .

Démonstration. Notons $\Delta(X)$ le polynôme caractéristique de f : $\det(f - X \text{Id})$. Choisissons une base pour E . Soit A la matrice représentant f dans la base choisie. Nous allons travailler dans les matrices à coefficients dans l'anneau des polynômes $k[X]$.

Notons $B(X)$ la comatrice (à coefficients dans $k[X]$) de la matrice $A - X I_n$, où n est la dimension de E . Le terme $b_{j,i}$ de la matrice $B(X)$ est obtenu comme suit : on enlève la i -ième colonne et la j -ième ligne à la matrice $A - X I_n$, on calcule le déterminant de cette matrice de taille $(n-1, n-1)$ et on le multiplie par $(-1)^{i+j}$.

Les formules de Cramer pour l'inverse d'une matrice expriment que :

$$(A - X I_n)B(X) = \Delta(X)I_n.$$

Le degré des coefficients de la matrice $B(X)$ en tant que polynômes à coefficients dans k est au plus $n-1$. En effet, en considérant une matrice obtenue à partir de $A - X I_n$ en enlevant une ligne et une colonne on obtient une matrice dont au plus

$n - 1$ termes sont des polynômes en X de degré 1. Tous les autres termes sont constants. Les termes de plus haut degré que l'on peut obtenir en développant sont produits de $n - 1$ polynômes de degré 1, donc de degré $n - 1$. Le déterminant comme polynôme est donc de degré au plus $n - 1$.

Écrivons donc $B(X)$ sous la forme

$$B_0 + XB_1 + \cdots + X^{n-1}B_{n-1},$$

où les matrices B_i sont à coefficients dans k . Le polynôme caractéristique $\Delta(X)$ s'écrit sous la forme

$$c_0 + c_1X + \cdots + (-1)^n X^n$$

où les c_i sont des éléments de k . On a donc la relation suivante :

$$\begin{aligned} (c_0 + c_1X + \cdots + c_iX^i + \cdots + (-1)^n X^n)I_n &= \cdots \\ \cdots &= (A - XI_n)(B_0 + \cdots + X^iB_i + \cdots + X^{n-1}B_{n-1}). \end{aligned}$$

Ce qui donne en développant le terme de droite :

$$\begin{aligned} (c_0 + c_1X + \cdots + c_iX^i + \cdots + (-1)^n X^n)I_n &= \cdots \\ \cdots &= AB_0 + (AB_1 - B_0)X + \cdots + (AB_i - B_{i-1})X^i + \cdots + (-B_{n-1})X^n. \end{aligned}$$

Les coefficients de X^i sont dans chacun des deux termes des matrices à coefficients dans k , on peut donc les identifier. On obtient alors les relations suivantes :

$$AB_0 = c_0I_n, \quad \dots, \quad AB_i - B_{i-1} = c_iI_n, \quad \dots, \quad -B_{n-1} = (-1)^n I_n.$$

Substituons à B_0 sa valeur donnée par la seconde équation, soit $AB_1 - c_1I_n$, dans la première. On obtient $A(AB_1 - c_1I_n) - c_0I_n = 0$, soit $A^2B_1 - c_1A - c_0I_n = 0$. On substitue alors dans cette dernière équation la valeur de B_1 donnée par $AB_2 - B_1 = c_2I_n$, et ainsi de suite. On obtient au i -ième pas l'équation suivante :

$$A^iB_{i-1} - c_{i-1}A^{i-1} - c_{i-2}A^{i-2} - \cdots - c_0I_n = 0,$$

et au dernier pas :

$$(-1)^{n+1}A^n - c_{n-1}A^{n-1} - \cdots - c_{i-1}A^{i-1} - \cdots - c_0I_n = 0.$$

Ceci démontre que le polynôme caractéristique s'annule quand on l'évalue sur la matrice de f dans une base quelconque, ou ce qui est équivalent sur f . ■

Soit E un espace vectoriel de dimension n sur un corps k algébriquement clos, et soit f une application linéaire. Soit $P = (-1)^n \prod_i (X - \lambda_i)^{\alpha_i}$ le polynôme caractéristique de f , les λ_i étant deux à deux distincts. Le théorème des noyaux implique que E est somme directe des sous-espaces $E_i = \text{Ker}(f - \lambda_i \text{Id})^{\alpha_i}$. Le polynôme caractéristique de la restriction de f à E_i est au signe près $(X - \lambda_i)^{\alpha_i}$.

Définition 2

Le sous-espace E_i est par définition *le sous-espace caractéristique* de f associé à la valeur propre λ_i .

Matrices compagnons et invariants de similitude

Étant donné un polynôme de degré n à coefficients dans un corps k il existe toujours une application linéaire d'un espace vectoriel de dimension n dans lui-même qui admet ce polynôme comme polynôme caractéristique.

Proposition 2. Soit $P = a_0 + \dots + a_i X^i + \dots + X^n$ un polynôme à coefficients dans k . La matrice (n, n) suivante admet P pour polynôme caractéristique :

$$\begin{pmatrix} 0 & \dots & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & -a_{n-2} \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Une matrice de ce type est appelée la matrice compagnon associée au polynôme P .

Soit $f \in \mathcal{L}_k(E)$. Supposons que E est de dimension finie sur k $A - XI_n$. Il est possible de montrer que l'on peut trouver une décomposition en somme directe de E , $E \cong \sum_{i=1, \dots, h} E_i$ telle que :

- Chaque E_i est stable par f , i.e. $f(E_i) \subset E_i$,
- on peut trouver une base de E_i dans laquelle la restriction de f à E_i a pour matrice une matrice compagnon de polynôme caractéristique P_i ,
- le polynôme P_i divise P_{i+1} si $i < h$.

On a :

Théorème 2. La décomposition précédente a les propriétés suivantes :

- La suite des polynômes P_i est uniquement déterminée par f , autrement dit, elle est la même pour toute décomposition satisfaisant aux propriétés ci-dessus,
- le polynôme P_h est égal au polynôme minimal de f ,
- le polynôme caractéristique de f est égal à $\prod_{1 \leq i \leq h} P_i$,
- si deux applications linéaires f et g sont conjuguées sous l'action du groupe linéaire, et si on note P_i et Q_i les polynômes qui leur sont associés on a $P_i = Q_i$ pour tout i ,
- inversement soient deux applications linéaires f et g , et soient P_i et Q_i les polynômes associés. Si on a $P_i = Q_i$ pour tout i les applications linéaires f et g sont conjuguées.

Nous n'allons pas démontrer ce théorème. La démonstration en sera faite en exercice. Notons seulement que les points deux et trois de l'énoncé sont faciles. Le troisième point résulte d'un calcul explicite, le second de la condition de divisibilité portant sur les polynômes. La cinquième condition est elle aussi facilement démontrée.

Les parties difficiles sont en premier lieu l'existence affirmée avant le théorème, puis l'unicité de la suite de polynômes. Tout ceci sera proposée en exercice.

Les polynômes définis plus haut sont appelés les invariants de similitude de l'application linéaire.

Notons enfin pour conclure l'analogue pour les matrices.

Deux matrices (n, n) A et B sont semblables s'il existe une matrice inversible (n, n) , soit P , telle que $A = P^{-1}BP$. Identifiant les matrices à des applications linéaires de k^n dans lui-même on peut associer à chacune une suite de polynômes avec les conditions indiquées plus haut.

Alors le théorème précédent affirme que les matrices sont semblables si et seulement si les deux suites de polynômes coïncident. Ceci permet donc de déterminer les classes de conjugaison du groupe linéaire.

3. Triangularisation des matrices et réduction de Jordan

On suppose dans toute cette section que l'on travaille sur un corps algébriquement clos. On pourra se restreindre au corps $\mathbb{C}$.

Théorème 1. *Soit E un espace vectoriel de dimension finie sur un corps algébriquement clos. Soit f une application linéaire de E dans lui-même, alors il existe une base de E dans laquelle la matrice de f est triangulaire supérieure.*

Démonstration. On raisonne par récurrence sur la dimension de E . Soit n la dimension de E . On admet que l'on a démontré le théorème pour toutes les applications linéaires sur des espaces vectoriels de dimension $n - 1$.

Le corps de base étant algébriquement clos l'application linéaire f a au moins une valeur propre λ . Soit $v \in E$ un vecteur propre non nul associé à λ et soit F le sous-espace vectoriel de dimension 1 de E engendré par v . L'espace vectoriel quotient E/F est de dimension $n - 1$. L'application f envoie par hypothèse le sous-espace F dans lui-même. Elle induit donc une application linéaire $\bar{f} : E/F \rightarrow E/F$ telle que $\pi \circ f = \bar{f} \circ \pi$, où π est la projection canonique de E sur E/F .

Par hypothèse de récurrence on peut trouver une base $(e'_2, \dots, e'_n)$ de E/F dans laquelle la matrice $\bar{f}$ est triangulaire supérieure, notons la B . Choisissons des

éléments $e_2, \dots, e_n \in E$ tels que $\pi(e_2) = e'_2, \dots, \pi(e_n) = e'_n$. Montrons que le système $(v, e_2, \dots, e_n)$ est une base de E . Comme il est constitué de n vecteurs il suffit de montrer que c'est un système libre. Supposons que l'on ait une relation :

$$\alpha v + \sum_{i=2, \dots, n} \alpha_i e_i = 0,$$

en appliquant π on obtient :

$$\sum_{i=2, \dots, n} \alpha_i e'_i = 0,$$

donc tous les coefficients α_i sont nuls. Il reste donc $\alpha v = 0$, mais comme v est non nul, α est nul.

La matrice de f dans cette base est alors de la forme souhaitée, à savoir :

$$\begin{pmatrix} \lambda & \alpha_2 & \dots & \alpha_n \\ 0 & (B) & & \end{pmatrix}.$$

En effet, on a d'un côté $f(v) = \lambda v$. Par ailleurs pour $2 \leq i \leq n$ on a $f(e_i) = \sum_{j \leq i} b_{j,i} e_j$ modulo un élément de F , soit $f(e_i) = \sum_{j \leq i} b_{j,i} e_j + \alpha_i v$. ■

De même que l'on a étudié la diagonalisation de deux applications linéaires (ou d'une famille d'applications linéaires) qui commutent on peut étudier la triangularisation de deux applications linéaires qui commutent :

Proposition 1. Soit E un espace vectoriel de dimension finie sur un corps algébriquement clos. Soient $f_i, i \in I$, une famille d'applications linéaires commutant deux à deux. Alors on peut trouver une base de E dans laquelle chacune de ces applications linéaires a une matrice triangulaire supérieure.

Démonstration. Elle est proposée en exercice. ■

Réduction de Jordan

La réduction de Jordan est une forme particulière de triangularisation qui est très utile dans certains cas, en particulier dans des études théoriques. Nous allons nous placer uniquement sur le corps $\mathbb{C}$ pour faire cette étude, et procéder en deux étapes.

Dans la première étape, on supposera que le polynôme minimal (ou le polynôme caractéristique, ce qui est équivalent) a une seule racine.

Supposons donc que le polynôme minimal de f soit de la forme $(\lambda - X)^k$. Pour tout entier ℓ compris entre 1 et k définissons un sous-espace vectoriel E_ℓ de E par la formule suivante :

$$E_\ell = \{v \in E \mid (\lambda \text{Id} - f)^\ell(v) = 0\}.$$

Le sous-espace E_1 est l'ensemble des vecteurs propres et par hypothèse on a $E_k = E$. On a la suite d'inclusion :

$$E_1 \subset \cdots \subset E_\ell \subset \cdots \subset E_k = E.$$

Notons g pour l'application linéaire $f - \lambda \text{Id}$. Par définition $g(E_i)$ est contenu dans E_{i-1} .

Choisissons un supplémentaire F_k de E_{k-1} dans $E_k = E$. Notons que par définition g^{k-1} est injective sur F_k .

Puis, choisissons un supplémentaire F_{k-1} de la somme $E_{k-2} + g(F_k)$ dans E_{k-1} . En fait, il est facile de voir que la somme précédente est directe. Nous donnerons un énoncé plus général dans le prochain lemme.

Par récurrence supposons avoir défini des sous-espaces vectoriels

$$F_k \subset E_k, \dots, F_{i+1} \subset E_{i+1}.$$

Définissons alors F_i comme un supplémentaire de

$$\sum_{i+1 \leq \ell \leq k} g^{\ell-i}(F_\ell) + E_{i-1}$$

dans E_i . Cette construction est justifiée car le sous-espace

$$\sum_{i+1 \leq \ell \leq k} g^{\ell-i}(F_\ell)$$

est contenu dans E_i . En effet par construction pour tout ℓ tel que $i+1 \leq \ell \leq k$ le sous-espace $g^{\ell-i}(F_\ell)$ est contenu dans E_i .

Par définition g^{i-1} est injective sur F_i , et en fait sur $F_i + \sum_{i+1 \leq \ell \leq k} g^{\ell-i}(F_\ell)$.

Lemme 1. *L'espace vectoriel E est isomorphe à la somme directe :*

$$\bigoplus_{1 \leq \ell \leq k} \left(\bigoplus_{0 \leq t \leq \ell-1} g^t(F_\ell) \right).$$

Démonstration. La démonstration consiste essentiellement à vérifier que les sous-espaces qui apparaissent dans la formule sont en somme directe.

Pour montrer que la somme est directe il nous faut établir que si on a une relation de la forme :

$$\sum_{1 \leq \ell \leq k} \sum_{0 \leq t \leq \ell-1} v_{\ell,t} = 0,$$

avec $v_{\ell,t} \in g^t(F_\ell)$, tous les éléments $v_{\ell,t}$ sont nuls.

Pour montrer que E est égal à cette somme directe, il faut aussi établir que tout élément de E s'écrit sous la forme $\sum_{1 \leq \ell \leq k} \sum_{0 \leq t \leq \ell-1} v_{\ell,t}$ avec $v_{\ell,t} \in g^t(F_\ell)$.

Considérons d'abord la première question.

Supposons donc que $k > 1$ (le cas $k = 1$ est clair) et appliquons g^{k-1} à la relation. Il nous reste $g^{k-1}(v_{k,0}) = 0$, tous les autres termes de la somme étant dans le noyau de g^{k-1} . Or g^{k-1} est injectif sur F_k comme on l'a dit plus haut. En effet F_k a par construction une intersection réduite à $\{0\}$ avec E_{k-1} , donc $v_{k,0} = 0$.

Appliquons maintenant g^{k-2} on obtient l'équation : $g^{k-2}(v_{k,1}) + g^{k-2}(v_{k-1,0}) = 0$. En effet tous les éléments $v_{i,j}$ de la somme pour $i \leq k-2$ sont annulés par g^{k-2} et on vient de montrer que $v_{k,0}$ est nul. Ainsi qu'on l'a dit plus haut l'application g^{k-2} est injective sur $F_{k-1} \oplus g(F_k)$, en effet cette somme directe (qui l'est par construction) a par définition une intersection avec E_{k-2} réduite à $\{0\}$ car $g(F_k) \cap E_{k-2} = \{0\}$ puis que g^{k-1} est injectif sur $g(F_k)$. Il reste donc $v_{k,1} + v_{k-1,0} = 0$.

Comme les sous-espaces F_{k-1} et $g(F_k)$ sont par construction en somme directe ceci implique que $v_{k,1} = v_{k-1,0} = 0$.

Passons au cas général, soit ℓ un entier fixé entre 0 et $k-1$. Supposons avoir montré que : $v_{h,i} = 0$ pour tous h tels que $h-i > \ell$, que les sous-espaces $g^i(F_h)$, pour $h-i > \ell$ sont en somme directe, enfin que l'application g^ℓ est injective sur

$$\bigoplus_{0 \leq t \leq k-\ell-1} g^{k-\ell-t-1}(F_{k-t}),$$

L' image du sous-espace précédent par g est aussi une somme directe :

$$\bigoplus_{0 \leq t \leq k-\ell-1} g^{k-\ell-t}(F_{k-t}),$$

qui est par construction d'intersection réduite à $\{0\}$ avec $E_{\ell-1}$ et donc par définition de F_ℓ on récupère la somme directe suivante

$$\bigoplus_{0 \leq t \leq k-\ell} g^{k-\ell-t}(F_{k-t}),$$

sur laquelle par construction $g^{\ell-1}$ est injective.

Appliquons $g^{\ell-1}$ à la relation considérée plus haut. Étant donné que les vecteurs $v_{h,i}$ sont nuls pour h tel que $h-i > \ell$ et que $g^{\ell-1}$ s'annule sur les sous-espaces

$g^i(F_h)$ pour $\ell > h - i$, on en déduit la relation :

$$g^{\ell-1}(v_{\ell,0} + \dots + v_{k,k-\ell}) = 0.$$

Mais comme $g^{\ell-1}$ est injective sur $\bigoplus_{0 \leq t \leq \ell} g^{\ell-t}(F_{k-t})$ il en résulte que tous les éléments $v_{k-\ell+t,t}$ sont nuls. Le résultat suit. ■

Construisons alors une base de E comme suit. Pour chaque F_i on choisit une base que l'on notera $(e_{i,1}, \dots, e_{i,d_i})$, où d_i est la dimension de F_i . Le résultat précédent montre que :

Proposition 2. *Le système de vecteurs*

$$\left\{ \begin{array}{l} (e_{1,1}, \dots, e_{1,d_1}), \\ \dots \\ (e_{i,1}, \dots, g^h(e_{i,1}), \dots, g^{i-1}(e_{i,1})), \dots, (e_{i,d_i}, \dots, g^h(e_{i,d_i}), \dots, g^{i-1}(e_{i,d_i})), \\ \dots \\ (e_{k,1}, \dots, g^h(e_{k,1}), \dots, g^{k-1}(e_{k,1})), \dots, (e_{k,d_k}, \dots, g^h(e_{k,d_k}), \dots, g^{k-1}(e_{k,d_k})) \end{array} \right\}$$

est une base de l'espace vectoriel E .

C'est par définition une base de Jordan pour f . C'est-à-dire une base dans laquelle la matrice f est triangulaire supérieure et ne comporte, éventuellement, comme termes non nuls au dessus de la diagonale principale que des 1 sur la diagonale qui lui est immédiatement supérieure.

L'énoncé est précisé par la démonstration ci-dessous.

Démonstration. Soit i tel que $1 \leq i \leq k$ et que $1 \leq d_i$. Soit un entier ℓ tel que $1 \leq \ell \leq d_i$, soit alors $E_{i,\ell}$, le sous-espace engendré par les vecteurs $(e_{i,\ell}), \dots, g^{i-1}(e_{i,\ell})$. On vient de voir que l'espace E est somme directe des sous-espaces $E_{i,\ell}$. Ces sous-espaces sont stables par f . On a en effet

$$f(g^h(e_{i,\ell})) = (-g + \lambda \text{Id})g^h(e_{i,\ell}) = -g^{h+1}(e_{i,\ell}) + \lambda g^h(e_{i,\ell}).$$

Étant entendu que $g^h(e_{i,\ell})$ est nul si $h \geq i$.

En fait le système $(e_{i,\ell}), \dots, g^{i-1}(e_{i,\ell})$ constitue même une base de $E_{i,\ell}$, et dans cette base la matrice de la restriction de f a la forme :

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$$

La matrice de f est donc en fin de compte une matrice diagonale par blocs :

$$\begin{pmatrix} J_{1,d_1} & 0 & \dots & \dots & 0 \\ 0 & J_{1,d_2} & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & J_{k,d_k} \end{pmatrix},$$

où $J_{i,\ell}$ est la matrice de la restriction de f au sous-espace vectoriel $E_{i,\ell}$.

Il nous reste à traiter le cas général de la jordanisation. Soit donc f une application linéaire d'un espace vectoriel complexe E de dimension finie dans lui-même. Puisque le corps sur lequel on travaille est $\mathbb{C}$ le polynôme minimal de f a toutes ses racines dans le corps et se scinde donc en produit de polynômes de la forme $(X - \lambda_i)^{\alpha_i}$.

Appliquons le théorème des noyaux au polynôme caractéristique, il implique que l'on peut écrire E comme somme directe de sous-espaces caractéristiques E_i stables par f . Sur ces sous-espaces, la restriction de f a pour polynôme minimal un diviseur de $(X - \lambda_i)^{\alpha_i}$. En fait ce dernier polynôme est le polynôme minimal d'après le théorème des noyaux.

On est donc ramené sur chaque E_i au cas précédent. Une base de jordanisation est alors par définition une base obtenue par réunion de bases de jordanisation sur chacun de ces sous-espaces. ■

Dans la fin de cette section nous allons donner des conséquences. Nous travaillerons uniquement dans un cadre matriciel.

Définition 1

Une matrice $N \in M_n(k)$ est dite nilpotente s'il existe un entier ℓ tel que $N^\ell = 0$.

Proposition 3. Une matrice triangulaire supérieure dont tous les éléments diagonaux sont nuls est nilpotente.

Démonstration. En fait si $T \in M_n(k)$ désigne la matrice en question on a $T^n = 0$. ■

Définition 2

Une matrice $U \in M_n(k)$ est dite unipotente si la matrice $U - I_n$ est nilpotente.

En particulier si la matrice N est nilpotente la matrice $I_n + N$ est unipotente.

Théorème 2. Soit A une matrice (n, n) à coefficients dans un corps k algébriquement clos. Alors il existe une matrice diagonalisable D et une matrice nilpotente N telles que :

- $A = D + N$,
- $DN = ND$.

Si de plus la matrice A est inversible alors les matrices D et N sont uniquement déterminées par ces conditions.

Démonstration. On commence par trouver une base de jordanisation pour A . Il existe une base B' et donc une matrice de changement de bases P telle que la matrice $P^{-1}AP = (\pi_{i,j})$ soit une réduite de Jordan. La matrice $P^{-1}AP$ peut donc s'écrire comme somme de la matrice $\Delta = (d_{i,j})$ définie par :

$$\begin{cases} d_{i,i} = \pi_{i,i}, \\ d_{i,j} = 0, & \text{si } i \neq j \end{cases}$$

et de la matrice $N = (n_{i,j})$ définie par :

$$\begin{cases} n_{i,i+1} = \pi_{i,i+1}, \\ \delta_{i,j} = 0, & \text{si } j \neq i+1. \end{cases}$$

La matrice N n'a de termes non nuls que sur la diagonale juste au dessus de la diagonale principale. Ces termes ne pouvant prendre que la valeur 1. Les matrices Δ et N commutent, donc les matrices $P\Delta P^{-1}$ et PNP^{-1} répondent à la question posée. Il reste à démontrer l'unicité de la décomposition.

Considérons une décomposition $A = D' + N'$ avec les propriétés requises. Les matrices A , D et N commutent ceci entraîne en particulier que les sous-espaces propres de D sont stables par A et N . Ces sous-espaces constituent une décomposition en somme directe de k^n . Ils coïncident avec les sous-espaces caractéristiques de A , ils sont donc bien déterminés et on peut se restreindre au cas où D' est une matrice diagonale. Ceci implique le résultat. ■

4. Applications topologiques de la diagonalisation et de la triangularisation

Dans cette section, le corps de base sera $\mathbb{C}$, sauf brièvement, à la fin, où on se placera sur $\mathbb{R}$. Rappelons que sur l'espace vectoriel complexe $M_n(\mathbb{C})$ de dimension finie, toutes les normes sont équivalentes. Dans la suite on choisira par exemple la norme donnée par $\|A\| = \sum_{i,j} |a_{i,j}|$.

Proposition 1. L'ensemble des matrices inversibles est ouvert et partout dense dans l'ensemble des matrices.

Démonstration. En effet une matrice est inversible si et seulement si son déterminant est non nul. Le déterminant est une application :

$$M_n(\mathbb{C}) \longrightarrow \mathbb{C}.$$

Cette application est donnée par une fonction polynomiale, elle est donc continue. L'ensemble des matrices inversibles $GL_n(\mathbb{C})$, qui est l'image inverse de l'ouvert $\mathbb{C} - \{0\}$, est ouvert.

Pour ce qui est de la densité, on considère une matrice A quelconque. Cette matrice peut être triangularisée. Il existe une matrice P inversible telle que $P^{-1}AP = T$ est triangulaire. La matrice T est inversible si et seulement si ses éléments diagonaux, qui sont les valeurs propres de A , sont tous non nuls.

Modifions T en la remplaçant par $T + \varepsilon I_n$. On peut trouver des $\varepsilon \in \mathbb{C}$ de module arbitrairement petit tels que $T + \varepsilon I_n$ soit inversible. En fait dès que le module de ε est strictement inférieur à celui des éléments diagonaux de T et non nul la matrice $T + \varepsilon I_n$ est inversible.

La matrice $P(T + \varepsilon I_n)P^{-1}$ est donc inversible dès que le module de ε est assez petit et non nul. Cette matrice est égale à $A + \varepsilon I_n$, quand $|\varepsilon|$ tend vers zéro cette matrice admet pour limite la matrice A . Le résultat suit. ■

Proposition 2. *L'ensemble des matrices diagonalisables est partout dense dans l'ensemble des matrices.*

Démonstration. Comme pour le résultat précédent on commence par triangulariser la matrice A : on a $T = P^{-1}AP$ pour une certaine matrice P . Mais au lieu d'ajouter à T un multiple de I_n on ajoute une matrice diagonale :

$$E = \begin{pmatrix} \varepsilon_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \varepsilon_n \end{pmatrix}.$$

Considérons la matrice $T + E$, on peut trouver des valeurs de $\varepsilon_1, \dots, \varepsilon_n$, de module aussi petit que l'on veut, telles que les éléments diagonaux de $T + E$ soient deux à deux distincts.

Soit en effet $x = (x_1, \dots, x_n)$ un vecteur dont les composantes sont deux à deux distinctes. Soit $v = (v_1, \dots, v_n)$ un vecteur quelconque.

Montrons que le vecteur $v + tx = (v_1, \dots, v_n) + t(x_1, \dots, x_n)$ a ses coordonnées deux à deux distinctes dès que le module de t est assez petit et non nul. Sinon on aurait une équation de la forme $v_i + tx_i = v_j + tx_j$. Mais comme $x_i \neq x_j$, si $i \neq j$, cette équation a une solution et une seule éventuellement nulle. Il faut qu'un nombre fini d'équations de ce type n'ait pas lieu. Il suffit donc de prendre t non nul et de

module inférieur à celui des solutions de ces équations pour garantir que $v + tx$ a ses coordonnées deux à deux distinctes.

Pour ces valeurs de t la matrice $T + tE$ est triangulaire avec des éléments diagonaux deux à deux distincts la matrice est diagonalisable. Il en est de même de $P(T + tE)P^{-1} = A + tE$. Quand t tend vers zéro cette matrice tend vers A . Le résultat suit. ■

Proposition 3. *L'ensemble des matrices (n, n) ayant des valeurs propres deux à deux distinctes est un ouvert partout dense dans l'ensemble des matrices.*

Démonstration. Une matrice A a ses valeurs propres deux à deux distinctes si et seulement si son polynôme caractéristique P n'a pas de racines communes avec son polynôme dérivé P' . Cette condition s'exprime par la non nullité du discriminant de P (chapitre 3, section 8). Mais le discriminant de P est un polynôme en les coefficients de P , c'est une application

$$\mathbf{M}_n(\mathbb{C}) \longrightarrow \mathbb{C}.$$

C'est une fonction polynomiale en les coefficients de la matrice A , elle est donc continue. L'image inverse de l'ouvert $\mathbb{C} - \{0\}$ qui est l'ensemble des matrices dont les valeurs propres sont deux à deux distinctes est donc un ouvert de l'ensemble des matrices.

Pour ce qui est de la densité l'argument donné pour la proposition 2 fonctionne à l'identique. ■

Proposition 4. *Le groupe des matrices (n, n) à coefficients complexes inversibles est connexe par arcs et donc connexe.*

Démonstration. On va en donner deux démonstrations.

La première utilise le lemme suivant :

Lemme 1. *Le plan complexe privé d'un nombre fini de points est connexe par arcs.*

La démonstration est laissée au lecteur.

Soient $A, B \in \mathbf{GL}_n(\mathbb{C})$, considérons le polynôme en la variable complexe z $\det(zA + (1 - z)B)$. Ce polynôme a un nombre fini de racines. On peut donc trouver un chemin $t \mapsto \gamma(t)$ de 0 à 1 dans $\mathbb{C}$ tel que pour toute valeur t comprise entre 0 et 1 le déterminant $\det \gamma(t)A + (1 - \gamma(t))B$ soit non nul donc que la matrice $\gamma(t)A + (1 - \gamma(t))B$ soit inversible. Ce qui démontre l'assertion.

La deuxième démonstration utilise la triangularisation. On commence par observer qu'il suffit pour démontrer la connexité par arcs de construire un chemin d'une matrice quelconque à la matrice identité. En effet pour construire un chemin

d'une matrice A à une matrice B il suffira alors de mettre bout à bout un chemin joignant A à la matrice identité, puis un chemin joignant la matrice identité à B .

Précisons, soit un chemin, c'est-à-dire un arc continu $\alpha : [0, 1] \rightarrow \text{GL}_n(\mathbb{C})$, de A à A' ($\alpha(0) = A$, $\alpha(1) = A'$) et un arc continu $\alpha' : [0, 1] \rightarrow \text{GL}_n(\mathbb{C})$, de A' à A'' ($\alpha'(0) = A'$, $\alpha'(1) = A''$).

On définit un arc continu γ de A à A'' par $\gamma(t) = \alpha(2t)$ si $0 \leq t \leq \frac{1}{2}$, et $\gamma(t) = \alpha'(2t - 1)$ si $\frac{1}{2} \leq t \leq 1$.

Soit A une matrice quelconque. On peut trouver une matrice P telle que $P^{-1}AP$ soit une matrice triangulaire supérieure T . Commençons donc par construire un chemin de T à la matrice identité. Le chemin est donné comme suit. On commence par relier la matrice triangulaire à une matrice diagonale par :

$$t \mapsto \begin{pmatrix} a_{1,1} & (1-t)a_{1,2} & \dots & (1-t)a_{1,n} \\ 0 & a_{2,2} & (1-t)a_{2,3} & (1-t)a_{2,n} \\ 0 & & \ddots & \vdots \\ 0 & & 0 & a_{n-1,n-1} & (1-t)a_{n-1,n} \\ 0 & & \dots & \dots & a_{n,n} \end{pmatrix},$$

t variant de 0 à 1.

Posons $a_{\ell,\ell} = \varrho_\ell e^{(i a_\ell)}$ et relierons la matrice diagonale à la matrice identité par :

$$t \mapsto \begin{pmatrix} \varrho_1^{(1-t)} e^{(1-t)ia_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \varrho_n^{(1-t)} e^{(1-t)ia_n} \end{pmatrix}.$$

En mettant ces chemins bout à bout on obtient un chemin γ de la matrice triangulaire T à la matrice identité. Soit alors γ un chemin de T à la matrice identité. Le chemin

$$t \mapsto P\gamma(t)P^{-1}$$

est un chemin de la matrice A à la matrice identité. ■

Proposition 5. *La classe de conjugaison d'une matrice est fermée si et seulement si elle est diagonalisable.*

Démonstration. Supposons la matrice A diagonalisable. Soit P le polynôme minimal de A . Il n'a que des racines simples et par définition il annule A . Supposons donné une suite de matrices A_n dans la classe de conjugaison de A convergeant dans $M_n(\mathbb{C})$ vers une matrice B . L'application qui à une matrice M associe la matrice $P(M)$ est une application continue de l'espace $M_n(\mathbb{C})$ dans lui-même. Elle

est donc nulle sur la matrice B , le polynôme minimal de B est donc un diviseur de P . Il a donc toutes ses racines simples et B est donc diagonalisable. En fait, toute valeur propre λ de A est valeur propre de B et le polynôme minimal de B est égal à P . En effet soit $v \in \mathbb{C}^n$ un vecteur non nul tel que $Av = \lambda v$. Supposons que $A_n = P_n^{-1}AP_n$; alors le vecteur propre $P_n^{-1}v$ est vecteur propre de A_n associé à la valeur propre λ . De la suite de vecteurs

$$\frac{P_n^{-1}v}{\|P_n^{-1}v\|}$$

on peut extraire une sous-suite convergente. En effet ces vecteurs sont dans la sphère de rayon 1 qui est compacte. Par construction cette suite converge vers un vecteur propre de B associé à la valeur propre λ . Il faut montrer encore que la dimension des sous-espaces propres est la même. Puisque la matrice est diagonalisable la dimension du sous-espace propre associé à une valeur propre λ est la multiplicité de λ comme racine du polynôme caractéristique. Or les coefficients du polynôme caractéristique dépendent continûment des coefficients de la matrice et sont constants sur une classe de conjugaison. La matrice B a donc même polynôme caractéristique que A , et lui est donc conjuguée.

Inversement soit A une matrice non-diagonalisable. Choisissons en une réduction de Jordan. Il y a donc au moins un terme non nul égal à 1 immédiatement au dessus de la diagonale principale. Les matrices A_ε obtenues en remplaçant les termes égaux à 1 immédiatement au dessus de la diagonale principale par ε quelconque sont toutes conjuguées. Admettons ce point, alors si on fait tendre ε vers zéro la matrice A_ε tend vers une matrice diagonale. La classe de conjugaison de A étant constituée uniquement de matrices non-diagonalisables n'est pas fermée.

Le point laissé en suspend sera démontré en exercice par le lecteur en faisant une homothétie sur les vecteurs de base, le rapport dépendant du vecteur. $\square$

Nous noterons pour conclure que :

Proposition 6. *Le groupe $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe.*

Démonstration. En effet on commence par considérer l'application déterminant :

$$\mathrm{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}.$$

Comme elle est donnée par une fonction polynomiale des coefficients elle est continue. Le groupe $\mathrm{GL}_n(\mathbb{R})$ qui est l'image inverse de l'ouvert $\mathbb{R} - \{0\}$ est ouvert. mais son image par le déterminant n'est pas connexe, donc il ne peut être connexe.

Ce résultat sera complété en exercice. $\square$

5. Générateurs du groupe linéaire

Dans cette section nous allons décrire un système de générateurs du *groupe linéaire*, puis en déduire des conséquences. On étudiera aussi le groupe dérivé du groupe linéaire.

Soient donc k un corps, un espace vectoriel E sur k , de dimension finie, et le groupe linéaire $\text{GL}(E)$. On se placera aussi dans un contexte matriciel. On considérera donc $E = k^n$ et le groupe $\text{GL}_n(k)$ dont on sait qu'il est isomorphe à $\text{GL}(E)$ si E est de dimension n .

Introduisons des classes particulières d'applications linéaires :

Définition 1

On appelle *dilation* de l'espace vectoriel E une application qui laisse fixe points par points un hyperplan H et qui laisse fixe globalement une droite vectorielle D supplémentaire, cette droite est la direction de la dilatation.

On appelle *transvection d'hyperplan* H , d'équation $\alpha \in E^*$, et de direction $v \in H$, $v \neq 0$, l'application linéaire f de E dans lui-même telle que :

$$f(x) = x + \alpha(x)v$$

pour tout $x \in E$.

Notons que α est donné dans la définition de la transvection. Mais il est clair que l'on peut remplacer v par av , et α par $a^{-1}\alpha$ avec $a \in k^*$.

On va démontrer :

Théorème 1. *Les transvections et les dilatations constituent un système générateur de $\text{GL}(E)$.*

On va donner une démonstration basée sur le calcul matriciel et on va donc considérer un énoncé analogue pour le groupe $\text{GL}_n(k)$. En fait on montrera que les matrices correspondant à certaines transvections et à certaines dilatations particulières engendrent $\text{GL}_n(k)$.

Commençons par décrire les matrices des dilatations et des transvections dans des bases privilégiées.

Pour ce qui est des dilatations, on choisit une base constituée de la réunion d'un vecteur v non nul de D et d'une base de H . On ne précise pas dans quel ordre sont pris ces vecteurs. La matrice de la dilatation est alors diagonale avec des 1 sur toutes les colonnes sauf une. Sur cette dernière colonne correspondant au vecteur porté par D on a le coefficient $a \in k^*$ de la dilatation, quelconque. Si v est

le premier vecteur de base on obtient la matrice suivante :

$$\begin{pmatrix} a & 0 & \dots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & & 1 & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Pour ce qui est des transvections, on choisit une base constituée, par une base de H à laquelle appartient le vecteur v , et d'un vecteur e quelconque n'appartenant pas à H . On ne précise pas dans quel ordre sont pris ces vecteurs. La matrice de la transvection a alors toutes ses colonnes, sauf une, nulles en dehors du terme diagonal égal à 1. Les colonnes nulles en dehors de la diagonale correspondent aux vecteurs de base qui appartiennent à l'hyperplan. La colonne à part correspond au vecteur e qui est égale à 1 sur la diagonale, égale à $\alpha(e) \in k^*$ sur la ligne correspondant à v et nulle partout ailleurs. Ici α est l'équation de H .

Si le vecteur v est le premier vecteur de base et si e est le second on obtient comme matrice :

$$\begin{pmatrix} 1 & \alpha(e) & \dots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Considérons des matrices particulières du type de celles que l'on vient de décrire.

Soit $1 \leq i, j \leq n$ avec $i \neq j$. Notons $E_{i,j}(\lambda)$, $\lambda \in k$, la matrice dont tous les termes diagonaux sont égaux à 1 et dont tous les autres termes sont nuls à l'exception du terme situé sur la i -ième ligne et la j -ième colonne qui est égal à λ . Une telle matrice est appelée une matrice élémentaire. Si $\lambda = 1$ elle sera notée simplement $E_{i,j}$. C'est une matrice de transvection.

Notons $D(\lambda)$, $\lambda \in k^*$ pour la matrice diagonale dont le premier terme (première ligne et première colonne) est égal à λ et dont tous les autres termes sont égaux à 1. C'est une matrice de dilatation.

Théorème 2. Les matrices $E_{i,j}$, pour $1 \leq i, j \leq n$ $i \neq j$, et $D(\lambda)$, $\lambda \in k$, constituent un système générateur de $\text{GL}_n(k)$.

Ceci implique le théorème 1 dans la mesure où on montre que $\text{GL}_n(k)$ est engendré par des matrices de dilatations et transvections particulières.

Démonstration. On va procéder en plusieurs étapes.

Démonstration dans le cas $n = 2$. Tout d'abord on commence par le cas $n = 2$. Notons H le sous-groupe engendré par les matrices $D(\lambda)$, $\lambda \in k$, et $E_{i,j}$, $i \neq j$.

Commençons par noter que :

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = D(-1)E_{1,2}^{-1}E_{2,1}E_{1,2}^{-1},$$

soit T cette matrice. Puis notons que :

$$TD(\lambda)T = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix},$$

notons cette matrice $D'(\lambda)$. Les matrices T et $D'(\lambda)$ sont obtenues par produit de matrices $D(\lambda)$ et $E_{i,j}$, elles sont donc dans le sous-groupe H de $\mathbf{GL}_2(k)$.

Enfin, si on calcule le produit $D'(\lambda)E_{2,1}D'(\lambda^{-1})$, avec $\lambda \neq 0$, on trouve $E_{2,1}(\lambda)$. Donc les matrices $E_{2,1}(\lambda)$ sont dans H . On montrerait de même que les matrices $E_{1,2}(\lambda)$ sont dans H .

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(k)$. On veut montrer qu'elle est dans H .

Supposons d'abord $a \neq 0$ et $b \neq 0$. Dans ce cas formons le produit : $D(a^{-1})AD'(b^{-1}a)$. Il est égal à :

$$B = \begin{pmatrix} 1 & 1 \\ c & adb^{-1} \end{pmatrix}.$$

Puis formons le produit $BE_{1,2}^{-1}$, il est égal à :

$$C = \begin{pmatrix} 1 & 0 \\ c & -c + adb^{-1} \end{pmatrix}.$$

On notera que $-c + adb^{-1}$ est non nul par hypothèse.

La matrice $CD'((-c + adb^{-1})^{-1})$ est égale à

$$E_{2,1}\left(\frac{c}{-c + adb^{-1}}\right)$$

et est donc dans le sous-groupe engendré par les matrices $D(\lambda)$ et $E_{i,j}$. Il en est donc de même des matrices C , B et A .

Si $b = 0$, on a d'emblée une matrice triangulaire inférieure. Les coefficients a et d sont nécessairement non nuls, car le déterminant est non nul. La matrice $D(a^{-1})AD'(d^{-1})$ est de la forme $E_{2,1}(\lambda)$ pour un certain λ et est donc dans H .

Si $a = 0$ on a nécessairement $b \neq 0$. Dans ce cas la matrice AT est triangulaire inférieure du type que l'on vient d'étudier est donc dans H . Il en est de même de A .

Ceci achève le cas $n = 2$.

□

Démonstration dans le cas général. On va faire une récurrence sur n et donc supposer le théorème démontré pour $n - 1$ et passer au cas n . Nous allons commencer par un lemme qui a son intérêt propre.

Définissons les matrices monomiales.

On dira qu'une matrice M est monomiale si et seulement si elle n'a qu'un seul terme non nul et égal à 1 dans chaque ligne et chaque colonne.

Autrement dit :

Lemme 1. *Soit M une matrice monomiale. Il existe une permutation $s \in \mathfrak{S}_n$ telle que les coefficients $m_{i,j}$ de la matrice M sont donnés par $m_{s(i),i} = 1$, $m_{i,j} = 0$ si $j \neq s(i)$. Le sous-ensemble constitué par ces matrices est un sous-groupe isomorphe à $\mathfrak{S}_n$.*

Démonstration. La démonstration de la première partie est laissée en exercice.

Notons $M(s)$ la matrice monomiale associée à la permutation s . C'est la matrice de l'application linéaire f_s de k^n dans lui-même donnée dans la base canonique par $f_s(e_i) = e_{s(i)}$. On a une application :

$$\mathfrak{S}_n \longrightarrow \mathrm{GL}_n(k), \quad s \longmapsto M(s).$$

C'est un homomorphisme de groupes. On le vérifie facilement au niveau des applications linéaires f_s . En effet le produit $f_s \circ f_{s'}$ est égal à $f_{ss'}$ car $f_{ss'}(e_i) = f_s(e_{s'(i)}) = e_{s(s'(i))}$. Il est clairement injectif et est donc un isomorphisme sur son image. On remarquera que le calcul précédent montre que les matrices monomiales sont bien inversibles. $\square$

Quand on multiplie à gauche une matrice A par la matrice monomiale $M(s)$ on obtient la matrice déduite de A par permutation des lignes selon s . C'est-à-dire que la i -ième ligne devient la $s(i)$ -ième. De façon analogue quand on multiplie à droite une matrice A par la matrice monomiale $M(s)$ on obtient la matrice déduite de A par permutation des colonnes selon s .

On va commencer par montrer que toutes les matrices monomiales se trouvent dans le sous-groupe engendré par matrices $D(\lambda)$, $\lambda \in k$, et $E_{i,j}$, $i \neq j$, notons encore H ce sous-groupe. Il suffit de montrer que toutes les matrices $M(\tau_{i,i+1})$, avec $1 \leq i \leq n - 1$ s'y trouvent. Ici $\tau_{i,i+1}$ désigne comme à l'ordinaire la transposition échangeant i et $i + 1$. En effet, on sait (théorème 2, section 5, chapitre 2) que ces transpositions engendrent le groupe symétrique.

Considérons le sous-ensemble de $\mathrm{GL}_n(k)$ constitué par les matrices dont la dernière ligne et la dernière colonne sont nulles, sauf en ce qui concerne le terme sur la n -ième et la n -ième colonne qui est égal à 1. Il s'agit des matrices telles que :

$$\begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}.$$

La matrice B est dans $\mathbf{GL}_{n-1}(k)$, le sous-ensemble en question est un sous-groupe isomorphe à $\mathbf{GL}_{n-1}(k)$. Notons comme à l'ordinaire $e_1, \dots, e_n$ la base canonique. Il s'agit du sous-groupe laissant fixe le vecteur e_n et laissant fixe globalement l'hyperplan engendré par $e_1, \dots, e_{n-1}$.

On peut appliquer l'hypothèse de récurrence à ce sous-groupe. Les matrices $D(\lambda)$ et $E_{i,j}$, $i \neq j$ et $0 \leq i, j \leq n-1$ engendrent $\mathbf{GL}_{n-1}(k)$ qui est donc contenu dans H . Les matrices $M(\tau_{i,i+1})$, $i \leq n-2$, sont dans $\mathbf{GL}_{n-1}(k)$, donc dans H .

Il en est de même des matrices de dilatations laissant fixes les vecteurs de base $e_1, \dots, e_{n-2}$ et e_n . C'est-à-dire des matrices de la forme :

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & \dots & a & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

L'ensemble des matrices inversibles A telle que $Ae_j = e_j$ si $j \leq n-1$ et telles que Ae_{n-1} et Ae_n soient combinaisons linéaires de e_{n-1} et e_n est isomorphe à $\mathbf{GL}_2(k)$. Autrement dit ce sont les matrices de la forme :

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & & 1 & 0 \\ 0 & \dots & & D \end{pmatrix},$$

où $D \in \mathbf{GL}_2(k)$.

On peut appliquer le cas $n=2$ à ce sous-groupe puisque que l'on sait par récurrence que H contient les dilatations de vecteur e_{n-1} . Ceci montre alors que $M(\tau_{n-1,n})$ est dans H . Donc toutes les matrices monomiales sont dans H .

Venons en au cas général. Soit alors A une matrice quelconque de $\mathbf{GL}_n(k)$. Comme tous les éléments de la première ligne ne sont pas nuls quitte à multiplier à gauche par une matrice monomiale on peut supposer que l'élément sur la première ligne et la première colonne est non nul. Quitte à multiplier par une matrice $D(\alpha)$ on peut supposer qu'il est égal à 1.

Soit $A = (a_{i,j})$ cette matrice. Multiplions la à gauche par les matrices $E_{i,1}(-a_{i,1})$ et à droite par les matrices $E_{1,j}(-a_{1,j})$, $i, j > 1$. On obtient une matrice dont tous les éléments sur la première ligne et tous les éléments sur la première colonne sont nuls, exception étant faite de celui sur la première ligne et la première colonne qui reste égal à 1. L'argument donné pour $n=2$ s'étend facilement pour montrer que les matrices $E_{i,1}(-a_{i,1})$ et $E_{1,j}(-a_{1,j})$, sont dans H . Il suffit donc de montrer que la dernière matrice est dans H .

Mais l'ensemble des matrices qui vérifient les conditions ci-dessus est isomorphe à $\mathbf{GL}_{n-1}(k)$. On est alors en mesure d'appliquer l'hypothèse de récurrence pour conclure. ■

Le cas du groupe spécial linéaire

Nous allons étudier le problème analogue pour le groupe spécial linéaire dont la définition suit :

Définition 2

Soit E un espace vectoriel de dimension finie sur un corps k . On appelle groupe spécial linéaire de E et on note $\mathrm{SL}(E)$ le sous-groupe de $\mathrm{GL}(E)$ constitué par les endomorphismes de déterminant 1.

On appelle groupe spécial linéaire et on note $\mathrm{SL}_n(k)$ le sous-groupe de $\mathrm{GL}_n(k)$ constitué par les matrices de déterminant 1.

Autrement c'est le noyau de l'homomorphisme :

$$\det : \mathrm{GL}_n(k) \longrightarrow k^*,$$

et de l'analogue dans le cas des endomorphismes.

Théorème 3. *Les matrices $E_{i,j}(\lambda)$ engendrent le groupe $\mathrm{SL}_n(k)$.*

On a en corollaire immédiat à cause des remarques faites plus haut :

Corollaire. *Les transvections engendrent le groupe spécial linéaire.*

On va maintenant démontrer le théorème 3.

Démonstration. Il s'agit d'adapter la démonstration du théorème 2. On raisonne donc par récurrence sur n et on suppose le théorème démontré pour $n - 1$.

Soit $A = (a_{i,j}) \in \mathrm{SL}_n(k)$ et soit H le sous-groupe engendré par les matrices $E_{i,j}(\lambda)$.

On va montrer qu'en multipliant A par des matrices de H on peut la remplacer par une matrice dont la première colonne et la première ligne sont nulles, exception faite du terme $a_{1,1}$ qui sera égal à 1. La sous-matrice $(n - 1, n - 1)$ qui apparaît alors est de déterminant 1 et donc dans $\mathrm{SL}_{n-1}(k)$, on peut donc lui appliquer l'hypothèse de récurrence et conclure comme plus haut.

Comme la première colonne est non nulle quitte à multiplier à gauche par une matrice $E_{i,1}(\lambda)$, pour i bien choisi, on peut supposer $a_{1,1} \neq 0$.

Soit maintenant $a \neq 0$, on a :

Lemme 2. *La matrice*

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

est dans le sous-groupe L de $\mathrm{GL}_2(k)$ engendré par les matrices $E_{1,2}(\lambda)$ et $E_{2,1}(\mu)$.

Démonstration. Cela résulte des deux formules suivantes :

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & a^{-1} \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix} \times \begin{pmatrix} 1 & a^{-1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & a^{-1} \\ -a & 0 \end{pmatrix}.$$

Les deux matrices exprimées comme produit sont dans L . Leur produit

$$\begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix}$$

l'est aussi. □

Pour achever la démonstration du théorème il suffit alors de multiplier la matrice à laquelle on s'est ramenée plus haut par :

$$\begin{pmatrix} a_{1,1}^{-1} & 0 & \dots & \dots & 0 \\ 0 & a_{1,1} & 0 & & 0 \\ & 0 & & 1 & 0 & \vdots \\ 0 & & & & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Le résultat suit ainsi qu'il a été indiqué. ■

Le centre et le groupe dérivé

Nous allons maintenant analyser le centre et le groupe dérivé (voir les exercices du chapitre 1) du groupe linéaire.

Proposition 1. *Le centre du groupe $\mathrm{SL}_n(k)$ est égal à l'ensemble des homothéties de déterminant 1.*

Démonstration. Notons A un élément du centre de $\mathrm{SL}_n(k)$. Un tel élément doit commuter à l'ensemble des transvections. Supposons que pour un certain vecteur $v \in k^n$ les vecteurs v et Av ne soient pas liés. Considérons alors une transvection d'hyperplan H , avec $v \notin H$ et de direction Av . Soit B sa matrice. Alors $ABv = A(v + \alpha Av)$, avec $\alpha \in k^*$, donc $ABv = Av + \alpha A^2(v)$. Par ailleurs, $BAv = Av$. Les deux quantités ne peuvent être égales. Donc pour tout $v \in k^n$ les vecteurs v et Av sont liés. Ce qui implique que la matrice A n'a d'éléments non-nuls que sur la diagonale. La commutation avec les transvections permet de montrer que tous les coefficients diagonaux sont égaux. Si en effet on s'est donné deux vecteurs propres v_1 et v_2 , associés aux valeurs propres λ_1 et λ_2 , indépendants il suffit de prendre une transvection de direction v_2 et telle que v_1 ne soit pas dans l'hyperplan associé

à la transvection. Le calcul est analogue à celui qu'on vient de faire : si B est la matrice associée on a

$$ABv_1 = A(v_1 + \alpha v_2) = \lambda_1 v_1 + \alpha \lambda_2 v_2$$

et

$$BAv_1 = \lambda_1 Bv_1 = \lambda_1 v_1 + \alpha \lambda_1 v_2$$

avec $\alpha \neq 0$. Le résultat suit de l'égalité de ces expressions. ■

Passons maintenant à l'étude du groupe dérivé. Nous ne la ferons ici que pour $n \geq 3$. Le cas $n = 2$ sera traité en exercice.

Théorème 4. *Supposons que $n \geq 3$. Le groupe dérivé de $\mathrm{GL}_n(k)$ est égal à $\mathrm{SL}_n(k)$. De même le groupe dérivé de $\mathrm{SL}_n(k)$ est $\mathrm{SL}_n(k)$ lui-même.*

Démonstration. Il suffit de démontrer que le groupe dérivé contient les matrices $E_{i,j}(\lambda)$. Puisque n est supérieur ou égal à 3 on peut choisir un indice k distinct de i et j . On forme alors le commutateur

$$E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(\lambda)^{-1}E_{j,k}(1)^{-1}.$$

Il est égal à

$$E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(-\lambda)E_{j,k}(-1),$$

puis tous calculs faits à $E_{j,i}(-\lambda)$. Le résultat suit. Pour faire le calcul on peut aussi calculer sur les applications linéaires associées. Dans la base canonique il suffit de calculer l'action sur e_i , e_j et e_k . Toutes les matrices dans le commutateur laissent e_i fixe. Sinon on a :

$$\begin{aligned} E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(-\lambda)E_{j,k}(-1)(e_j) &= E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(-\lambda)(e_j - e_k) = \dots \\ \dots &= E_{k,i}(\lambda)E_{j,k}(1)(e_j - e_k + \lambda e_i) = E_{k,i}(\lambda)(e_j - \lambda e_i) = e_j - \lambda e_i, \end{aligned}$$

et

$$\begin{aligned} E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(-\lambda)E_{j,k}(-1)(e_k) &= E_{k,i}(\lambda)E_{j,k}(1)E_{k,i}(-\lambda)(e_k) \\ &= E_{k,i}(\lambda)(e_k - \lambda e_i) = e_k. \end{aligned}$$

Le résultat suit. ■

Sous-groupes distingués du groupe linéaire

✎ Pour achever cette section nous allons démontrer :

Théorème 5. *Supposons que n soit supérieur ou égal à 3. Soit alors H un sous-groupe distingué du groupe spécial linéaire $\mathrm{SL}_n(k)$, alors H est soit contenu dans le centre de H qui est l'ensemble des matrices diagonales de déterminant 1, soit égal à $\mathrm{SL}_n(k)$ tout entier.*

Démonstration. On va supposer que l'on travaille dans le groupe $\text{SL}(E)$ et démontrer l'énoncé équivalent. C'est-à-dire démontrer qu'un sous-groupe distingué est, soit contenu dans le sous-groupe des homothéties, soit égal au groupe tout entier.

Soit donc H un sous-groupe distingué. Supposons donné un élément $f \in H$ et supposons que f ne soit pas une homothétie. Il existe donc un vecteur v tel que v et $f(v)$ forment un système libre. Considérons alors une transvection non-triviale g d'hyperplan H , d'équation α , contenant v et $f(v)$ et de direction parallèle à v (rappelons que $n \geq 3$).

On vérifie par un calcul direct que f et g ne commutent pas. En effet soit $w \notin H$, on a $f(w) = aw + u$, $u \in H$, $a \in k$. Calculons $g \circ f(w)$, on obtient $g \circ f(w) = aw + u + a\alpha(w)v$. D'un autre côté on a : $f \circ g(w) = f(w + \alpha(w)v) = aw + u + \alpha(w)f(v)$. Si ces deux expressions sont égales comme $\alpha(w) \neq 0$ on a $av = f(v)$, ce qui est exclu.

L'application linéaire $f \circ g \circ f^{-1} \circ g^{-1}$ est dans le sous-groupe H , car f et $g \circ f^{-1} \circ g^{-1}$ sont dans H .

C'est le produit des deux transvections g^{-1} et $f \circ g \circ f^{-1}$, car :

Lemme 3. Soit E un espace vectoriel de dimension finie sur un corps k . Soit f une transvection de E d'hyperplan H et g un automorphisme linéaire de E . Alors l'application $g \circ f \circ g^{-1}$ est une transvection d'hyperplan $g(H)$.

Démonstration. On a $f(x) = x + \alpha(x)v$, avec $v \in H$, pour tout $x \in E$. Donc

$$g \circ f \circ g^{-1}(x) = g(g^{-1}(x) + \alpha(g^{-1}(x))v) = x + (\alpha \circ g^{-1}(x))g(v). \quad \square$$

Le résultat va alors suivre du corollaire et du lemme suivant sur les transvections que nous énonçons et démontrerons après.

Lemme 4. Si la dimension de E est supérieure ou égale à 3, deux transvections quelconques de E sont conjuguées dans $\text{SL}(E)$.

Démonstration. Raisonnons matriciellement, il est suffisant de le faire pour $n = 3$. Soient A et A' les matrices de deux transvections. On peut pour chacune d'entre elle trouver une base dans laquelle leur matrice est :

$$J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ceci implique que les matrices A et A' sont conjuguées dans $\text{GL}_3(k)$. Si P et P' sont les matrices de changement de bases introduites plus haut on a $P^{-1}AP = J$ et $P'^{-1}A'P' = J$. Donc on a $P'P^{-1}A(P^{-1}P')^{-1} = A'$. Il reste à montrer que l'on peut choisir $Q = P'P^{-1}$ de déterminant 1. Soit donc $a \in k^*$ son déterminant. On

vérifie sans peine que l'on peut la remplacer dans l'équation précédente par QD où

$$D = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^{-1} \end{pmatrix}.$$

La matrice est de déterminant 1. □

On a aussi :

Lemme 5. *Le produit de deux transvections d'hyperplan commun est encore une transvection.*

Démonstration. Soit f et g les deux transvections d'un espace vectoriel E . Le produit laisse fixe l'hyperplan H commun aux deux transvections. Si on choisit un vecteur quelconque w n'appartenant pas à H on a $g \circ f(w) = aw + u$, avec $a \in k$, et $u \in H$. Comme f et g sont de déterminant 1 leur produit est de déterminant 1. Si on calcule le déterminant du produit dans une base de E constitué d'une base de H et de w on trouve que $a = 1$. On en déduit que le produit est une transvection. □

Fin de la démonstration du théorème. Puisque f et g ne commutent pas le sous-groupe H contient nécessairement une transvection non-triviale. Comme elles sont toutes conjuguées il les contient toutes et est donc égal à $SL(E)$. ■

Exercices

I. Diagonalisation, triangularisation

1. a) On rappelle que étant donné une décomposition en somme directe $F \oplus G$ d'un espace vectoriel E le projecteur sur F parallèlement à G est l'application linéaire qui est l'identité sur F et qui est nulle sur G . Montrer qu'une application linéaire p est un projecteur si et seulement si $p^2 = p$. Quelles sont les valeurs propres de p ainsi que ces sous-espaces propres ?
- b) Dans cette question la caractéristique du corps est supposée différente de 2. Démontrer que pour que la somme de deux projecteurs f et g soit un projecteur, il faut et il suffit que $f \circ g + g \circ f = 0$, puis qu'il faut et suffit que $f \circ g = g \circ f = 0$. Démontrer que pour que deux projecteurs f et g aient même noyau, il faut et il suffit que $f = f \circ g$ et $g = g \circ f$.
2. a) Soit E un espace vectoriel de dimension finie sur un corps de caractéristique 0. Soient $p_1, \dots, p_n$ des projecteurs de E . On pose :

$$p = p_1 + \dots + p_n.$$

Montrer que si $p_i \circ p_j = 0$ si $i \neq j$, alors p est un projecteur de E .

- b) On revient au cas général, montrer que $\text{Imp} \subset \text{Imp}_1 + \dots + \text{Imp}_n$. On suppose de plus p est un projecteur. Montrer que :

$$\text{rg}(p) = \text{Tr}(A)$$

où A est la matrice de p dans une base quelconque et on rappelle que la trace $\text{Tr}(A)$ d'une matrice $A = (a_{i,j})$ est la somme $a_{1,1} + a_{2,2} + \dots + a_{n,n}$.

- c) Sous les hypothèses de la question précédente montrer que

$$\text{rg}(p) = \text{rg}(p_1) + \dots + \text{rg}(p_n).$$

En déduire que la somme Imp est somme directe des sous-espaces Imp_i .

- d) Démontrer que si p est un projecteur

$$p_i \circ p_j = 0 \quad \forall i \neq j.$$

3. a) (Décomposition de Dunford) Soient E un espace vectoriel sur $\mathbb{C}$ et u un endomorphisme de E . Soient F_λ les sous-espaces caractéristiques de u et soit π_λ la projection sur F_λ parallèlement à la somme directe des sous-espaces caractéristiques de u distinct de F_λ . On rappelle que E est somme directe des sous-espaces caractéristiques.

Montrer que π_λ peut s'exprimer comme un polynôme en u .

- b) Montrer que u peut s'écrire comme somme d'un endomorphisme d diagonalisable et d'un endomorphisme nilpotent n tels que $d \circ n = n \circ d$.

- c) Montrer qu'une telle décomposition est unique (on utilisera a).

4. Soit p un nombre premier, et soit $\mathbb{F}_{p^n}$ le corps à p^n éléments. Déterminer le polynôme caractéristique de l'automorphisme de Frobenius vu comme application linéaire d'espaces vectoriels sur $\mathbb{F}_p$ de $\mathbb{F}_{p^n}$ sur lui-même.

5. Y-a-t-il un élément d'ordre 5 (d'ordre 7) dans le groupe $GL_2(\mathbb{Q})$?

6. Montrer qu'une matrice à coefficients complexes (n,n) d'ordre fini est diagonalisable et que ses valeurs propres sont de module 1.

Montrer qu'un ensemble fini de matrices à coefficients complexes d'ordre fini et commutant deux à deux sont diagonalisables dans une même base. Que peut on dire si l'ensemble de matrices n'est pas fini ?

7. La trace d'une matrice $A = (a_{i,j})$ est la somme $a_{1,1} + a_{2,2} + \dots + a_{n,n}$. On la note $\text{Tr}(A)$. Montrer que c'est aussi la somme des valeurs propres de la matrice. Montrer que pour toute paire de matrices de taille (n,n) , X et Y , on a $\text{Tr}(XY) = \text{Tr}(YX)$. Montrer que si P est inversible $\text{Tr}(P^{-1}AP) = \text{Tr}(A)$.

8. a) Soit E un espace vectoriel complexe de dimension finie et soit T un sous-ensemble de l'espace $\mathcal{L}(E)$ des applications linéaires de E constitué par des applications commutant deux à deux.

Soit f un élément de S , soit λ une valeur propre de f et soit E_λ le sous-espace propre associé. Montrer qu'il est stable par tout élément de T , c'est-à-dire que pour tout élément $g \in T$ on a $g(E_\lambda) \subset E_\lambda$.

b) Montrer que les éléments de T ont au moins un vecteur propre en commun.

c) En déduire par récurrence que l'on peut trouver une base de E dans laquelle tous les éléments de T ont une matrice triangulaire supérieure.

9. Soit E un espace vectoriel sur un corps k , et soit $\mathcal{L}_\mathbb{C}(E)$ l'espace des applications linéaires de E dans lui-même. Soient $f, g \in \mathcal{L}_k(E)$, on appelle crochet de Lie de f et g et on note $[f, g]$ l'application linéaire $fg - gf$. Montrer que pour tous $f, g, h \in \mathcal{L}_k(E)$ on a l'identité de Jacobi :

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0.$$

10. Soit E un espace vectoriel complexe de dimension finie, et soient $f, g, h \in \mathcal{L}(E)$ l'espace des applications linéaires de E dans lui-même. On suppose que $[f, g] = h$, $[f, h] = [g, h] = 0$. Montrer que f , g et h sont triangularisables dans une base commune.

11. Soit E un espace vectoriel complexe de dimension finie, et soient $f, g, h \in \mathcal{L}(E)$ l'espace des applications linéaires de E dans lui-même. On suppose que $[f, g] = h$, $[f, h] = f$ et $[g, h] = 0$. Montrer que f , g et h sont triangularisables dans une base commune.

12. Soit L un sous-espace vectoriel de $\mathcal{L}(E)$ stable par crochet de Lie, c'est-à-dire tel que : $\forall f, g \in L$ $[f, g] \in L$. Soit de plus K un sous-espace vectoriel de codimension 1 de L (i.e. on a $\dim_\mathbb{C} L - \dim_\mathbb{C} K = 1$). On suppose que pour tout $f \in L$ et tout $g \in K$ $[f, g] \in K$. On suppose enfin que les éléments de K ont un vecteur propre commun. On veut montrer qu'il en est de même pour les éléments de L .

a) Soit v un vecteur propre commun aux éléments de K : on a $k(v) = \lambda(k)v$, montrer que λ est une forme linéaire sur K .

b) Soit $W = \{u \mid k(u) = \lambda(k)u, \quad \forall k \in K\}$. Montrer que W est sous-espace vectoriel.

c) Montrer que l'on peut trouver un élément $x \in L - K$ tel que $L = K \oplus \mathbb{C}x$ (utiliser l'hypothèse sur la codimension).

d) On va montrer que W est stable par x . Pour cela on introduit le sous-espace W_i de E engendré par les éléments $x^a v$, $0 \leq a \leq i - 1$, $v \in W$.

Montrer que $W_i \subset W_{i+1}$ et que la suite des W_i stabilise, c'est-à-dire que l'inclusion est une égalité pour tout i assez grand. Soit $\mathcal{W}$ le sous-espace $\cup_i W_i$.

- e) Montrer qu'il existe une base bien choisie de $\mathcal{W}$ dans laquelle tout élément k de K a une matrice triangulaire supérieure dont les éléments diagonaux sont égaux à $\lambda(k)$.
- f) Montrer que $\mathcal{W}$ est stable par x (c'est-à-dire que $x(\mathcal{W}) \subset \mathcal{W}$), par tout élément de K , et donc par tout élément de L . En déduire que $\lambda([k, l]) = 0$, où $k \in K$, $l \in L$.
- g) En déduire que W est stable par x , puis le résultat.

13. Déduire de l'exercice précédent le théorème de Lie dont l'énoncé suit. Soit E un espace vectoriel complexe de dimension finie. Soit $L \subset \mathcal{L}(E)$ un sous-espace stable par crochet de Lie. On définit par récurrence $L^{(n)}$ par $L^{(0)} = L$ et $L^{(n)} = \{[f, g] \mid f, g \in L^{(n-1)}\}$.

On suppose que pour n assez grand $L^{(n)} = 0$. Montrer que l'on peut trouver une base dans laquelle tous les éléments de L ont une matrice diagonale. On fera une récurrence.

14. ** Soit U le groupe multiplicatif des nombres complexes de module 1, et soit $\alpha : U \rightarrow \mathrm{GL}_n(\mathbb{C})$ un homomorphisme. On suppose de plus que l'application α est continue. Montrer qu'il existe une matrice inversible A tel que pour tout $s \in U$ la matrice $A^{-1}\alpha(s)A$ soit diagonale.

15. Soient A et B deux matrices (n, n) à coefficients réels. On suppose qu'il existe une matrice (n, n) inversible à coefficients complexes P telle que $A = P^{-1}BP$. On veut montrer que l'on peut trouver une matrice (n, n) inversible à coefficients réels Q telle que $A = Q^{-1}BQ$.

- a) Montrer que l'ensemble des matrices S (n, n) à coefficients complexes telle que $AS = SB$ est un sous-espace de $M_n(\mathbb{C})$.
- b) On considère l'application déterminant sur ce sous-espace. C'est une fonction polynôme à valeurs dans $\mathbb{C}$. Montrer que si elle est nulle sur toutes les matrices à coefficients réels elle est nulle sur tout le sous-espace, conclure.

16. Formuler l'exercice analogue du précédent en remplaçant $\mathbb{R}$ par un corps infini k quelconque et $\mathbb{C}$ par un corps K contenant k .

17. Une matrice symétrique à coefficients complexes est-elle diagonalisable ?

18. Théorème de Maschke. Soient G un groupe fini et E un espace vectoriel complexe de dimension finie. Et soit α un homomorphisme de G dans $\mathrm{GL}(E)$. Soit F un sous-espace vectoriel de E stable par toutes les applications linéaires $\alpha(g)$, $g \in G$. On va montrer que F a un supplémentaire F' stable par toutes les applications linéaires $\alpha(g)$, $g \in G$.

- a) Soit p un projecteur sur F , c'est-à-dire une application linéaire telle que $\text{Im}(p) = F$ et $p^2 = p$. On considère l'application linéaire :

$$\frac{1}{\#G} \sum_{g \in G} \alpha(g^{-1}) \circ p \circ \alpha(g).$$

Montrer que $\pi^2 = \pi$ et que pour tout g on a $\pi \circ \alpha(g) = \alpha(g) \circ \pi$.

- b) Montrer que $\text{Im}(\pi) = F$, conclure.

19. Soit E un espace vectoriel de dimension finie sur un corps k . Soit f une application linéaire de E dans E et soit P son polynôme minimal.

- a) Montrer qu'il existe $x \in E$ tel que $Q(f)(x) \neq 0$ pour tout polynôme $Q \in k[X]$ tel que $Q \neq 0$ et $|Q| < |P|$. Le vecteur x est fixé dans la suite de l'exercice.

- b) On note E_x le plus petit sous-espace de E stable par f et contenant x . En donner une base. Montrer que si F est un sous-espace vectoriel de E_x stable par f il existe un polynôme $H_F \in k[X]$ tel que :

- $H_F \mid P$,
- et que F est le plus petit sous-espace contenant $H_F(f)(x)$.

- c) Soit $y \in E$. On appelle polynôme annulateur de y et on note P_y un générateur de l'idéal $\{Q \in k[X] \mid Q(f)(y) = 0\}$. Montrer que $P_y \mid P$. Soit E_y le plus petit sous-espace de E stable par f et contenant y . Montrer qu'il existe un polynôme $H \in k[X]$ tel que :

- $H \mid P_y$,
- $E_x \cap E_y$ est de la forme $E_{H(f)(x)}$. Ce dernier espace étant le plus petit contenant $H(f)(x)$ et stable par f .

- d) Soit $E_x \subset E$ et soit F un sous-espace stable par f et tel que $E_x \subset F \subset E$. Supposons donnée une application linéaire ψ de F sur E_x qui commute aux restrictions de f et qui, restreinte à E_x , est l'identité. Soit $y \in E - F$, et soit F' le plus petit sous-espace stable par f contenant F et y . Montrer que l'on peut étendre l'application ψ en une application $\bar{\psi}$ de F' sur E_x qui possède les mêmes propriétés. C'est-à-dire qui est l'identité sur E_x et telle $\bar{\psi} \circ i = \text{Id}_{E_x}$, ici i désigne l'inclusion de E_x dans F' .

On pourra s'inspirer de l'analyse des groupes abéliens de type fini de torsion.

- e) En déduire que E_x admet un supplémentaire stable par f .

20. À l'aide de l'exercice 19. démontrer les énoncés sur les invariants de similitude des applications linéaires. On commencera par écrire $E = E_x \oplus E'$, le vecteur x ayant les propriétés de l'exercice 16 et E' étant stable par f . Puis on fera une hypothèse de récurrence sur la dimension de E . Ceci pour l'existence. Pour l'unicité on s'inspirera du cas des groupes abéliens de type fini.

21. À l'aide des exercices précédents, déterminer les classes de conjugaison dans $\text{GL}_3(\mathbb{F}_2)$ et dans $\text{GL}_4(\mathbb{F}_2)$.
22. Montrer que les entiers d_i qui apparaissent dans la description d'une matrice de Jordan générale (voir après la démonstration de la proposition 1 dans la section concernée) sont déterminés par les dimensions de $E_1, \dots, E_{[\frac{k}{2}]}$, gardant les notations de la section.

II. Applications topologiques

1. Soit n un entier, et soit $\text{GL}_n^+(\mathbb{R})$ le groupe des matrices (n, n) à coefficients réels de déterminant strictement positif. Montrer que ce groupe est connexe.
2. a) Soit $A \in \text{GL}_n(\mathbb{C})$ une matrice telle qu'il existe un sous-groupe compact K de $\text{GL}_n(\mathbb{C})$ qui contienne A . Montrer que toutes les valeurs propres de A sont de module 1.
- b) Montrer que A est diagonalisable. On utilisera la réduction de Jordan.
3. Montrer que le groupe $\text{GL}_1(\mathbb{C})$ est homéomorphe au produit de l'ensemble des nombres complexes de module 1 par l'espace vectoriel $\mathbb{R}$. Est-ce un isomorphisme de groupes ?
4. a) Soit $A = (a_{i,j}) \in \text{M}_n(\mathbb{C})$, montrer que si pour tout $i = 1, \dots, n$ on a $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$ la matrice A est inversible.
- b) Soit $A = (a_{i,j}) \in \text{M}_n(\mathbb{C})$, et soit λ une valeur propre de A . Montrer qu'il existe i tel que $|a_{i,i} - \lambda| \leq \sum_{j \neq i} |a_{i,j}|$.
5. a) Soit $A = (a_{i,j}) \in \text{M}_n(\mathbb{C})$, montrer que la série de terme général $\frac{A^n}{n!}$ converge dans l'espace vectoriel normé $\text{M}_n(\mathbb{C})$. On note $\exp(A)$ sa somme, on l'appelle l'exponentielle de A . Montrer que cette matrice est inversible et calculer son inverse. On montrera que si A et B commutent il en est de même de $\exp(A)$ et $\exp(B)$.
- b) Déterminer les valeurs propres de $\exp(A)$ en fonction de celles de A .
- c) Montrer que l'application $t \mapsto \exp(tA)$ est un homomorphisme de $\mathbb{R}$ dans $\text{GL}_n(\mathbb{C})$.
- d) On suppose donné un homomorphisme γ de $\mathbb{R}$ dans $\text{GL}_n(\mathbb{C})$ qui est une fonction de classe C^1 de t . Établir une équation différentielle satisfaite par γ et en déduire que γ est de la forme $\exp(tA)$.
- e) Par un procédé de régularisation montrer que tout homomorphisme continu γ de $\mathbb{R}$ dans $\text{GL}_n(\mathbb{C})$ est une fonction de classe C^1 donc de la forme $\exp(tA)$.

6. a) Montrer que l'application exponentielle de $M_n(\mathbb{C})$ dans $GL_n(\mathbb{C})$ est différentiable et calculer sa différentielle en l'origine.
- b) Calculer le développement limité à l'ordre 5 en zéro de l'application de $\mathbb{R}$ dans $GL_n(\mathbb{C})$ donnée par :

$$t \mapsto \exp(tA) \exp(tB) \exp(-tA) \exp(-tB),$$

avec $A, B \in M_n(\mathbb{C})$.

7. Le théorème de Cayley-Hamilton par la topologie.

- a) Démontrer le théorème de Cayley-Hamilton directement pour les matrices diagonalisables.
- b) Si le corps de base est $\mathbb{C}$ démontrer le théorème de Cayley-Hamilton à l'aide de a) et de la densité des matrices diagonalisables.

III. Groupe linéaire

1. Soit $E \cong E_1 \oplus \cdots \oplus E_k$ une décomposition en somme directe d'un espace vectoriel E . Déterminer le sous-groupe de $GL(E)$ qui laisse fixe cette décomposition, c'est-à-dire le sous-groupe des applications linéaires qui laissent fixes chaque E_i globalement. De même, déterminer le sous-groupe de $GL(E)$ qui laisse fixe globalement chacun des sous-espaces $\oplus_{i=1, \dots, \ell} E_i$. On donnera dans ces deux cas une description matricielle.
2. Soit k un corps fini à q éléments. Trouver le nombre d'éléments du groupe $GL_n(k)$.
3. Soit k un corps fini à q éléments. Trouver le nombre de sous-espaces vectoriels de dimension d d'un espace vectoriel E de dimension n .
4. Soit E un espace vectoriel de dimension n sur un corps fini $\mathbb{K}$ à q éléments. Soient k et ℓ des entiers positifs de somme n . Trouver le nombre de décompositions de E en somme directe d'un sous-espace de dimension k et d'un sous-espace de dimension ℓ .
5. Soit k un corps fini à q éléments de caractéristique impaire p . Trouver le nombre de matrices A de $GL_n(k)$ telles que $A^2 = I_n$. Puis si p est de la forme $4\ell + 1$ le nombre de matrices telles que $A^4 = I_n$.
6. a) Montrer que le groupe linéaire $GL_2(\mathbb{F}_2)$ est isomorphe à $\mathfrak{S}_3$. On étudiera l'action du groupe sur $\mathbb{F}_2 \times \mathbb{F}_2$.
- b) Étudier le groupe dérivé de $GL_2(\mathbb{F}_2)$.

7. a) Déterminer le nombre d'éléments de $\mathrm{SL}_2(\mathbb{F}_3)$. Puis déterminer le centre Z de ce groupe.
 b) Montrer que le quotient $\mathrm{SL}_2(\mathbb{F}_3)/Z$ est isomorphe à $\mathcal{A}_4$.
 c) Montrer que le groupe dérivé de $\mathrm{SL}_2(\mathbb{F}_3)$ est isomorphe au quaternionien Q .
 d) Montrer que $\mathrm{SL}_2(\mathbb{F}_3)$ est un produit semi-direct Q et $\mathbb{Z}/3\mathbb{Z}$.
 e) Écrire l'équation des classes pour $\mathrm{SL}_2(\mathbb{F}_3)$.
8. Montrer que le quotient du groupe $\mathrm{GL}_2(\mathbb{F}_5)$ par son centre est isomorphe à $\mathfrak{S}_5$.
 Montrer que le quotient du groupe $\mathrm{GL}_2(\mathbb{F}_4)$ par son centre est isomorphe à $\mathcal{A}_5$.
9. Montrer que le groupe $\mathrm{GL}_n(k)$ est un produit semi-direct de $\mathrm{SL}_n(k)$ et de k^* .
10. Montrer que le sous-groupe dérivé de $\mathrm{SL}_2(k)$ est $\mathrm{SL}_2(k)$ lui-même dès que le corps a strictement plus de 3 éléments.
11. Soit E un espace vectoriel sur un corps k . Montrer que deux dilatations quelconques sont conjuguées.
12. Soit E un espace vectoriel sur un corps k . Soit f une application linéaire de E dans lui-même laissant invariant points par points un hyperplan H . Montrer que f est une transvection si et seulement si $\det(f) = 1$.
13. Le groupe $(\mathbb{Z}/2\mathbb{Z})^3$ est-il un sous-groupe de $\mathrm{GL}_3(\mathbb{C})$? On pose la même question pour $(\mathbb{Z}/2\mathbb{Z})^4$.
14. Soit n un entier donné. Pour quelles valeurs de k le groupe $(\mathbb{Z}/2\mathbb{Z})^k$ est-il un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$?
15. Soit n un entier donné. Pour quelles valeurs de k le groupe $(\mathbb{Z}/3\mathbb{Z})^k$ est-il un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$?
16. Invariants de Dickson.
 On fait agir le groupe $\mathrm{GL}_2(\mathbb{F}_2)$ sur les polynômes à coefficients dans $\mathbb{F}_2$ en 2 indéterminées X et Y par la formule :
- $$A(P)(X, Y) = P(aX + bY, cX + dY),$$
- où $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_2)$.
- a) Montrer que les polynômes $w_2 = X^2 + XY + Y^2$ et $w_3 = XY(X + Y)$ sont invariants par cette action.
- b) Soit $P \in \mathbb{F}_2[X, Y]$ invariant par cette action. Montrer qu'il existe un unique polynôme R à deux indéterminées à coefficients dans $\mathbb{F}_2$ tel que $P = R(w_2, w_3)$.
17. Montrer que les matrices triangulaires supérieures forment un 2-sous-groupe de Sylow de $\mathrm{GL}_n(\mathbb{F}_2)$.

I. Diagonalisation, triangularisation

1. L'exercice est très classique. Notons qu'un endomorphisme tel que $p^2 = p$ est diagonalisable car il annule un polynôme à racines simples. On notera aussi que l'on passe de $f \circ g = -g \circ f$ à $f \circ g = g \circ f = 0$ en multipliant à gauche par f et en utilisant $f^2 = f$ et $f \circ g = -g \circ f$.
2. On utilise l'exercice 1. En particulier on note que dans une base de diagonalisation, et comme la seule valeur propre non nulle est 1 la somme des éléments diagonaux de la matrice de p est égale à la dimension de l'image de p . On notera que ceci n'est vrai qu'en caractéristique 0. Pour conclure on précompose par p_i la relation $p = p_1 + \dots + p_n$.
4. L'automorphisme de Frobenius annule le polynôme $X^n - 1$. Considérons alors un générateur α du groupe des éléments inversibles du corps, les éléments $F^i(\alpha)$, $1 \leq i \leq n$ constituent une base $\mathbb{F}_{p^n}$ sur $\mathbb{F}_p$. Comme le polynôme minimal de F divise $X^n - 1$ on en déduit facilement qu'il est égal à $X^n - 1$, qui est donc (au signe prêt et à cause du degré) le polynôme caractéristique.
5. Étudier le polynôme caractéristique, dont on observera qu'il divise $X^4 + X^3 + X^2 + X + 1$ et montrer qu'il ne peut être à coefficients dans $\mathbb{Q}$.
6. Si l'ordre de la matrice est n elle annule $X^n - 1$, et est donc diagonalisable car ce polynôme n'a que des racines simples.
10. On montre que f , g et h ont un vecteur propre en commun. Pour ce faire on considère un sous-espace propre F de h . On montre qu'il est stable par f et g . Puis on considère la trace de h restreinte à ce sous-espace (c'est-à-dire la trace de n'importe quelle matrice associée). Comme h est égal à $f \circ g - g \circ f$ cette trace est nulle. Comme h sur ce sous-espace est une homothétie, on en déduit que h est nulle sur ce sous-espace, c'est-à-dire que f et g commutent sur ce sous-espace. Le résultat suit.
11. On montre que f , g et h ont un vecteur propre en commun. On commence par résoudre le problème pour f et h en considérant un sous-espace propre associé à f . Puis on introduit g .
12. c) Un supplémentaire de K est de dimension 1.
 - d) La suite des W_i stabilise car $W_i \subset W_{i+1}$ et la dimension de W_i est bornée supérieurement. Donc la suite $\dim W_i$ est constante à partir d'un certain rang. À partir de ce rang l'inclusion $W_i \subset W_{i+1}$ devient une égalité.
 - e) On choisit une base de W , soit $v_1, \dots, v_k$ puis on introduit les vecteurs $x\hat{a}(v_i)$. Montrer qu'il existe une base bien choisie de $\mathcal{W}$ dans laquelle tout élément k de K a une matrice triangulaire supérieure dont les éléments diagonaux sont égaux à $\lambda(k)$.
 - f) Introduire une trace (voir exercice 7).
14. Montrer que pour s quelconque la matrice $\alpha(s)$ est diagonalisable, puis appliquer le théorème de diagonalisation simultanée pour les matrices commutant.
15. b) On raisonne par l'absurde. S'il n'y a pas de solutions l'application déterminant sur le sous-espace du a) est une fonction polynôme qui s'annule sur toute matrice à coefficients réels appartenant au sous-espace.

Les équations du sous-espace sont des formes linéaires à coefficients réels. On peut donc choisir une base du sous-espace constituée par des matrices à coefficients réels. Désignons par $s_1, \dots, s_k$ les coordonnées d'une solution S de $AS = SB$ dans cette base. La fonction $\det(S)$ est polynomiale en les s_i et nulle quand tous les s_i sont réels.

Ceci est impossible par hypothèse, car il y a une solution inversible à valeurs complexes.

16. Adapter la démonstration précédente.

17. Non, considérer des matrices $(2, 2)$ ayant une valeur propre double.

18. b) Considérer $\text{Ker}(\pi) = F$.

21. Les décrire à l'aide de leur invariants de similitude. Par exemple dans le cas $n = 3$ les invariants de similitude possibles sont :

- $X^3 + X + 1$, $X^3 + X^2 + 1$, qui correspondent aux éléments d'ordre 7,
- $X^3 + 1$ qui correspond aux éléments d'ordre 3,
- $(X + 1)^3$ qui correspond aux éléments d'ordre 4,
- $X^2 + 1, X + 1$ qui correspond aux éléments d'ordre 2,
- $X + 1, X + 1, X + 1$ qui correspond à l'élément neutre.

À l'aide des matrices compagnons on détermine les centralisateurs et donc le nombre d'éléments des classes de conjugaison. On peut utiliser d'autres renseignements.

Par exemple dans le premier cas l'ordre de la classe de conjugaison est un diviseur de 24. En effet le centralisateur de l'élément contient le sous-groupe engendré par l'élément qui est d'ordre 7, et l'ordre du groupe linéaire considéré est 7×24 . En utilisant le fait qu'il y a 1 ou 8 ou 15, ..., 7-sous-groupes de Sylow on déduit que la classe de conjugaison a 24 éléments (on observe facilement que le 7-Sylow contenant la matrice compagnon de $X^3 + X + 1$ n'est pas distingué).

Dans le second on montre que le nombre d'éléments d'ordre 3 est aussi le nombre de décomposition de $\mathbb{F}_2^3$ en somme directe d'un sous-espace de dimension 2 et de dimension 1, soit 28, multiplié par 2 car on a le choix de deux applications d'ordre 3 sur le sous-espace de dimension 2.

Dans le troisième on trouve que le centralisateur est d'ordre 4, ce qui est fait plus facilement en utilisant une forme de Jordan. La classe de conjugaison a donc 42 éléments.

Enfin pour ce qui est des éléments d'ordre 2 on trouve un centralisateur à 8 éléments, soit une classe de conjugaison avec 21 éléments.

II. Applications topologiques

1. Commencer par traiter le problème pour $n = 2$. Puis faire un raisonnement par récurrence en montrant que l'on peut relier une matrice quelconque par un arc continu à une matrice dont le seul terme non nul sur la première colonne soit celui situé sur la première ligne. Montrer que l'on peut supposer ce terme positif. On pourra multiplier par une matrice $(2, 2)$.
2. a) b) Étudier la convergence des suites A^n et $(A^{-1})^n$. Le fait que la première admette une valeur d'adhérence implique que toutes les valeurs propres sont de modules inférieur ou égal à 1. Pour démontrer cela on peut supposer A triangulaire et observer que la norme de la matrice doit rester bornée. La seconde suite donne l'autre inégalité. En utilisant la réduction de Jordan on montre que si A n'est pas diagonalisable la suite A^n n'admet pas de valeur d'adhérence car la norme de la matrice tend vers l'infini.

4. a) Montrer que l'image d'un vecteur non nul a au moins une composante non nulle en considérant la composante non nulle de plus grand module du vecteur d'origine.
5. d) Dériver $\gamma(s+t) = \gamma(s)\gamma(t)$ par rapport à t , et établir une équation différentielle que l'on résoudra.
6. a) La différentielle est l'identité.
7. On fait un calcul direct pour les matrices diagonales, qui passent au matrice diagonalisables. Puis, si on note c_A le polynôme caractéristique de A on montrera que la fonction qui à A associe $c_A(A)$ est continue, en interprétant l'espace d'arrivée comme un espace vectoriel sur $\mathbb{C}$ de dimension finie.

III. Groupe linéaire

2. Il s'agit de trouver le nombre de bases de k^n . On trouve

$$(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1}).$$

3. On trouve

$$\frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{d-1})}{(q^d - 1)(q^d - q) \cdots (q^d - q^{d-1})}.$$

4. Si $k \neq \ell$ on trouve

$$\frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})}{(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1})(q^\ell - 1)(q^\ell - q) \cdots (q^\ell - q^{\ell-1})}.$$

- Si $k = \ell$ on trouve

$$\frac{1}{2} \times \frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})}{((q^k - 1)(q^k - q) \cdots (q^k - q^{k-1}))^2}.$$

5. Montrer qu'une matrice telle que $A^2 = I$ est diagonalisable et est déterminée par la décomposition de l'espace vectoriel en somme directe des sous-espaces propres. On est donc ramenée à chercher le nombre de décomposition de k^n en somme directe de sous-espaces propres. Dans le cas $A^4 = I$ montrer que de même que A est diagonalisable et reprendre le même raisonnement.

7. a) 24, le centre est isomorphe $\mathbb{Z}/2\mathbb{Z}$.

b) Faire agir le groupe $\mathrm{SL}_2(\mathbb{F}_3)$ sur les droites de $\mathbb{F}_3^2$.

c) On commence par montrer qu'un 2-sous-groupe de Sylow de $\mathrm{SL}_2(\mathbb{F}_3)$ est isomorphe au quaternionien Q . Pour cela on peut utiliser la classification des groupes d'ordre 8.

8. Étudier l'action sur l'ensemble des droites de $\mathbb{F}_5^2$ et $\mathbb{F}_4^2$.

13. Pour cet exercice et les deux suivants on utilisera la diagonalisation simultanée des matrices.

On montrera que le sous-groupe s'identifie, à un changement de base près à un ensemble de matrices diagonales. Puis on tiendra compte des valeurs propres. La réponse est $k \leq n$.

16. b) Soit $P \in \mathbb{F}_2[X, Y]$ invariant par l'action considérée. Montrer que $P(0, Y)$ s'écrit sous la forme $Q(Y^2)$, puis montrer que $P - Q(w_2)$ est divisible par w_3 . Conclure en utilisant un argument de récurrence sur le degré.

17. Calculer la plus grande puissance de 2 divisant l'ordre de $\mathrm{GL}_n(\mathbb{F}_2)$ en utilisant le premier exercice de cette section. Puis calculer le nombre de matrices triangulaires supérieures inversibles.

Chapitre 6

Formes bilinéaires et sesquilinéaires Groupes orthogonaux et unitaires

Dans tout ce chapitre, tous les espaces vectoriels seront supposés de dimension finie. On étudiera les formes bilinéaires, sesquilinéaires et quadratiques sur ces espaces, ainsi que celle des groupes qui leur sont associés. Dans une large mesure c'est un chapitre de révision, quelques pistes de développement y sont ouvertes. La première section consistera donc en rappels fondamentaux, la seconde à une étude un peu plus détaillée du groupe orthogonal euclidien, et en particulier de ses systèmes de générateurs. La troisième est consacrée à une étude particulière de la dimension 3 et 4. Enfin, la dernière section est consacrée au groupe unitaire.

1. Formes bilinéaires et sesquilinéaires

Soit k un corps quelconque, et soit E un espace vectoriel de dimension finie sur k . Nous commençons par définir les formes bilinéaires sur l'espace vectoriel E .

Définition 1

On appelle *forme bilinéaire* sur E une application b de $E \times E$ dans k telle que pour tous $x, y, z \in E$ et $\lambda \in k$ on ait :

- $b(\lambda x, y) = b(x, \lambda y) = \lambda b(x, y)$,
- $b(x + y, z) = b(x, z) + b(y, z)$, et
- $b(z, x + y) = b(z, x) + b(z, y)$.

La forme b est *symétrique* si $b(x, y) = b(y, x)$ pour tous $x, y \in E$. Elle est dite *antisymétrique* si $b(x, y) = -b(y, x)$ pour tous $x, y \in E$.

On notera que l'ensemble des formes bilinéaires sur E est un espace vectoriel sur k . Le sous-ensemble des formes symétriques est un sous-espace vectoriel. Il en est de même du sous-ensemble des formes antisymétriques.

La notion de forme sesquilinéaire peut être introduite dans un contexte plus général que celui que nous allons choisir. La définition générale se place dans le contexte où est donné un corps k muni d'un automorphisme σ . Nous nous contenterons de considérer le cas où k est le corps $\mathbb{C}$ et σ la conjugaison complexe. Soit donc E un espace vectoriel complexe de dimension finie.

Définition 2

Une *forme sesquilinéaire* sur E est une application b de $E \times E$ dans $\mathbb{C}$ telle que pour tous $x, y, z \in E$ et $\lambda \in \mathbb{C}$ on ait :

- ▶ $b(\lambda x, y) = \bar{\lambda} b(x, y)$, et
- ▶ $b(x, \lambda y) = \lambda b(x, y)$,
- ▶ $b(x + y, z) = b(x, z) + b(y, z)$, et
- ▶ $b(z, x + y) = b(z, x) + b(z, y)$.

La forme est dite *hermitienne* si pour tous $x, y \in E$ on a $b(x, y) = \overline{b(y, x)}$.

On notera que dans le cas hermitien, $b(x, x)$ étant égal à $\overline{b(x, x)}$, est toujours réel.

L'ensemble des formes hermitiennes est un espace vectoriel *réel*.

Représentations matricielles

Décrivons les représentations matricielles de formes bilinéaires et sesquilinéaires. Supposons donnée une base $\{e_1, \dots, e_n\}$ de l'espace vectoriel E . On associe à la forme bilinéaire (resp. sesquilinéaire) b la matrice $B = (b_{i,j})$ définie par

$$b_{i,j} = b(e_i, e_j).$$

Soient alors $x, y \in E$ donnés dans la base considérée par les vecteurs colonnes

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

On a :

$$b(x, y) = {}^t X B Y$$

dans le cas bilinéaire, et :

$$b(x, y) = {}^t \bar{X} B Y$$

dans le cas sesquilinéaire.

La matrice d'une forme bilinéaire symétrique est évidemment symétrique, celle d'une forme antisymétrique est antisymétrique.

Enfin, celle d'une forme sesquilinéaire est hermitienne, c'est-à-dire que

$${}^t \bar{B} = B,$$

où $\bar{B}$ désigne la matrice complexe conjuguée de la matrice B . C'est-à-dire la matrice dont les coefficients sont les nombres complexes conjugués des coefficients de B .

Soit une seconde base $\{e'_1, \dots, e'_n\}$ de l'espace vectoriel E , et soit P la matrice de passage de la première base à la seconde base.

Proposition 1. La matrice B' de la forme bilinéaire b dans la seconde base est ${}^t PBP$.

Démonstration. Soit $P = (c_{i,j})$ la matrice de changement de base. Il suffit de calculer $b(e'_i, e'_j)$. Cette quantité est égale à

$$b\left(\sum_{\ell} c_{\ell,i} e_{\ell}, \sum_k c_{k,j} e_k\right),$$

soit égale à

$$\sum_{\ell,k} c_{\ell,i} c_{k,j} b_{\ell,k}.$$

Cette dernière quantité est bien le coefficient correspondant de la matrice ${}^t PBP$. ■

Dans le cas sesquilinéaire la formule de changement de base est donnée par :

Proposition 1 bis. La matrice B' de la forme sesquilinéaire b dans la seconde base est ${}^t \bar{P}BP$, où $\bar{P}$ désigne la matrice complexe conjuguée.

Démonstration. Cette propriété se démontre de la même façon que la propriété 1. ■

Formes quadratiques associées

Introduisons maintenant les formes quadratiques associées. Nous supposons dans cette sous-section que le corps n'est pas de caractéristique 2.

Définition 3

Soit b une forme bilinéaire symétrique, respectivement sesquilinéaire hermitienne. On appelle forme quadratique associée à b , respectivement quadratique hermitienne, l'application q , respectivement h , de E dans $\mathbb{C}$, respectivement de

E dans $\mathbb{R}$, donnée par

$$v \mapsto b(v, v).$$

Définissons alors une forme quadratique comme étant une application q de E dans k , telle que l'application de $E \times E$ dans k donnée par :

$$b(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

est une forme bilinéaire, appelée la forme polaire de q .

Notons que l'application donnée par la formule précédente est symétrique par construction. La donnée d'une forme quadratique q détermine donc une unique forme bilinéaire symétrique. Énonçons :

Proposition 2. *Soit k un corps de caractéristique différente de 2, et soit E un espace vectoriel sur k . La formule précédente détermine une bijection entre l'espace vectoriel des formes quadratiques sur l'espace vectoriel E et l'espace vectoriel des formes bilinéaires symétriques sur l'espace vectoriel E .*

Définissons une forme quadratique hermitienne comme étant une application h de E dans $\mathbb{R}$, telle que l'application b de $E \times E$ dans $\mathbb{C}$ donnée par :

$$b(x, y) = \frac{1}{4}(h(x + y) - h(x - y)) - \frac{i}{4}(h(x + iy) - h(x - iy))$$

soit une forme sesquilinéaire hermitienne, appelée la forme polaire de h .

Notons qu'elle est hermitienne par construction. La donnée d'une forme quadratique hermitienne h détermine donc une unique forme sesquilinéaire hermitienne. Énonçons :

Proposition 3. *Soit E un espace vectoriel complexe. La formule précédente détermine une bijection entre l'espace vectoriel réel des formes quadratiques hermitiennes et l'espace vectoriel réel des formes sesquilinéaires hermitiennes.*

Définition 4

Soit b une forme bilinéaire symétrique, respectivement une forme sesquilinéaire hermitienne. On appelle noyau de b , et on note $\text{Ker}(b)$, l'ensemble des vecteurs $v \in E$ tels que $b(v, y) = 0$ pour tout $y \in E$.

On dit qu'une forme bilinéaire symétrique, respectivement sesquilinéaire hermitienne, est une forme non dégénérée si et seulement si son noyau est réduit au vecteur nul.

On remarquera que le noyau est un sous-espace vectoriel. Voici une propriété fondamentale des formes non dégénérées.

Proposition 4. Soit b une forme bilinéaire symétrique non dégénérée sur un espace vectoriel E sur un corps k . L'application $E \rightarrow E^*$ qui à $v \in E$ associe l'élément $\bar{v} \in E^*$ donné par :

$$w \longrightarrow b(w, v)$$

est un isomorphisme d'espaces vectoriels sur k .

Démonstration. Vérifions d'abord que l'application ainsi déterminée est linéaire. On a $\overline{v+w}(x) = b(x, v+w) = b(x, v) + b(x, w) = \bar{v}(x) + \bar{w}(x)$, et $\overline{\lambda v}(x) = b(x, \lambda v) = \lambda b(x, v) = \lambda \bar{v}(x)$. D'où le résultat. Maintenant comme E est de dimension finie il est de même dimension que son dual E^* . Il suffit donc de montrer que l'application $w \mapsto \bar{w}$ est injective. Mais son noyau est constitué par l'ensemble des vecteurs w tels que pour tout $x \in E$ on ait $b(w, x) = 0$. C'est-à-dire que son noyau est celui de b qui est réduit à 0 par hypothèse. Le résultat suit. ■

Définition

On appelle rang d'une forme bilinéaire symétrique, respectivement d'une forme sesquilinéaire hermitienne, b ou de la forme quadratique associée, respectivement quadratique hermitienne associée, la quantité $\dim_k(E) - \dim_k(\text{Ker}(b))$, respectivement la quantité $\dim_{\mathbb{C}}(E) - \dim_{\mathbb{C}}(\text{Ker}(b))$.

Proposition 5. Soit b une forme bilinéaire symétrique sur un espace vectoriel E de dimension n . Le rang de b est aussi le rang de la matrice associée A dans une base quelconque.

Démonstration. Observons d'abord que le rang de la matrice associée reste le même quand on change la base.

En représentation matricielle le noyau de la forme bilinéaire symétrique est l'ensemble des X tels que pour tout Y on ait ${}^tYAX = 0$. C'est donc le sous-espace des X tels que $AX = 0$. Si on note $\varrho(A)$ le rang de la matrice A la dimension de ce sous-espace est donc $n - \varrho(A)$. Le résultat suit. ■

Groupes orthogonaux et unitaires

Définition 5

Étant donné un espace vectoriel E sur un corps k et b une forme bilinéaire symétrique sur E on appelle groupe orthogonal de b et on note $O(b)$ le sous-

groupe de $\mathbf{GL}(E)$ constitué par les isomorphismes linéaires f tels que pour tous $v, w \in E$ on ait $b(v, w) = b(f(v), f(w))$.

Étant donné un espace vectoriel E sur un corps k et q une forme quadratique sur E on appelle groupe orthogonal de q et on note $\mathbf{O}(q)$ le sous-groupe de $\mathbf{GL}(E)$ constitué par les isomorphismes linéaires f tels que pour tous $v, w \in E$ on ait $q(v) = q(f(v))$.

Proposition 6. *Si le corps k est de caractéristique différente de 2. Le groupe orthogonal d'une forme bilinéaire symétrique coïncide avec celui de la forme quadratique associée.*

Démonstration. Soit q une forme quadratique sur un espace vectoriel E et b la forme polaire. Il est clair que le groupe de la forme bilinéaire est contenu dans celui de la forme quadratique. En effet, si on a $u \in \mathbf{O}(b)$ on a $b(u(x), u(y)) = b(x, y)$ pour tous $x, y \in E$. Donc en particulier pour $x = y$, ce qui implique que $u \in \mathbf{O}(q)$. Inversement, si $u \in \mathbf{O}(q)$ on a :

$$\begin{aligned} b(u(x), u(y)) &= \frac{1}{2}(q(u(x+y)) - q(u(x)) - q(u(y))) \\ &= \frac{1}{2}(q(x+y) - q(x) - q(y)) = b(x, y). \end{aligned}$$

Donc on a bien $u \in \mathbf{O}(b)$. ■

Définition 6

Étant donné un espace vectoriel complexe E et b une forme sesquilinéaire hermitienne sur E , on appelle groupe unitaire de b et on note $\mathbf{U}(b)$ le sous-groupe de $\mathbf{GL}(E)$ constitué par les isomorphismes linéaires f tels que pour tous $v, w \in E$ on ait $b(v, w) = b(f(v), f(w))$.

Étant donné un espace vectoriel complexe E et h une forme quadratique hermitienne sur E , on appelle groupe unitaire de q et on note $\mathbf{U}(h)$ le sous-groupe de $\mathbf{GL}(E)$ constitué par les isomorphismes linéaires f tels que pour tous $v, w \in E$ on ait $h(v) = h(f(v))$.

Proposition 7. *Le groupe unitaire d'une forme sesquilinéaire hermitienne coïncide avec celui de la forme quadratique hermitienne associée.*

Démonstration. La démonstration se fait comme plus haut. ■

La réduction en carrés et la loi d'inertie

Nous passons maintenant à l'étude de la réduction en carrés d'une forme quadratique, réduction dite de Gauss. Voici le premier théorème de cette sous-section. Le corps de base est toujours supposé de caractéristique différente de 2.

Théorème 1. Soit E un espace vectoriel de dimension finie sur un corps k et soit q une forme quadratique sur E . On peut trouver une famille de formes linéaires $\lambda_1, \dots, \lambda_u$ et des éléments du corps $a_1, \dots, a_u$, tels que pour tout $x \in E$ on ait :

$$q(x) = \sum_i a_i \lambda_i(x)^2.$$

Si on suppose les formes linéaires en question linéairement indépendantes et les éléments a_i tous non nuls, l'entier u est bien déterminé, c'est le rang de la forme quadratique.

Démonstration. Soit $\alpha_1, \dots, \alpha_n$ une base du dual E^* de E . Soit $x \in E$, $q(x)$ est alors un polynôme homogène de degré 2 en les $\alpha_i(x)$. C'est-à-dire que :

$$q(x) = \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x),$$

avec $a_{i,j} \in k$. On va faire un raisonnement par récurrence.

Distinguons deux cas.

Dans le premier on suppose que le coefficient $a_{i,i}$ apparaissant dans l'expression de $q(x)$ est non nul pour un certain indice i . Quitte à réindexer on peut supposer que c'est pour $i = 1$. Dans ce cas on écrit :

$$\begin{aligned} q(x) = & a_{1,1} \left(\alpha_1(x)^2 + \frac{1}{2a_{1,1}} \alpha_1(x) \left(\sum_{j>1} a_{1,j} \alpha_j(x) \right) + \frac{1}{4a_{1,1}^2} \left(\sum_{j>1} a_{1,j} \alpha_j(x) \right)^2 \right) - \dots \\ & \dots - \frac{1}{4a_{1,1}} \left(\sum_{j>1} a_{1,j} \alpha_j(x) \right)^2 + \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x). \end{aligned}$$

On pose $\lambda_1 = \alpha_1 + \frac{1}{2a_{1,1}} \left(\sum_{j>1} a_{1,j} \alpha_j \right)$.

On constate que $\lambda_1, \alpha_2, \dots, \alpha_n$ est encore une base de l'espace dual. La matrice de passage étant triangulaire avec des coefficients non nuls sur la diagonale. On obtient :

$$q(x) = a_{1,1} \lambda_1(x)^2 - \left(\frac{1}{4a_{1,1}} \left(\sum_{j>1} a_{1,j} \alpha_j(x) \right)^2 - \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x) \right).$$

Dans le second cas tous les termes $a_{i,i}$ sont nuls. Quitte à réindexer on peut supposer que le terme $a_{1,2}$ est non nul. Dans ce cas réécrivons $q(x)$ comme suit :

$$\begin{aligned} & a_{1,2} \left(\alpha_1(x) + \sum_{j>2} \frac{a_{2,j}}{a_{1,2}} \alpha_j(x) \right) \left(\alpha_2(x) + \sum_{j>2} \frac{a_{1,j}}{a_{1,2}} \alpha_j(x) \right) - \dots \\ & \dots - a_{1,2} \left(\sum_{j>2} \frac{a_{2,j}}{a_{1,2}} \alpha_j(x) \right) \left(\sum_{j>2} \frac{a_{1,j}}{a_{1,2}} \alpha_j(x) \right) + \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x). \end{aligned}$$

Posons alors

$$\lambda_1 = \alpha_1 + \sum_{j>2} \frac{a_{1,j}}{a_{1,2}} \alpha_j$$

et

$$\lambda_2 = \alpha_2 + \sum_{j>2} \frac{a_{2,j}}{a_{1,2}} \alpha_j.$$

On obtient :

$$\begin{aligned} & q(x) = a_{1,2} \lambda_1(x) \lambda_2(x) - \dots \\ & \dots - a_{1,2} \left(\sum_{j>2} \frac{a_{2,j}}{a_{1,2}} \alpha_j(x) \right) \left(\sum_{j>2} \frac{a_{1,j}}{a_{1,2}} \alpha_j(x) \right) + \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x), \end{aligned}$$

soit

$$\begin{aligned} & q(x) = \frac{a_{1,2}}{4} \left((\lambda_1(x) + \lambda_2(x))^2 - (\lambda_1(x) - \lambda_2(x))^2 \right) - \dots \\ & \dots - a_{1,2} \left(\sum_{j>2} \frac{a_{2,j}}{a_{1,2}} \alpha_j(x) \right) \left(\sum_{j>2} \frac{a_{1,j}}{a_{1,2}} \alpha_j(x) \right) + \sum_{i,j} a_{i,j} \alpha_i(x) \alpha_j(x). \end{aligned}$$

La famille $\lambda_1 + \lambda_2, \lambda_1 - \lambda_2, \alpha_3, \dots, \alpha_n$ constitue encore une base de l'espace dual. Le déterminant de la matrice de passage est égal à 2.

On peut alors dans un cas comme dans l'autre appliquer l'hypothèse de récurrence au terme de droite du membre de droite des équations. On admet donc qu'une forme qui s'écrit comme polynôme homogène de degré 2 en $n-1$ formes linéaires peut s'écrire comme combinaison linéaire de carrés. Or on vient bien de se réduire à une forme qui s'exprime comme polynôme homogène de degré 2 en au plus $n-1$ formes linéaires.

Le résultat peut être précisé. Dans la construction ci-dessus on substitue pas à pas à la base initiale de l'espace dual une nouvelle base. On peut donc supposer que les formes linéaires λ_i du théorème constituent un système libre. ■

Le théorème s'étend évidemment aux formes quadratiques hermitiennes. Par exactement le même type d'arguments on montre qu'elles peuvent se mettre sous la forme :

$$q(x) = \sum_i a_i |\lambda_i(x)|^2,$$

les coefficients a_i étant réels.

Nous allons maintenant préciser l'étude dans le cas réel.

Théorème 2. Soit E un espace vectoriel réel de dimension finie et soit q une forme quadratique. On peut trouver une famille de formes linéaires $\lambda_1, \dots, \lambda_r$ et une famille de formes linéaires $\mu_1, \dots, \mu_s$ telles que pour tout $x \in E$ on ait :

$$q(x) = \sum_{i=1, \dots, r} \lambda_i(x)^2 - \sum_{j=1, \dots, s} \mu_j(x)^2.$$

Si la famille $\{\lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_s\}$ est libre, l'entier $r + s$ est le rang de la forme.

Sous l'hypothèse précédente les entiers r et s sont eux aussi uniquement déterminés.

La dernière propriété est connue sous le nom de loi d'inertie de Sylvester. Le couple (r, s) est la signature de la forme quadratique.

Démonstration. Pour ce qui est de la première partie on commence, suivant le théorème précédent par écrire $q(x)$ sous la forme $\sum_i a_i \lambda_i(x)^2$. On suppose tous les a_i non nuls et on suppose que les λ_i forment un système libre. On regroupe suivant le signe des a_i . Supposons que les r premiers soient positifs et les s derniers soient négatifs. On peut alors réécrire $q(x)$ sous la forme :

$$\sum_{1 \leq i \leq r} (\sqrt{a_i} \lambda_i(x))^2 - \sum_{r+1 \leq i \leq r+s} (\sqrt{-a_i} \lambda_i(x))^2.$$

Montrons l'unicité de r et s . La somme $r + s$ est ainsi qu'on l'a vu bien déterminée, c'est le rang de la forme. Notons n la dimension de l'espace E . Soit donc deux décompositions :

$$q(x) = \sum_{i=1, \dots, r} \lambda_i(x)^2 - \sum_{j=1, \dots, s} \mu_j(x)^2.$$

et

$$q(x) = \sum_{i=1, \dots, r'} \lambda'_i(x)^2 - \sum_{j=1, \dots, s'} \mu'_j(x)^2.$$

Complétons la famille constituée par les λ_i et par les μ_j en une base de E^* , notons $\varepsilon_1, \dots, \varepsilon_t$ les formes linéaires adjointes. On a $r + s + t = n$.

Faisons de même pour la famille constituée par les λ'_i et les μ'_j . Soient $\varepsilon'_1, \dots, \varepsilon'_t$ les formes linéaires adjointes. On a $r' + s' + t = n$.

Supposons alors que $r < r'$. Considérons le sous-espace F de E d'équations $\lambda_i = 0$, $i = 1, \dots, r$. Il est de dimension $n - r$. La forme quadratique y prend des valeurs négatives ou nulles. Considérons le sous-espace F' d'équations $\mu'_j = 0$, $j = 1, \dots, s'$ et $\varepsilon'_\ell = 0$, $\ell = 1, \dots, t$. Il est de dimension $n - s' - t = r'$. La forme quadratique y

prend des valeurs positives non nulles en dehors de 0. Comme :

$$(n - r) + r' = n + (r' - r) > n$$

l'intersection de F et F' est de dimension au moins 1 et contient donc un vecteur non nul sur lequel q devrait être négative ou nulle, ou strictement positive. C'est impossible. La condition $r < r'$ menant à une contradiction on en déduit que $r = r'$ (la condition $r > r'$ est éliminée de façon analogue) et donc $s = s'$. ■

Pour conclure rappelons qu'une forme quadratique sur un espace vectoriel réel E est appelée positive si elle ne prend que des valeurs positives ou nulles. Elle est définie positive si elle est positive et ne prend la valeur 0 que pour le vecteur nul. Si l'espace E est de dimension n elle est de signature $(n, 0)$. On définit de même les formes négatives et définies négatives.

Enfin, ajoutons que ce que l'on vient de faire s'étend aux formes quadratiques hermitiennes :

Théorème 3. Soit E un espace vectoriel complexe de dimension finie et soit h une forme quadratique hermitienne. On peut trouver une famille de formes linéaires $\lambda_1, \dots, \lambda_r$ et une famille de formes linéaires $\mu_1, \dots, \mu_s$ telles que pour tout $x \in E$ on ait :

$$q(x) = \sum_{i=1, \dots, r} |\lambda_i(x)|^2 - \sum_{j=1, \dots, s} |\mu_j(x)|^2.$$

Si la famille $\{\lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_s\}$ est libre, l'entier $r + s$ est le rang de la forme.

Sous l'hypothèse précédente les entiers r et s sont eux aussi uniquement déterminés.

On définit également des formes définies positives et négatives.

Vecteurs isotropes

Définition 7

Soit q une forme quadratique sur un espace vectoriel E . On dit qu'un vecteur v est un vecteur isotrope si et seulement si $q(v) = 0$. On dit qu'un sous-espace F est totalement isotrope si tout $v \in F$ est isotrope.

Voici un exemple important. Considérons l'espace vectoriel k^2 et sa base canonique $\{e_1, e_2\}$. Considérons la forme quadratique déterminée par la matrice

$$J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Les vecteurs e_1 et e_2 sont isotropes. L'espace k^2 muni de cette forme est appelé le plan hyperbolique. Plus généralement on appelle plan hyperbolique tout espace

de dimension 2 muni d'une forme quadratique q pour laquelle on peut trouver une base où sa matrice soit J_2 .

Par exemple, c'est le cas de k^2 muni de la forme quadratique $x^2 - y^2$.

Définition 8

Soit E un espace vectoriel sur un corps k , et soit q une forme quadratique sur E . On appelle sous-espace totalement isotrope maximal un sous-espace totalement isotrope F tel que si F' est un sous-espace totalement isotrope contenant F , alors $F = F'$.

Tout sous-espace totalement isotrope est contenu dans un sous-espace totalement isotrope maximal car E est de dimension finie.

Si on considère le cas d'une forme quadratique, sur un espace vectoriel réel de dimension n , de signature (r, s) on montre que les sous-espaces totalement isotropes maximaux sont de dimension $n - (r + s) + \max(r, s)$. En fait on a :

Théorème 4. *Soit E un espace vectoriel de dimension finie sur un corps k , et soit q une forme quadratique sur E . Tous les sous-espaces totalement isotropes maximaux ont même dimension.*

Ce résultat est un corollaire du théorème de Witt qui suit et qui sera démontré en exercice :

Théorème 5 (Witt). *Soit E un espace vectoriel de dimension finie sur un corps k de caractéristique différente de 2, et soit q une forme quadratique sur E . Soient F et F' des sous-espaces vectoriels de E . Alors les deux conditions suivantes sont équivalentes :*

- *il existe un élément u du groupe $O(q)$ tel que $u(F) = F'$,*
- *il existe une application linéaire bijective f de F sur F' telle que pour tout $v \in F$ on ait $q(v) = q(f(v))$.*

2. Structure du groupe orthogonal euclidien

Dans cette section nous allons étudier plus en détails le cas où on a un espace vectoriel réel de dimension finie muni d'une forme quadratique définie positive q . Un tel espace est appelé un espace vectoriel euclidien. La forme polaire associée b est le produit scalaire euclidien.

L'application $x \mapsto \|x\| = \sqrt{q(x)}$ est alors une norme, rappelons que l'on a l'inégalité de Cauchy-Schwarz [LM], [LA].

Notons qu'une application qui préserve la norme préserve aussi le produit scalaire et est linéaire.

Une base $\{e_1, \dots, e_n\}$ de l'espace telle que $q(e_i) = 1$ et $b(e_i, e_j) = 0$ si $i \neq j$ est dite orthonormée.

Le théorème 2 de la section 1 montre que l'on peut trouver une base de E dans laquelle la forme q s'écrit $x_1^2 + \dots + x_n^2$. On en déduit que si q et q' sont deux formes définies positives sur un même espace vectoriel de dimension n les groupes $O(q)$ et $O(q')$ sont isomorphes. En effet, ils sont isomorphes tous les deux au groupe de la forme $x_1^2 + \dots + x_n^2$ qui est noté $O(n)$.

Une matrice de $O(n)$ est dite orthogonale.

Proposition 1. *Une matrice (n, n) A est orthogonale si et seulement si ses vecteurs forment une base orthonormée de $\mathbb{R}^n$, ou encore si et seulement si ${}^tAA = I_n$.*

Démonstration. Soient X et Y deux vecteurs colonnes. Leur produit scalaire est égal au produit ${}^tXY$. Pour qu'une matrice A soit orthogonale il faut et suffit que pour tous X, Y on ait ${}^tX({}^tAA)Y = {}^tXY$. Ceci est équivalent à l'égalité ${}^tAA = I_n$. Cette identité exprime exactement le fait que les vecteurs colonnes de A constituent une base orthonormée de $\mathbb{R}^n$. En effet, si $A = (a_{i,j})$ cette identité équivaut aux relations $\sum_i a_{i,j}^2 = 1$, pour $j = 1, \dots, n$, et $\sum_i a_{i,k}a_{i,\ell} = 0$ si $k \neq \ell$, relations qui expriment que le système des vecteurs lignes est orthonormé. □

Corollaire. *Une matrice orthogonale est de déterminant 1 ou -1 .*

Démonstration. En effet, on a $\det({}^tAA) = \det({}^tA)\det(A) = \det(A)^2 = \det(I_n) = 1$. D'où le résultat. □

■ Définition 1

On appelle groupe spécial orthogonal et on note $SO(n)$ le sous-groupe constitué par les matrices orthogonales de déterminant 1.

En particulier le groupe $SO(2)$ est constitué par les matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telles que :

- $a^2 + c^2 = b^2 + d^2 = 1$,
- $ab + cd = 0$,
- $ad - bc = 1$.

De la première équation on conclut qu'il existe θ tel que $a = \cos \theta$ et $c = \sin \theta$. De la seconde et de la première on déduit que $b = -\sin \theta$ et $d = \cos \theta$ ou $b = \sin \theta$ et $c = -\cos \theta$. Enfin de la troisième on déduit que la première solution est la bonne.

La matrice est donc égale à la matrice R_θ suivante

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Autrement dit que c'est une rotation d'angle θ . Notons que $R_\theta R_{\theta'} = R_{\theta+\theta'}$.

Notons enfin que l'application $i : \mathbf{U} \rightarrow \mathbf{SO}(2)$ qui à $e^{i\theta}$ associe R_θ est un isomorphisme de groupes. En effet, c'est un homomorphisme car on a

$$i(e^{i\theta} e^{i\theta'}) = e^{i\theta+\theta'} = R_{\theta+\theta'} = R_\theta R_{\theta'}.$$

C'est un isomorphisme à cause de la formule de Moivre, en fait cet énoncé n'est rien d'autre qu'une présentation pédante de cette formule !

Les autres éléments du groupe $\mathbf{O}(2)$ sont de déterminants -1 . L'analyse faite ci-dessus montre que la matrice d'un tel élément est de la forme :

$$\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix},$$

pour un certain réel θ . Il s'agit de la symétrie orthogonale autour de la droite faisant un angle $\frac{\theta}{2}$ avec l'axe des abscisses.

Dans la section suivante on reviendra plus en détails sur le cas de $\mathbf{SO}(3)$ et de $\mathbf{SO}(4)$.

Symétries et générateurs du groupe orthogonal

Soit E un espace vectoriel euclidien et b son produit scalaire.

Définition 2

Soit F un sous-espace vectoriel de E . Son orthogonal est par définition l'ensemble des vecteurs $v \in E$ tels que $b(x, v) = 0$ pour tout $x \in F$. On le note $F^\perp$.

C'est évidemment un sous-espace vectoriel de E .

Proposition 2. *L'espace vectoriel E est la somme directe de F et de $F^\perp$.*

Démonstration. En fait il faut commencer par montrer que $F^\perp$ est un sous-espace. Mais si on a $v, w \in F^\perp$ et $\lambda \in \mathbb{R}$, on a $b(x, v+w) = b(x, v) + b(x, w) = 0$ pour tout $x \in F$, et on a $b(x, \lambda v) = \lambda b(x, v) = 0$. Donc $v+w$ et λv sont éléments de $F^\perp$.

Montrons maintenant que E est somme directe de F et $F^\perp$. Si on a $v \in F \cap F^\perp$ cela implique que $b(v, v) = 0$. Mais comme b est un produit scalaire cela implique $v = 0$.

Pour démontrer que E est somme directe de F et $F^\perp$ il suffit maintenant de montrer que

$$\dim_{\mathbb{R}}(F) + \dim_{\mathbb{R}}(F^\perp) \geq \dim_{\mathbb{R}}(E).$$

Mais si d est la dimension de F et si $e_1, \dots, e_d$ est une base de F , le sous-espace $F^\perp$ est l'ensemble des x annulant les d formes linéaires $x \mapsto b(e_i, x)$. Il est donc de dimension au moins $n - d$. Le résultat suit. ■

On dit que F et $F^\perp$ forment une décomposition de E en somme directe orthogonale, et on note $E \cong F \perp F^\perp$. On notera que l'on a $(F^\perp)^\perp = F$. Le lemme élémentaire suivant est très utile :

Lemme 1. Soit E un espace vectoriel euclidien, soit f une application orthogonale, et soit F un sous-espace tel que $f(F) \subset F$. Alors on a $f(F^\perp) \subset F^\perp$.

Démonstration. L'application f restreinte à F induit un isomorphisme de F dans lui-même, étant injective et prenant ses valeurs par hypothèse dans F . Soit $v \in F^\perp$, on a donc pour tout $x \in F$ la relation $b(f(v), x) = b(v, f^{-1}(x)) = 0$, comme $f^{-1}(x)$ décrit tout F on a $f(v) \in F^\perp$. ■

Définition 3

Soit E un espace vectoriel euclidien, et soit $E \cong H \perp G$ une décomposition en somme directe orthogonale de E (i.e. G est égal à $H^\perp$). Soit $x \in E$, et soit $x = x_H + x_G$ sa décomposition, $x_H \in H$ et $x_G \in G$.

On appelle projection orthogonale de E sur H l'application linéaire qui à x associe x_H .

On appelle symétrie orthogonale autour de H , et on note s_H , l'application linéaire qui à x associe $x_H - x_G$.

Les symétries sont des éléments du groupe orthogonal. Il suffit de calculer

$$b(x_H + x_G, y_H + y_G) = b(x_H, y_H) + b(x_G, y_G) = b(x_H - x_G, y_H - y_G).$$

La conjuguée par un élément $r \in O(n)$ d'une symétrie s_H est la symétrie $s_{r(H)}$. En effet, soit $s' = r \circ s_H \circ r^{-1}$ et soit $v = r(x) \in r(H)$. Alors

$$s'(v) = r \circ s_H \circ r^{-1}(r(x)) = r \circ s_H(x) = r(x) = v.$$

Par ailleurs si w est dans l'orthogonal de $r(H)$ alors $r^{-1}(w)$ est dans l'orthogonal de H , donc $s(r^{-1}(w)) = -r^{-1}(w)$ et $s'(w) = -w$. Donc on a bien :

$$r \circ s_H \circ r^{-1} = s_{r(H)}.$$

Inversement on a :

Proposition 3. Soit E un espace vectoriel euclidien, et soient H et H' deux sous-espaces de même dimension. Les symétries s_H et $s_{H'}$ sont conjuguées.

Démonstration. Il suffit de prendre comme facteur de conjugaison une transformation orthogonale quelconque envoyant H sur H' . Une telle transformation existe. On choisit une base orthogonale de F , on la complète en une base orthogonale de E . Pour ce faire on peut utiliser le procédé de Gram-Schmidt [LA], [LM]. On fait la même chose pour F' , puis on considère la transformation qui envoie la première base de E sur la seconde, et celle de F sur celle de F' . ■

Définition 4

On appelle une symétrie orthogonale autour d'un hyperplan H , H est appelé l'hyperplan de réflexion. On appelle une symétrie orthogonale autour d'un sous-espace de dimension $n - 2$.

Proposition 4. Les retournements sont dans le groupe spécial orthogonal. Les réflexions sont de déterminants -1 .

Démonstration. Plus généralement considérons une symétrie autour d'un sous-espace H . Choisissons une base de E qui soit réunion d'une base de H et d'une base de son orthogonal G . La matrice de la symétrie est alors diagonale avec des 1 et des -1 . Le nombre de -1 est la dimension du sous-espace G . Le déterminant est donc $(-1)^{\dim(G)}$. Le résultat suit. ■

Théorème 1. Le groupe orthogonal est engendré par les réflexions. En fait, un élément du groupe orthogonal est produit d'au plus n réflexions.

Démonstration. Soit $f \in O(q)$. Supposons que le sous-espace propre H de f associé à la valeur propre 1 soit de dimension k . On va montrer qu'en multipliant f par une réflexion r on peut supposer que le sous-espace propre associé à la valeur propre 1 de $r \circ f$ est de dimension au moins $k + 1$.

Supposons que $k < n$, l'application f étant distincte de l'identité il existe donc un vecteur $v \in H^\perp$ tel que $v \neq f(v)$, notons que $f(v) \in H^\perp$. Considérons alors la réflexion r autour de l'hyperplan orthogonal à $v - f(v)$. On a $r(v) = f(v)$, et donc $r \circ f$ laisse v fixe.

L'application r laisse fixe point par point l'hyperplan orthogonal à $v - f(v)$, donc H qui est contenu dedans. Comme H est fixe point par point par l'application $r \circ f$, elle laisse fixe point par point le sous-espace vectoriel $H \perp \mathbb{R}v$ qui est de dimension $k + 1$.

En itérant au plus n fois cette construction on obtient un produit $r_h \circ \cdots \circ r_1 \circ f$, $h \leq n$, qui est égal à l'identité. Ceci donne le résultat. $\square$

Nous passons maintenant à l'étude du groupe dérivé.

Théorème 2. *Le groupe dérivé de $O(q)$ est $SO(q)$ si la dimension de E est supérieure ou égale à 2. Le groupe dérivé de $SO(q)$ est $SO(q)$ si la dimension de E est supérieure ou égale à 3.*

Démonstration. Un commutateur est de déterminant 1, donc le groupe dérivé est sous-groupe du groupe spécial orthogonal.

Traitions d'abord le cas $n = 2$. Rappelons que toute rotation α du plan est produit de deux réflexions r et r' . Par ailleurs deux réflexions quelconques sont conjuguées par l'unique rotation ρ qui envoie la droite de réflexion de la seconde sur celle de la première : $r' = \rho^{-1} \circ r \circ \rho$. Donc on a $\alpha = r \circ r' = r \circ \rho^{-1} \circ r^{-1} \rho$. C'est donc un commutateur car $r^{-1} = r$.

Pour ce qui est du groupe $SO(2)$, il est commutatif, son groupe dérivé est donc trivial.

Passons au cas général. Soit t un retournement, il est produit de deux réflexions s_H et $s_{H'}$. En effet, si t est le retournement autour du sous-espace G , il suffit de prendre pour H et H' deux hyperplans distincts contenant G . Comme précédemment s_H et $s_{H'}$ sont conjugués par un élément du groupe spécial orthogonal envoyant H sur H' . On montre alors exactement comme précédemment que le retournement est un commutateur. Maintenant l'argument montre que le produit de deux réflexions est un commutateur dans le groupe orthogonal. Or un élément du groupe spécial orthogonal est produit d'un nombre pair de réflexions :

$$s = r_1 \circ r_2 \circ \cdots \circ r_{2k}$$

car son déterminant est égal à 1. Donc

$$s = (r_1 \circ r_2) \circ (r_3 \circ r_4) \circ \cdots \circ (r_{2k-1} \circ r_{2k})$$

et chaque parenthèse est un commutateur.

Ceci détermine le groupe dérivé du groupe orthogonal. Dans le cas du groupe spécial orthogonal commençons par étudier le cas $n = 3$. Comme deux retournements quelconques sont conjugués et que les retournements engendrent le groupe spécial orthogonal il suffit de montrer qu'un retournement est un commutateur. On renvoie à [Pe]. $\square$

Pour terminer cette sous-section nous allons donner une forme « diagonale » des matrices orthogonales.

Théorème 3. Soit E un espace vectoriel euclidien de forme quadratique q . Soit $f \in \mathcal{O}(q)$. Il existe une base orthonormée de E dans laquelle f a une matrice diagonale par blocs de la forme :

$$\begin{pmatrix} C_1 & & & & & \\ & C_k & & & & \\ & & -1 & & & \\ & & & \ddots & & \\ & & & & -1 & \\ 0 & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}.$$

Dans cette matrice les sous-matrices C_i sont des matrices $(2, 2)$ spéciales orthogonales, c'est-à-dire que $C_i = \begin{pmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{pmatrix}$. On suppose de plus que les θ_i ne sont pas multiples entiers de π .

La traduction matricielle de cet énoncé est la suivante. Soit $A \in \mathcal{O}(n)$, alors il existe une matrice orthogonale P telle que la matrice $P^{-1}AP$ soit diagonale par blocs de la forme décrite plus haut.

Démonstration. On va raisonner par récurrence sur la dimension n de E .

Soit P le polynôme minimal de f . Toutes les valeurs propres de f sont de module 1. Le polynôme minimal est produit de facteurs $X - 1$, $X + 1$, et de facteurs P_i de degré 2 ayant des racines complexes conjugués :

$$P = (X + 1)^k (X - 1)^\ell \prod_i P_i^{\alpha_i}.$$

Distinguons alors deux cas.

Dans le premier cas on suppose que $k > 0$ ou que $\ell > 0$. On choisit alors un vecteur propre, de norme 1, v associé à la valeur propre 1 ou à la valeur propre -1 suivant le cas. Soit H le sous-espace vectoriel orthogonal à v . L'application f laisse invariant globalement H . La restriction de f à H est encore orthogonale. Dans une base orthonormée constituée de v et d'une base orthonormale de H la matrice de f est diagonale par blocs de la forme

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & K \end{pmatrix}$$

où K est orthogonale de taille $(n - 1, n - 1)$. On peut alors appliquer l'hypothèse de récurrence et trouver une base de H où la restriction de f à H a une matrice

du type décrit dans l'énoncé. La matrice de f sera alors, à une permutation près des vecteurs de base, du type cherché.

Dans le second cas on a $k = \ell = 0$. Choisissons alors un vecteur non nul v de norme 1 et tel que $P_1(f)(v) = 0$. Il en existe nécessairement un par définition du polynôme minimal. Le système $\{v, f(v)\}$ est libre, en effet si ce n'était pas le cas v serait vecteur propre (associé à la valeur propre 1 ou -1) ce qui est exclu par hypothèse.

Le sous-espace L engendré par v et $f(v)$ est stable par f . Il suffit de montrer qu'un vecteur de la forme $f(\alpha v + \beta f(v))$, $\alpha, \beta \in \mathbb{R}$, est de la forme $\lambda v + \mu f(v)$, $\lambda, \mu \in \mathbb{R}$. Mais ceci résulte du choix de v , en effet si on note $P_1 = X^2 + bX + c$ avec $b, c \in \mathbb{R}$ on a $(f^2 + f + \text{Id})(v) = 0$ soit $f^2(v) = -bf(v) - v$. Ce qui donne le résultat.

Mais f , laissant stable L , laisse stable son orthogonal $L^\perp$. Dans une base orthonormée constituée, d'une base orthonormée de L , et d'une base orthonormée de $L^\perp$, f a une matrice diagonale par blocs de la forme

$$\begin{pmatrix} C & \\ & K \end{pmatrix}$$

où K est orthogonale de taille $(n-2, n-2)$. On conclut alors en utilisant l'hypothèse de récurrence comme plus haut.

Notons que si l'application f (ou la matrice A) est de déterminant 1 la matrice du théorème a forcément un nombre pair de termes -1 sur la diagonale. Mais une matrice de la forme $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ est évidemment de la forme $C_i = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$ avec $\theta = \pi$. Donc en regroupant les termes diagonaux -1 deux par deux la matrice peut se mettre sous la forme

$$\begin{pmatrix} C_1 & & & & \\ & \ddots & & & \\ & & C_k & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix},$$

en supprimant la restriction sur les θ_i . ■

Ces résultats appliqués en dimension 3 restituent la description classique des éléments de $\text{SO}(3)$:

- un élément de $\text{SO}(3)$ a toujours 1 pour valeur propre et est une rotation autour de l'axe parallèle à un vecteur propre associé.

Notons qu'ils impliquent aussi que toute valeur propre d'une matrice orthogonale est de module 1.

Topologie des groupes orthogonaux

Nous étudions ici quelques propriétés topologiques des groupes orthogonaux.

Proposition 5. $O(n)$ est compact.

Démonstration. Munissons l'espace vectoriel $M_n(\mathbb{R})$ de la norme définie par $\|A\| = \sqrt{\sum a_{i,j}^2}$, pour $A = (a_{i,j})$.

Observons d'abord que $O(n)$ est borné. En effet, si $A \in O(n)$ on a $\sum_i a_{i,j}^2 = 1$, pour $j = 1, \dots, n$, et donc $\|A\| = n^2$.

Par ailleurs $O(n)$ est fermé. La caractérisation des matrices orthogonales nous dit que c'est l'image inverse du point $(\underbrace{1, \dots, 1}_{n \text{ fois}}, \underbrace{0, \dots, 0}_{n(n-1)/2 \text{ fois}})$ par l'application continue

de $M_n(\mathbb{R})$ dans $\mathbb{R}^n \times \mathbb{R}^{\frac{n(n-1)}{2}}$ donnée par

$$A \mapsto \left(\left(\sum_i a_{i,j}^2, \quad j = 1, \dots, n \right), \left(\sum_i a_{i,k} a_{i,\ell} = 0, \quad k \neq \ell \right) \right).$$

Proposition 6. Le groupe $SO(n)$ est connexe par arcs.

Démonstration. Il suffit de relier une matrice $A \in SO(n)$ quelconque à la matrice identité par un arc continu. Pour cela effectuons un changement de base qui mette sous la forme décrite à la fin de la sous-section précédente. Soit P la matrice de changement de base. La matrice $P^{-1}AP$ est de la forme :

$$\begin{pmatrix} C_1 & & & & \\ & \ddots & & & \\ & & C_k & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix},$$

avec

$$C_i = \begin{pmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{pmatrix}.$$

Introduisons alors les matrices :

$$C_i(t) = \begin{pmatrix} \cos t\theta_i & \sin t\theta_i \\ -\sin t\theta_i & \cos t\theta_i \end{pmatrix}.$$

Puis la matrice

$$\begin{pmatrix} C_1(t) & & & \\ & C_k(t) & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}.$$

Pour $t = 0$ cette matrice est égale à la matrice identité et pour $t = 1$ elle est égale à la matrice $P^{-1}AP$. Ceci nous fournit un arc continu joignant ces deux matrices. En multipliant par P et P^{-1} on a un arc continu joignant la matrice identité à A . ■

On observera que $\mathrm{SO}(n)$ est donc la composante connexe de l'élément neutre de $\mathrm{O}(n)$.

3. Les quaternions, les groupes $\mathrm{SO}(3)$ et $\mathrm{SO}(4)$

Cette section est consacrée à une étude particulière des groupes $\mathrm{SO}(3)$ et $\mathrm{SO}(4)$ en liaison avec les quaternions que nous commençons par définir.

On considère l'espace vectoriel $\mathbb{C}^2$ et on note les vecteurs de la base canonique $\{e_1, e_2\}$ par 1 et j , 1 est donc le vecteur $(1, 0)$ et j le vecteur $(0, 1)$. On y définit un produit par la formule suivante :

$$(a_0 + b_0 j)(a_1 + b_1 j) = a_0 a_1 - b_0 \bar{b}_1 + (a_0 b_1 + \bar{a}_1 b_0)j.$$

Dans cette formule a_0, b_0, a_1, b_1 sont des nombres complexes.

Théorème 1. *L'espace vectoriel $\mathbb{C}^2$ muni de ce produit est un corps non commutatif. On le note $\mathbf{H}$ et on l'appelle le corps des quaternions.*

Démonstration. La vérification de la distributivité ne pose pas de problèmes. La vérification de l'associativité de la multiplication est un peu fastidieuse mais sans problème. Voici les calculs :

$$\begin{aligned} ((a_0 + b_0 j)(a_1 + b_1 j))(a_2 + b_2 j) &= (a_0 a_1 - b_0 \bar{b}_1 + (a_0 b_1 + \bar{a}_1 b_0)j)(a_2 + b_2 j) = \cdots \\ \cdots &= a_0 a_1 a_2 - b_0 \bar{b}_1 a_2 - a_0 b_1 \bar{b}_2 - \bar{a}_1 b_0 \bar{b}_2 + (a_0 a_1 b_2 - b_0 \bar{b}_1 b_2 + a_0 \bar{a}_2 b_1 - \bar{a}_1 \bar{a}_2 b_0)j. \end{aligned}$$

D'un autre côté on trouve aussi :

$$\begin{aligned} (a_0 + b_0 j)((a_1 + b_1 j)(a_2 + b_2 j)) &= (a_0 + b_0 j)(a_1 a_2 - b_1 \bar{b}_2 + (a_1 b_2 + \bar{a}_2 b_1)j) = \cdots \\ \cdots &= a_0 a_1 a_2 - b_0 \bar{b}_1 a_2 - a_0 b_1 \bar{b}_2 - \bar{a}_1 b_0 \bar{b}_2 + (a_0 a_1 b_2 - b_0 \bar{b}_1 b_2 + a_0 \bar{a}_2 b_1 - \bar{a}_1 \bar{a}_2 b_0)j. \end{aligned}$$

Ceci démontre l'associativité.

Il reste à démontrer que tout élément admet un inverse à gauche et à droite. Pour ce faire on va commencer par introduire quelques notations puis un lemme. On note k l'élément ij . Si bien que tout élément peut s'écrire d'une manière unique soit $a + bj$ avec $a, b \in \mathbb{C}$, soit $\alpha + \beta i + \gamma j + \delta k$ avec $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ et, $a = \alpha + \beta i$ et $b = \gamma + \delta i$. À partir de la formule donnée plus haut on établit les relations :

- $i^2 = j^2 = k^2 = -1$,
- $ij = k, jk = i, ki = j$,
- $ij = -ji, jk = -kj, ki = -ik$.

Le sous-ensemble des éléments tels que $\beta = \gamma = \delta = 0$ est un sous-corps isomorphe à $\mathbb{R}$, par abus de langage on l'identifiera à $\mathbb{R}$. De même le sous-ensemble des éléments tels que $\gamma = \delta = 0$ est un sous-corps isomorphe à $\mathbb{C}$. On déduit de la description précédente que :

Proposition 1. *Le centre de $\mathbf{H}$, c'est-à-dire l'ensemble des éléments de $\mathbf{H}$ commutant à tous les éléments de $\mathbf{H}$ est égal à $\mathbb{R}$.*

Démonstration. En effet un élément $\alpha + \beta i + \gamma j + \delta k \in \mathbf{H}$ commute à tous les éléments de $\mathbf{H}$ si et seulement si il commute à i, j et k . Ceci a lieu si et seulement si $\beta = \gamma = \delta = 0$. □

L'espace $\mathbb{R}^4$ est muni du produit scalaire standard que l'on notera $(x|y)$. La norme euclidienne est notée $\|z\|$.

Définissons alors le conjugué $\bar{z}$ d'un élément $z \in \mathbf{H}$, avec $z = a + bj$ comme étant $\bar{a} - bj$, ou si on écrit $z = \alpha + \beta i + \gamma j + \delta k$ comme étant $\alpha - \beta i - \gamma j - \delta k$.

Lemme 1. *Avec les notations précédentes on a :*

$$z\bar{z} = \alpha^2 + \beta^2 + \gamma^2 + \delta^2.$$

Démonstration. Elle se fait par calcul direct. Il est plus facile d'écrire $z = a + bj$ avec $a = \alpha + \beta i$ et $b = \gamma + \delta j$. Puis de calculer

$$z\bar{z} = (a + bj)(\bar{a} - bj) = a\bar{a} + b\bar{b} = \alpha^2 + \beta^2 + \gamma^2 + \delta^2. \quad \square$$

Le produit $z\bar{z}$ est donc égal au carré de la norme euclidienne $\|z\|^2$.

On en déduit que l'inverse à gauche et à droite d'un élément z non nul est égal à $\frac{\bar{z}}{\|z\|^2}$. ■

Lemme 2. *Pour tous $y, z \in \mathbf{H}$ on a :*

$$\overline{(yz)} = \bar{z} \bar{y}.$$

■ Définition 1

On appellera quaternion pur un élément de $\mathbf{H}$ de la forme $\beta i + \gamma j + \delta k$.

Considérons maintenant un quaternion pur de norme 1. Il s'écrit donc $z = \beta i + b j$, on a donc $z^2 = -\beta^2 - b^2 + \beta b(ij + ji) = -1$. On observera donc que dans le corps non commutatif $\mathbf{H}$ le polynôme $X^2 + 1$ a une infinité de zéros.

Proposition 2. Soit $v \in \mathbf{H}$ tel que $\|v\| = 1$. Alors l'application R_v de $\mathbf{H}$ dans lui-même qui à un élément z associe $z\bar{v}$ est orthogonale. De plus elle laisse invariant le sous-espace des quaternions purs.

Démonstration. Il suffit de vérifier qu'elle préserve la norme euclidienne. Il suffit donc de vérifier que pour tout $z \in \mathbf{H}$ on a $z\bar{z} = (vz\bar{v})(\overline{vz\bar{v}})$. Or on a :

$$\begin{aligned}(vz\bar{v})(\overline{vz\bar{v}}) &= vz\bar{v}v\bar{z} = vz\|v\|^2\bar{z} = \dots \\ \dots &= vz\bar{z}\bar{v} = v\|z\|^2\bar{v} = \|z\|^2v\bar{v} = \|z\|^2\|v\|^2 = \|z\|^2.\end{aligned}$$

Pour ce qui est de la seconde partie on observe que l'axe réel, c'est-à-dire le sous-corps $\mathbb{R}$, est stable par R_v . En effet, un réel α commute à tout élément de $\mathbf{H}$ donc $v\alpha\bar{v} = \alpha\|v\|^2 = \alpha$. Donc R_v laisse stable l'orthogonal de l'axe réel qui est par définition l'ensemble des quaternions purs et isomorphe à l'espace euclidien $\mathbb{R}^3$. La restriction de R_v à ce sous-espace est donc un élément du groupe orthogonal $\mathbf{O}(3)$. ■

Notons S^3 l'ensemble des quaternions de norme 1. C'est un groupe pour la multiplication de $\mathbf{H}$ car si on a $\|u\| = \|v\| = 1$ on a $\|uv\| = \|u\|\|v\| = 1$.

Théorème 2. L'application $v \mapsto R_v$ est un homomorphisme de groupes de S^3 sur $\mathbf{SO}(3)$. Son noyau a deux éléments et est égal à $\{1, -1\}$.

Démonstration. L'application est un homomorphisme car

$$R_{uv}(x) = (uv)x(\overline{uv}) = u(vx\bar{v})\bar{u} = R_u((vx\bar{v})) = R_u(R_v(x)).$$

Le noyau de l'homomorphisme est constitué par les éléments de norme 1 tels que pour tout $x \in \mathbf{H}$ on ait $vx\bar{v} = x$, soit $vx = xv$. Donc les éléments de norme 1 qui commutent à tout élément de $\mathbf{H}$. C'est-à-dire à 1 et -1 .

L'application prend ses valeurs dans $\mathbf{SO}(3)$. En effet l'ensemble S^3 est un sous-espace connexe de $\mathbb{R}^4$. Ce point est laissé en exercice au lecteur. Alors l'application $v \mapsto R_v$ est une fonction polynomiale des coordonnées et est donc continue. Son image est donc connexe, et comme elle doit contenir l'élément neutre, c'est-à-dire la matrice I_3 , elle est contenue dans $\mathbf{SO}(3)$.

Il reste à montrer que l'application est surjective sur $\mathbf{SO}(3)$.

Soit q un quaternion pur de norme 1, et soit $q_\theta = \cos \theta + \sin \theta q$.

Lemme 3. Soit $q \in S^3$ un quaternion pur fixé. L'ensemble des quaternions q_θ , $0 \leq \theta < 2\pi$, est un sous-groupe de S^3 .

Démonstration. Il suffit d'observer que $q_\theta q_{\theta'} = q_{\theta+\theta'}$. On notera C_q ce sous-groupe. □

L'application R_{q_θ} laisse fixe la droite de direction q . Cette application est donc une rotation de $\mathbb{R}^3$ d'axe q .

Montrons enfin que l'application $q_\theta \mapsto R_{q_\theta}$ est surjective sur l'ensemble des rotations de $\mathbb{R}^3$ d'axe q . Ceci achèvera la démonstration du théorème.

Nous savons que R_{q_θ} est une rotation d'axe q . Pour calculer l'angle de la rotation, au signe près, il suffit de connaître la trace. En effet on sait que la trace d'une rotation de $\mathbb{R}^3$ d'angle a est $1 + 2 \cos a$. Soit donc $q = \beta i + \gamma j + \delta k$. Pour calculer sa trace il faut faire la somme du coefficient de i dans $q_\theta i \bar{q}_\theta$, de j dans $q_\theta j \bar{q}_\theta$ et de k dans $q_\theta k \bar{q}_\theta$. On trouve tous calculs faits : $\cos^2 \theta + \sin^2 \theta (\beta^2 - \gamma^2 - \delta^2)$, $\cos^2 \theta + \sin^2 \theta (-\beta^2 + \gamma^2 - \delta^2)$, $\cos^2 \theta + \sin^2 \theta (-\beta^2 - \gamma^2 + \delta^2)$. La somme est égale $3 \cos^2 \theta - \sin^2 \theta = 1 + 2 \cos 2\theta$.

L'image contient donc une rotation d'angle 2θ ou -2θ . Mais comme elles sont inverses l'une de l'autre et que l'image est un sous-groupe elle contient les deux. Donc elle contient toutes les rotations d'axe q .

Le résultat suit. ■

Soient maintenant $u, v \in S^3$. On considère l'application :

$$\mathbf{H} \longrightarrow \mathbf{H}, \quad x \longmapsto ux\bar{v}.$$

Proposition 3. Cette application est un homomorphisme de $S^3 \times S^3$ vers $\text{SO}(4)$. Cet homomorphisme est surjectif et de noyau le sous-groupe $\{(1, 1), (-1, -1)\}$.

Démonstration. On démontre comme plus haut que cette application est à valeurs dans $\text{O}(4)$ en montrant qu'elle préserve la norme euclidienne. La démonstration est identique à celle faite plus haut. De même on montre que c'est un homomorphisme de groupes. On montre que cette application est à valeurs dans $\text{SO}(4)$ en observant que son image est connexe, puisque sa source l'est, qu'elle est continue, et qu'elle contient la matrice identité, et enfin que $\text{SO}(4)$ est la composante connexe de la matrice identité dans le groupe orthogonal.

Il nous faut déterminer le noyau de cette application et montrer qu'elle est surjective.

Pour ce qui est du noyau on cherche u et v tels que $ux\bar{v} = x$ pour tout $x \in \mathbf{H}$. En faisant $x = v$ on obtient $u = v$. Puis l'argument du théorème précédent nous dit que $u = 1$ ou $u = -1$.

Pour ce qui est de la surjectivité on se ramène au cas précédent comme suit. Soit $f \in \text{SO}(4)$, posons $\bar{v} = f(1)$, $v \in S^3$. Alors l'application $x \mapsto vf(x)$ est orthogonale et fixe 1. Elle laisse donc fixe globalement l'orthogonal, c'est-à-dire l'ensemble des quaternions purs. Il existe donc $u \in S^3$ tel que pour tout x on ait $vf(x) = ux\bar{u}$. Il en résulte que pour tout $x \in \mathbf{H}$ on a $f(x) = \bar{v}ux\bar{u}$. ■

Simplicité de $\text{SO}(3)$

Cette section est consacrée à la simplicité du groupe $\text{SO}(3)$. Ainsi qu'on le verra le résultat correspondant n'est pas vrai pour $\text{SO}(4)$, cela résulte de ce que l'on a fait plus haut. Par contre pour $n \geq 5$ il y a un résultat correspondant qui sera énoncé aussitôt après, on renverra aux exercices pour la démonstration.

Théorème 3. *Le groupe $\text{SO}(3)$ n'a pas de sous-groupes distingués non-triviaux.*

Démonstration. Soit H un sous-groupe distingué non trivial. Il contient donc une matrice R de rotation d'angle θ . Considérons maintenant une matrice réelle antisymétrique $(3,3)$ quelconque soit A , on a donc ${}^tA = -A$.

Lemme 4. *Pour tout $u \in \mathbb{R}$ la matrice $\exp(uA)$ est orthogonale réelle.*

Démonstration du lemme. On renvoie aux exercices du chapitre précédent pour des informations sur l'exponentielle. Il nous suffit de vérifier le calcul :

$$\begin{aligned} {}^t(\exp(uA))(\exp(uA)) &= (\exp(u{}^tA))(\exp(uA)) = \dots \\ \dots &= (\exp(-uA))(\exp(uA)) = (\exp(-uA + uA)) = I_3. \end{aligned} \quad \square$$

Revenons à la démonstration du théorème.

Choisissons une matrice antisymétrique $(3,3)$ dont le noyau soit de dimension 1 (il est nécessairement de dimension au moins 1) et distinct de l'axe de rotation de R . Alors, si $u \neq 0$, la rotation $\exp(uA)$ ne commute pas avec R .

Pour tout $u \in \mathbb{R}$ l'élément $R\exp(uA)R^{-1}\exp(-uA)$ est dans H . Cet élément est une rotation distincte de l'identité à cause de la remarque précédente. Mais l'application :

$$u \longmapsto R\exp(uA)R^{-1}\exp(-uA)$$

est une fonction continue de u . Il en est de même de la fonction :

$$u \longmapsto \text{Tr}(R\exp(uA)R^{-1}\exp(-uA)).$$

Comme $R\exp(uA)R^{-1}\exp(-uA)$ tend vers I_3 quand u tend vers 0 la fonction précédente tend vers 3.

Rappelons que la trace d'une rotation d'angle θ dans $\mathbb{R}^3$ est égale à $1 + 2\cos\theta$. L'angle de la rotation considérée plus haut tend vers 0 quand u tend vers 0.

La trace ne détermine l'angle qu'au signe près mais si H contient un élément il contient aussi son inverse, qui est une rotation d'angle opposée. Plus précisément, la fonction, de $\mathbb{R}$ dans $\mathbb{R}$, qui à u associe la valeur absolue de l'angle de rotation de $R \exp(uA) R^{-1} \exp(-uA)$ est continue sur un voisinage de 0 dans $\mathbb{R}$, et tend vers 0 quand u tend vers 0. Le théorème des valeurs intermédiaires nous dit qu'elle prend toute valeur positive proche de zéro.

En conclusion nous venons de montrer que H contient des rotations d'angle assez petit arbitraire. Comme il est distingué, et que deux rotations d'angle donné sont conjugués, il contient toutes les rotations d'angle assez petit. Mais soit R est une rotation quelconque, d'axe D et d'angle θ , et soit n un entier positif non nul. La rotation r d'axe D et d'angle $\frac{\theta}{n}$ est dans H pour n assez grand, et on a $r^n = R$, donc $R \in H$. ■

Passons au groupe $\text{SO}(4)$, on a montré qu'il était quotient de $S^3 \times S^3$. Ce dernier groupe admet par exemple le groupe $S^3 \times 1$ comme sous-groupe distingué. Comme l'homomorphisme vers $\text{SO}(4)$ est surjectif l'image de ce sous-groupe est distinguée dans $\text{SO}(4)$. On remarquera qu'elle est isomorphe à S^3 .

Énonçons le cas général :

Théorème 4. *Si $n \geq 5$, un sous-groupe distingué de $\text{SO}(n)$ est, soit contenu dans le centre, soit égal au groupe tout entier.*

Notons que le centre de $\text{SO}(n)$ est constitué de l'ensemble des matrices orthogonales diagonales. Si n est impair il est réduit à la matrice I_n , si n est pair il est égal à $\{I_n, -I_n\}$.

On renvoie aux exercices pour les détails.

Sous-groupes finis de $\text{SO}(3)$

Dans cette section on détermine les sous-groupes finis de $\text{SO}(3)$ et on les relie aux polyèdres réguliers.

Soit $H \subset \text{SO}(3)$ un sous-groupe fini. Le groupe $\text{SO}(3)$ laisse invariante la sphère S^2 de rayon 1. Étant donné $g \in H$ soit $\text{Fix}(g)$ l'ensemble des points fixes de g appartenant à la sphère S^2 . Un élément distinct de l'identité dans H est une rotation. Il y a donc deux points fixes sur S^2 qui sont les points d'intersection de la sphère et de l'axe de rotation. Pour déterminer les sous-groupes finis H de $\text{SO}(3)$ on va étudier l'ensemble de ces points fixes quand g décrit $H - \{\text{Id}\}$.

À cette fin on définit un ensemble F par la formule suivante :

$$F = \{(g, s) \in H \times S^2 \mid g \neq \text{Id}, s \in \text{Fix}(g)\}.$$

On va calculer le cardinal de cet ensemble de deux manières différentes et en déduire des restrictions sur la structure de H .

Considérons d'abord l'application $p: \mathbf{F} \rightarrow H - \{\text{Id}\}$ qui à (g, s) associe g . On vient de voir que l'image inverse par p d'un g quelconque est constituée de deux points. Donc, on a

$$\#\mathbf{F} = 2(\#H - 1).$$

Considérons ensuite l'application $\pi: \mathbf{F} \rightarrow S^2$ qui à (g, s) associe s . Son image est l'ensemble $\mathcal{S}$ des points de S^2 qui sont point fixe d'au moins un élément de H , élément distinct de Id . Si $s \in \mathcal{S}$ son image inverse par π est l'ensemble des (g, s) où $g \in H_s$, où H_s est le sous-groupe des éléments de H laissant s fixe. Le sous-groupe H_s est un sous-groupe de rotations autour de l'axe passant par s et l'origine. L'image inverse $\pi^{-1}(s)$ a pour cardinal $\#H_s - 1$. Soit s' un point de l'orbite de s sous l'action H , le fixateur $H_{s'}$ est conjugué de H_s dans H . Le lemme qui suit est une spécialisation d'un résultat démontré de manière plus générale dans le chapitre 2 (Proposition 2, Section 2).

Lemme 5. *Le cardinal de $H_{s'}$, dans l'orbite de s sous l'action de H , ne dépend que de s . Cette valeur est l'ordre d'une rotation engendrant H_s*

Donc l'image inverse par π de s' a aussi $\#H_s - 1$ éléments. Il y a $\frac{\#H}{\#H_s}$ éléments dans l'orbite de s . Soit $\mathcal{R}$ un système de représentants pour ces orbites, on obtient :

$$\#(\mathbf{F}) = \sum_{s \in \mathcal{R}} \frac{\#H}{\#H_s} (\#H_s - 1).$$

Soit finalement :

$$2(\#H - 1) = \sum_{s \in \mathcal{R}} \frac{\#H}{\#H_s} (\#H_s - 1)$$

ou encore :

$$2 - \frac{2}{\#H} = \sum_{s \in \mathcal{R}} \left(1 - \frac{1}{\#H_s}\right).$$

Cette identité impose des bornes aux valeurs $\#\mathcal{R}$ et $\#H_s$.

Comme le groupe H_s est toujours par hypothèse de définition non trivial, la quantité $1 - \frac{1}{\#H_s}$ est supérieure ou égale à $\frac{1}{2}$, en comparant à l'équation ci-dessus on en déduit que $\#\mathcal{R} \leq 3$. Considérons donc les cas possibles, on notera H_i pour H_{s_i} dans la suite.

Cas 1 : $\#\mathcal{R} = 1$.

En considérant l'équation ci-dessus on constate que ce cas est impossible.

Cas 2 : $\#\mathcal{R} = 2$.

On a alors $2 - \frac{2}{\#H} = 2 - \frac{1}{\#H_1} - \frac{1}{\#H_2}$. On a soit $2\#H_i \leq \#H$ soit $\#H_i = \#H$, on en déduit facilement que $H_i = H$. Comme le groupe H laisse fixe un vecteur non-nul

les éléments de H sont des rotations autour de l'axe déterminé par ce vecteur. Le sous-groupe H est donc un groupe cyclique constitué de rotations autour d'un axe donné. Les points fixes sont les points d'intersection de la sphère avec l'axe de rotation, ils forment deux orbites.

Cas 3 : $\# \mathcal{R} = 3$.

On commence par observer que le groupe est d'ordre pair $2n$. La formule ci-dessus devient :

$$1 + \frac{2}{\#H} = \frac{1}{\#H_1} + \frac{1}{\#H_2} + \frac{1}{\#H_3}.$$

Considérant cette équation on observe que les trois entiers $\#H_1$, $\#H_2$, et $\#H_3$ ne peuvent être tous les trois supérieurs ou égaux à 3. L'un d'eux au moins est égal à 2. Il est clair aussi que les deux autres ne peuvent être "grands" simultanément. En fait ils ne peuvent tous les deux être supérieurs ou égaux à 4. À partir de ces remarques on déduit que le cas 3 se subdivise en quatre sous-cas que l'on décrit ci-dessous en donnant leur réalisation géométrique, détaillant en particulier les ensembles de points fixes.

Le premier sous-cas dépend d'un entier n :

Cas 3.1, les polygones réguliers, $\#H_1 = \#H_2 = 2$, et $2\#H_3 = \#H = 2n$.

Ce cas est celui du groupe diédral d'ordre $2n$ (Chapitre 1, Section 2). Le groupe est engendré par les rotations d'ordre n autour d'un axe donné D , et par la symétrie autour d'un axe de symétrie d'un polygone régulier à n côtés tracé dans le plan P orthogonal à D , centré à l'origine, et inscrit sur la sphère S^2 .

Les points fixes sont les sommets du polygone et les intersections des droites passant par l'origine et le milieu des côtés avec la sphère S^2 , ainsi que les points d'intersection de la sphère avec D . Les points d'intersection de la sphère avec D forment une orbite. La seconde est formée par les sommets d'un polygone régulier, la dernière par les intersections des droites passant par l'origine et le milieu des côtés avec S^2 .

Les sous-cas suivants ne dépendent plus d'un entier n :

Cas 3.2, le tétraèdre, $\#H_1 = 2$, et $\#H_2 = \#H_3 = 3$, et $\#H = 12$.

Le groupe des isométries directes laissant fixe un tétraèdre centré en l'origine et inscrit sur la sphère S^2 réalise ce cas. Ce groupe est isomorphe au groupe alterné $\mathcal{A}_4$. On renvoie à l'exercice 7 du chapitre 2. On montrera en exercice qu'un groupe qui réalise ce cas a nécessairement cette structure (voir aussi l'exercice 6 du Chapitre 2).

Les points fixes sont les sommets du tétraèdre, leurs points antipodaux sur la sphère, ainsi que les points d'intersection des droites joignant les milieux des côtés opposés du tétraèdre avec S^2 .

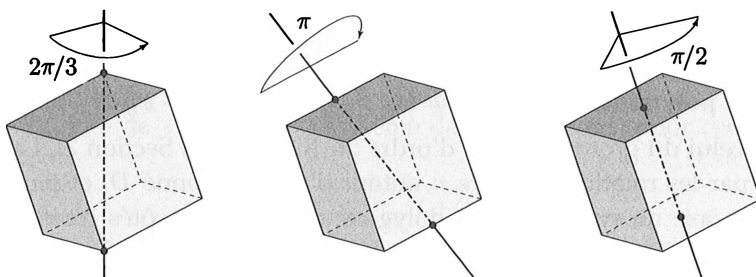
Les sommets et leurs antipodaux constituent deux orbites à 4 éléments. Les points d'intersection des droites joignant les milieux des côtés opposés avec S^2 en constituent une à 6 éléments.

Cas 3.3, le cube, $\#H_1 = 2$, et $\#H_2 = 3$, $\#H_4 = 4$, et $\#H = 24$.

Le groupe des isométries directes laissant fixe un cube centré en l'origine, inscrit sur S^2 , réalise ce cas. Ce groupe est isomorphe au groupe symétrique $\mathfrak{S}_4$ (voir exercice 8 Chapitre 2).

Les points fixes sont les sommets du cube, les points d'intersection des droites joignant les milieux des côtés opposés avec S^2 , ainsi que les points d'intersection des droites joignant les faces opposées du cube avec S^2 .

Les sommets constituent une orbite de points fixes à 8 éléments. Les points d'intersection des droites joignant les milieux des côtés opposés avec S^2 constituent une orbite à 12 éléments. Les intersections des droites joignant les faces opposées du cube avec S^2 en constituent une à 4 éléments. On renvoie à la figure pour les détails.



On renvoie à l'exercice 8 du chapitre 2. On montrera en exercice qu'un groupe qui réalise ce cas a nécessairement cette structure.

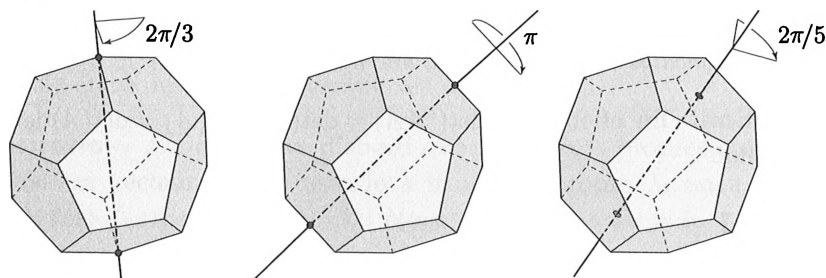
Cas 3.4, on a $\#H_1 = 2$, et $\#H_2 = 3$, $\#H_3 = 5$, et $\#H = 60$.

Ce cas est le plus difficile à réaliser. Nous ne donnerons pas les détails complets. Il est réalisé par le groupe laissant fixe un dodécaèdre (polyèdre régulier dont les faces sont des pentagones), ou un icosaèdre. Le dodécaèdre (voir figure) a 12 faces, 30 arêtes et 20 sommets. On renvoie au cours de géométrie de Berger pour tous les détails. On notera les points suivants. Les orbites de points fixes sont d'abord les sommets il y en a 20 ; les points d'intersections des droites passant par les centres de faces opposées avec la sphère il y en a 12 ; et enfin les points d'intersections des droites passant par les centres d'arêtes opposées avec la sphère il y en a 30.

Cette orbite permet de construire un isomorphisme du groupe avec le groupe $\mathcal{A}_5$. Si on considère les droites passant par les centres d'arêtes opposées on constate qu'elles se regroupent par 3 pour former des systèmes d'axes orthogonaux (on ne tient pas compte de l'ordre des axes). Il y a 5 tels systèmes d'axes. Les isométries

du dodécaèdre les échangent. On peut aussi observer qu'il y a 5 cubes inscrits dans le dodécaèdre qui sont échangés par les isométries.

Clairement le groupe du dodécaèdre contient des rotations d'ordre 3 et 5. Il contient aussi des éléments d'ordre 2, les rotations d'angle π autour des droites passant par les centres d'arêtes opposées. On en déduit facilement que l'homomorphisme est sur $\mathcal{A}_5$. L'injectivité est facile.



Il convient enfin de vérifier que tout groupe vérifiant les conditions initiales est de ce type. Ceci est laissé au lecteur.

4. Structure du groupe unitaire

Dans cette section nous allons étudier plus en détails le cas d'un espace vectoriel complexe de dimension finie muni d'une forme hermitienne définie positive h . Un tel espace est appelé un espace vectoriel hermitien. La forme polaire associée b est le produit hermitien.

L'application $x \mapsto \sqrt{h(x)}$ est alors une norme [LM], [LA]. Une base $\{e_1, \dots, e_n\}$ de l'espace telle que $q(e_i) = 1$ et $b(e_i, e_j) = 0$ si $i \neq j$ est dite orthonormée ou hermitienne.

Le théorème 3 de la section 1 montre que l'on peut trouver une base de E dans laquelle la forme h s'écrit $x_1 \bar{x}_1 + \dots + x_n \bar{x}_n$. On en déduit que si h et h' sont deux formes hermitiennes définies positives sur un même espace vectoriel complexe de dimension n , les groupes $U(q)$ et $U(q')$ sont isomorphes. En effet, ils sont isomorphes tous les deux au groupe de la forme hermitienne $x_1 \bar{x}_1 + \dots + x_n \bar{x}_n$ qui sera noté $U(n)$.

Une matrice de $U(n)$ est dite unitaire.

Proposition 1. Une matrice (n, n) A est unitaire si et seulement si ses vecteurs forment une base hermitienne de $\mathbb{C}^n$, ou encore si et seulement si ${}^t A A = I_n$.

Démonstration. Soient X et Y deux vecteurs colonnes. Leur produit scalaire hermitien est égal au produit ${}^tX\bar{Y}$. Pour qu'une matrice A soit unitaire il faut et suffit que pour tous X, Y on ait ${}^tX({}^tA\bar{A})\bar{Y} = {}^tX\bar{Y}$. Ceci est équivalent à l'égalité ${}^tA\bar{A} = I_n$. Cette identité exprime exactement le fait que les vecteurs colonnes de A constituent une base orthonormée de $\mathbb{R}^n$. En effet, si $A = (a_{i,j})$ cette identité équivaut aux relations $\sum_i a_{i,j}\bar{a}_{i,j} = 1$, pour $j = 1, \dots, n$, et $\sum_i a_{i,k}\bar{a}_{i,\ell} = 0$ si $k \neq \ell$, relations qui expriment que le système des vecteurs colonnes est hermitien. ■

Corollaire. *Le module du déterminant d'une matrice unitaire est 1.*

Démonstration. En effet on a $\det({}^tA\bar{A}) = \det({}^tA)\det(\bar{A}) = \det(A)\overline{\det(A)} = 1$. D'où le résultat. ■

Définition 1

On appelle groupe spécial unitaire et on note $SU(n)$ le sous-groupe constitué par les matrices orthogonales de déterminant 1.

Le groupe $U(1)$ est isomorphe au groupe U des nombres complexes de module 1.

Matrices unitaires et diagonalisation

Définition 2

Soit E un espace vectoriel hermitien, et soit F un sous-espace. L'orthogonal de F est par définition l'ensemble des vecteurs $v \in E$ tels que $b(x, v) = 0$ pour tout $x \in F$. On le note $F^\perp$. C'est un sous-espace vectoriel de E .

Proposition 2. *L'espace vectoriel E est la somme directe de F et de $F^\perp$.*

Démonstration. Elle est analogue à celle donnée dans le cas réel. ■

On dit que F et $F^\perp$ forment une décomposition de E en somme directe orthogonale, et on note $E \cong F \perp F^\perp$. On notera que l'on a $(F^\perp)^\perp = F$. Le lemme suivant, analogue à celui du cas réel, est très utile :

Lemme 1. *Soit E un espace vectoriel hermitien, soit f une application unitaire, et soit F un sous-espace tel que $f(F) \subset F$. Alors on a $f(F^\perp) \subset F^\perp$.*

Démonstration. L'application f restreinte à F induit un isomorphisme de F dans lui-même, étant injective et prenant ses valeurs par hypothèse dans F . Soit $v \in F^\perp$, on a donc pour tout $x \in F$ la relation $b(f(v), x) = b(v, f^{-1}(x)) = 0$, comme $f^{-1}(x)$ décrit tout F on a $f(v) \in F^\perp$. ■

Nous allons maintenant montrer que les matrices unitaires sont diagonalisables.

Théorème 1. *Soit E un espace vectoriel hermitien de forme quadratique h . Soit $f \in \mathbf{U}(h)$. Toutes les valeurs propres de f sont de module 1. De plus il existe une base orthonormée de E dans laquelle f a une matrice diagonale.*

La traduction matricielle de cet énoncé est la suivante. Soit $A \in \mathbf{U}(n)$, alors il existe une matrice unitaire P telle que la matrice $P^{-1}AP$ soit diagonale, les éléments de la diagonale étant des nombres complexes de module 1.

Démonstration. On démontre d'abord que les valeurs propres sont de module 1. Si v est un vecteur propre associé à la valeur propre λ , on a $f(v) = \lambda v$ et $b(v, v) = b(f(v), f(v)) = b(\lambda v, \lambda v) = |\lambda|^2 b(v, v)$. Donc on a $|\lambda| = 1$.

On va maintenant raisonner par récurrence sur la dimension n de E .

Soit v un vecteur propre de f . Comme $\mathbb{C}$ est algébriquement clos il en existe. Le sous-espace vectoriel F orthogonal à v est stable par f . La restriction de f à F est unitaire, et donc on peut par hypothèse de récurrence trouver une base de F dans laquelle f a pour matrice une matrice diagonale.

Dans la base de E constituée par cette base de F et v l'application f a une matrice diagonale. ■

Topologie des groupes unitaires

Nous étudions ici quelques propriétés topologiques des groupes unitaires.

Proposition 3. *Le groupe $\mathbf{U}(n)$ est compact.*

Démonstration. Munissons l'espace vectoriel $\mathbf{M}_n(\mathbb{C})$ de la norme définie par $\|A\| = \sqrt{\sum |a_{i,j}|^2}$, pour $A = (a_{i,j})$.

Observons d'abord que $\mathbf{U}(n)$ est borné. En effet, on a $\sum_i |a_{i,j}|^2 = 1$, pour $j = 1, \dots, n$, et donc $\|A\| = n$.

Par ailleurs $\mathbf{U}(n)$ est fermé. La caractérisation des matrices unitaires nous dit que c'est l'image inverse du point $(\underbrace{1, \dots, 1}_{n \text{ fois}}, \underbrace{0, \dots, 0}_{n(n-1)/2 \text{ fois}})$ par l'application continue de

$\mathbf{M}_n(\mathbb{C})$ dans $\mathbb{C}^n \times \mathbb{C}^{\frac{n(n-1)}{2}}$ donnée par

$$A \mapsto \left(\left(\sum_i |a_{i,j}|^2, \quad j = 1, \dots, n \right), \left(\sum_i a_{i,k} \overline{a_{i,\ell}}, \quad k \neq \ell \right) \right).$$

■

Proposition 4. *Le groupe $U(n)$ est connexe par arcs.*

Démonstration. Il suffit de relier une matrice $A \in U(n)$ quelconque à la matrice identité par un arc continu. Pour cela effectuons un changement de base qui mette A sous forme diagonale. Soit P la matrice de changement de base. Considérons la matrice

$$\begin{pmatrix} e^{i\lambda_1 t} & & \\ & \ddots & \\ & & e^{i\lambda_n t} \end{pmatrix}.$$

Pour $t = 0$ cette matrice est égale à la matrice identité et pour $t = 1$ elle est égale à la matrice $P^{-1}AP$. En multipliant par P et P^{-1} on obtient un arc continu joignant I_n et A . ■

Exercices

I. Formes bilinéaires et quadratiques

1. a) Soit p un nombre premier impair, et soit E un espace vectoriel de dimension finie sur le corps $\mathbb{F}_p$ à p éléments. Soit q une forme quadratique sur E . Montrer qu'on peut la mettre sous la forme $x_1^2 + \cdots + x_i^2 + \alpha x_{i+1}^2$, avec $\alpha \in \mathbb{F}_p$. On utilisera l'exercice 6 du II du chapitre 4.
- b) Montrer que i est bien déterminé, et que α est bien déterminé, à multiplication près par le carré d'un élément du corps.
2. a) Vérifier que la formule $(Aq)(x, y) = q(ax + by, cx + dy)$, où $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est une matrice $(2, 2)$ de $GL(\mathbb{F}_5)$, et où q est un polynôme homogène de degré 2 en deux variables x et y à coefficients dans le corps $\mathbb{F}_5$, définit une action du groupe $GL(\mathbb{F}_5)$ sur l'espace vectoriel P des polynômes homogènes de degré 2 en deux variables x, y à coefficients dans $\mathbb{F}_5$.
- b) Calculer le cardinal de P , ceux de $GL(\mathbb{F}_5)$ et de $SL(\mathbb{F}_5)$. Enfin celui de $G = PSL(\mathbb{F}_5) = SL(\mathbb{F}_5)/\text{cong}SL(\mathbb{F}_5)/D$ où D est le sous-groupe des homothéties de déterminant 1.
- c) En utilisant l'exercice précédent montrer que l'action de $GL(\mathbb{F}_5)$ sur P à 5 orbites, correspondant aux polynômes $0, x^2, 2x^2, x^2 + y^2, x^2 + 2y^2$. Calculer les cardinaux de ces orbites.
- d) En considérant l'espace projectif associé à P , c'est-à-dire le quotient de $P - \{0\}$ par la relation d'équivalence qui identifie deux polynômes (homogènes de degré 2 en x et y) si ils sont multiple scalaire (non-nul) l'un de l'autre, déduire une action de G sur des ensembles à 3, 10 et 15 éléments.

- e) En identifiant entre $q(x, y)$ et $q(2x, 2y)$ dans l'orbite à 10 éléments, déduire une action de G sur un ensemble à 5 éléments. En déduire un homomorphisme de G dans $\mathfrak{S}_5$. Montrer que G est isomorphe à $\mathcal{A}_5$.
3. Soit un espace vectoriel complexe E muni d'une forme sesquilinéaire b dont la forme hermitienne associée est définie positive. On appelle adjoint d'une application linéaire f l'unique application linéaire f^* telle que pour tous x et y on ait $b(x, f(y)) = b(f^*(x), y)$.
- a) Montrer l'existence et l'unicité de l'adjoint.
- b) En donner une description matricielle.
- c) Une application est dite normale si et seulement si elle commute avec son adjointe. Montrer que les applications normales sont diagonalisables.
4. On considère l'espace vectoriel $\mathbb{R}^2$. Décrire le groupe orthogonal de la forme quadratique $x_1^2 - x_2^2$.
5. Montrer que l'application qui aux matrices A et B associe $\text{Tr}(AB)$ est une forme bilinéaire sur $M_n(\mathbb{R})$. Déterminer sa signature.
6. a) On considère un espace vectoriel complexe E muni d'une forme hermitienne définie positive h . Soit f un endomorphisme de E , hermitien, défini positif, ce qui signifie que :
- f est égal à son adjoint,
 - pour tout $v \in E$, $v \neq 0$, on a $h(f(v)) > 0$.
- Montrer qu'il existe un unique endomorphisme hermitien défini positif g tel que $g^2 = f$.
- b) En déduire que tout élément de $\text{GL}(E)$ s'écrit sous la forme us , où u est un endomorphisme unitaire et s est un endomorphisme hermitien positif. Montrer que cette écriture est unique.
- c) Donner une traduction matricielle du résultat précédent.
7. ** On considère une forme bilinéaire symétrique b sur un espace vectoriel E sur un corps k de caractéristique différente de 2. On suppose que la forme b est non dégénérée. Soit F un sous-espace vectoriel de E . Soit U le noyau de la forme b restreint à F , et soit $\{x_1, \dots, x_k, \dots, x_r\}$ une base de F , telle que $\{x_1, \dots, x_k\}$ soit une base de U . Montrer par récurrence qu'il existe des vecteurs $y_1, \dots, y_k$ tels que :
- $b(x_i, y_i) = 1$ pour $i = 1, \dots, k$,
 - $b(x_i, y_j) = 0$ pour $i \neq j$, $i = 1, \dots, r$ et $j = 1, \dots, k$,
 - $b(y_i, y_i) = 0$ pour $i, j = 1, \dots, k$,
 - le système $\{x_1, \dots, x_k, \dots, x_r, \dots, y_1, \dots, y_k\}$ est libre,
 - enfin l'intersection du sous-espace vectoriel engendré par les x_i et les y_j avec F est égal à U .

8. On considère une forme bilinéaire symétrique b sur un espace vectoriel E sur un corps k de caractéristique différente de 2. Démontrer à partir du théorème de Witt que des sous-espaces totalement isotropes maximaux ont même dimension.
9. a) ** On considère une forme bilinéaire symétrique b sur un espace vectoriel E sur un corps k de caractéristique différente de 2. On suppose que la forme b est non dégénérée. Soit F et F' des sous-espaces vectoriels de E . On suppose donc donnée une bijection linéaire de F vers F' respectant la forme b . En utilisant l'exercice 6 montrer que l'on peut supposer que la restriction de la forme b sur F est non dégénérée.
- b) On reprend les notations précédentes et on suppose la restriction de b non dégénérée sur F et F' . Démontrer le théorème par récurrence sur la dimension de F .
10. Démontrer le théorème de Witt dans le cas général.
11. On dit qu'une forme bilinéaire b sur un espace vectoriel réel E est alternée si pour tous $v, w \in E$ on a $b(v, w) = -b(w, v)$. On dit qu'elle est non dégénérée si $b(v, w) = 0$ pour tout w implique que $v = 0$.
On suppose que E est de dimension finie sur un corps de caractéristique différente de 2. Montrer que l'on peut en trouver une base $x_1, \dots, x_n, y_1, \dots, y_n$ telle que, que :
- $b(x_i, y_i) = 1$,
 - $b(x_i, y_j) = 0$, si $i \neq j$.

II. Applications topologiques

1. On considère l'espace vectoriel $\mathbb{R}^3$. Le groupe orthogonal de la forme quadratique $x_1^2 + x_2^2 - x_3^2$ est-il compact ?
2. On se place dans $M_n(\mathbb{C})$. Montrer que l'application exponentielle est un homéomorphisme de l'ensemble des matrices antihermitiennes (telle que ${}^t \bar{A} = -A$) vers l'ensemble des matrices hermitiennes définies positives, c'est-à-dire telle que pour tout vecteur colonne non nul X on ait ${}^t \bar{X} S X > 0$.
3. On identifie le groupe $SO(n-1)$ au sous-groupe H de $SO(n)$ constitué par les matrices dont le terme sur la n -ième ligne et la n -ième colonne vaut 1. On considère l'action par translation de ce sous-groupe sur $SO(n)$. Construire une application continue de $SO(n)$ vers la sphère S^{n-1} qui détermine une bijection de l'espace quotient de cette action vers S^{n-1} .

4. a) Montrer que le groupe $SU(2)$ est isomorphe et homéomorphe au groupe des quaternions de norme 1, et est donc homéomorphe à la sphère S^3 .
- b) On considère l'homomorphisme du groupe U des nombres complexes de module 1 vers $SU(2)$ qui à λ associe $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$. Cet homomorphisme identifie U avec un sous-groupe de $SU(2)$, on considère l'action par translation de ce sous-groupe sur $SU(2)$. Construire une application continue de $SU(2)$ vers la sphère S^2 qui détermine une bijection de l'espace quotient de cette action vers S^2 .
5. Montrer que le groupe $U(n)$ est homéomorphe, mais n'est pas isomorphe au produit $U(1) \times SU(n)$.
6. a) Soit $A \in SO(3)$, et soit $v \in \mathbb{R}^3$. Étudier l'adhérence de l'ensemble $\{Uv \mid U \in \langle A \rangle\}$, $\langle A \rangle$ désignant le sous-groupe de $SO(3)$ engendré par A .
- b) Généraliser à $SO(n)$.

III. Groupes orthogonaux et unitaires

1. Montrer que le quotient de $SO(4)$ par son centre est isomorphe à $SO(3) \times SO(3)$. Décrire les sous-groupes distingués de $SO(4)$.
2. On considère la forme quadratique $q = x_1^2 + \cdots + x_n^2$ sur $\mathbb{F}_p^n$, p premier impair. Calculer le nombre d'éléments du groupe $O(q)$.
3. Montrer que tout automorphisme du corps des quaternions est de la forme $x \mapsto qx\bar{q}$ où q est un quaternion de norme 1.
4. Montrer que tout automorphisme du groupe $SO(3)$ est un automorphisme intérieur.
5. ** Déterminer le centre du groupe $O(n)$. On suppose que $n \geq 5$. Montrer qu'un sous-groupe distingué de $SO(n)$ est, soit contenu dans le centre, soit égal au groupe tout entier.
6. Dans le groupe unitaire $U(n)$ on considère le groupe T des matrices diagonales. Soit N le sous-ensemble de $U(n)$ constitué par les matrices A telle que $ATA^{-1} \subset T$.
- a) Montrer que N est un sous-groupe de $U(n)$ et que T est un sous-groupe distingué de N .
- b) Montrer que le groupe quotient N/T est isomorphe au groupe symétrique $\mathfrak{S}_n$.

7. a) ** Soit α un homomorphisme du groupe U des nombres complexes de module 1 vers le groupe $U(n)$. Soit $z \in U$, on pose $c_\alpha(z) = \text{Tr}(\alpha(z))$. Montrer qu'il existe des entiers positifs non nuls $a_1, \dots, a_k$ dont la somme est égale à n , et des entiers relatifs $e_1, \dots, e_k$ tels que $c(z) = \sum_i a_i z^{e_i}$. Montrer que ces entiers sont bien déterminés.
- b) Montrer que si α et β sont deux homomorphismes conjugués, c'est-à-dire tels qu'il existe une matrice unitaire U avec $\alpha(z) = U\beta(z)U^{-1}$ pour tout z , alors les entiers a_i et e_i sont les mêmes pour α et β . Étudier la réciproque.
8. Montrer que toute matrice orthogonale est le produit d'au plus n réflexions.
9. ** Compléter la classification des sous-groupes finis de $SO(3)$, (voir [A], [Ar], [Be]).
10. Le groupe de Lorentz. On considère $\mathbb{R}^4$ munit de la forme quadratique $x^2 + y^2 + z^2 - t^2$ et le groupe orthogonal associé.
- a) Montrer que ce groupe n'est pas compact comme sous-espace des matrices $(4, 4)$ réelles (comparer avec l'exercice 1).
- b) Montrer que sa composante connexe est homéomorphe à $SO(3) \times \mathbb{R}$.
11. a) Déterminer l'équation des classes du groupe $\mathcal{A}_5$.
- b) Montrer que le groupe du dodécaèdre est simple (raisonner géométriquement). Trouver son équation des classes.
- c) Montrer qu'un sous-groupe d'ordre 60 de $SO(3)$ à la même équation des classes que $\mathcal{A}_5$ (utiliser les cardinaux des orbites). En déduire qu'ils sont isomorphes.

■ Quelques réponses ou indications ■

I. Formes bilinéaires et quadratiques

1. a) b) Commencer par démontrer le résultat en dimension 2 pour une forme que l'on aura déjà réduite comme une combinaison linéaire de carrés. On cherchera un vecteur sur lequel la forme quadratique prend la valeur 1, puis on considère un orthogonal. Puis faire une récurrence sur la dimension. On raisonnera en utilisant la forme bilinéaire symétrique associée.
2. a) b) Le cardinal de P est 125, celui de $GL(\mathbb{F}_5)$ est 480, celui de $SL(\mathbb{F}_5)$ 120. Enfin celui de G est 60 car celui de D est 2.
- c) d) L'exercice précédent montre que par un changement de base on peut ramener le polynôme q (la forme quadratique) à l'un de ceux qui suivent : 0 , x^2 , $2x^2$, $x^2 + y^2$, $x^2 + 2y^2$. Les cardinaux des orbites sont 1, 12, 12, 60 et 40. Pour vérifier cela il suffit de calculer le sous-groupe fixant le polynôme donné.

2. e) L'espace projectif $\tilde{P}$ associé à P a 31 éléments. On considère la restriction de l'action de $\mathrm{GL}(\mathbb{F}_5)$ sur P à $\mathrm{SL}(\mathbb{F}_5)$. Puis que si on note $\tilde{q}$ la classe d'un polynôme q dans $\tilde{P}$ alors la formule $A\tilde{q} = \tilde{A}q$ définit une action de $\mathrm{SL}(\mathbb{F}_5)$ sur $\tilde{P}$. Enfin on montrera que le sous-groupe D agit trivialement, et on en déduira par quotient une action de G sur $\tilde{P}$. On montrera que les orbites ont 3, 3, 15 et 10 éléments.

Notons $\bar{q} = q(2x, 2y)$. Montrer que pour tout h dans l'orbite à 10 éléments et pour $A \in G$ on a $\bar{A}h = A\bar{h}$. On en déduit que l'ensemble quotient à une action de G .

L'homomorphisme associé à cette action est un homomorphisme de G dans $\mathfrak{S}_5$. On montre qu'il est surjectif sur $\mathfrak{A}_5$ et on compare les ordres. Pour montrer la surjectivité on trouve dans G des éléments d'ordre 2, 3 et 5 d'image non nulle. Par exemple les classes des

$$\text{matrices } \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

3. Cet exercice est très classique, on exhibe un sous-espace propre de f et on constate que l'orthogonal est stable par f . La restriction de l'endomorphisme à l'orthogonal est encore normale, on fait une récurrence sur la dimension.

4. On cherchera un paramétrage à l'aide des fonctions sinus et cosinus hyperboliques.

6. a) On diagonalisera l'endomorphisme. Comme il est positif les termes diagonaux de la matrice sont positifs. On en considère les racines carrées. Pour l'unicité on utilisera le fait que f et g commutent par construction et donc laissent leurs sous-espaces vectoriels respectifs stables.

7. On raisonnera par récurrence sur k .

9. b) On fera un raisonnement par récurrence sur la dimension de F . Si la dimension de F est 1 on construira une réflexion appropriée. Puis en dimension supérieure on cherchera à écrire F comme somme directe de deux sous-espaces F_0 et F_1 non triviaux et orthogonaux (on a $b(u, v) = 0$ pour tout $u \in F_0$ et tout $v \in F_1$). Puis on appliquera l'hypothèse de récurrence à F_0 .

11. Raisonner par récurrence.

II. Applications topologiques

1. Non, voir l'exercice 3. I.

2. On pourra diagonaliser les matrices.

3. L'application est celle qui à une matrice associe sa dernière colonne.

4. a) b) Montrer qu'un élément général de $\mathrm{SU}(2)$ s'écrit $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$ où u et v sont des nombres complexes tels que $|u|^2 + |v|^2 = 1$, conclure. On identifie la sphère S^2 avec $\mathbb{C}$ complété d'un point à l'infini par projection stéréographique [Be]. Puis on considère l'application qui à la matrice $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$ associe le complexe $\frac{u}{v}$ si $v \neq 0$, le point à l'infini si $v = 0$. On écrira la formule qui prend directement valeurs dans S^2 .

5. Multiplier une colonne de la matrice par le déterminant pour obtenir une matrice de déterminant 1.
6. a) On obtient soit un ensemble fini de points si A est d'ordre fini, soit un cercle si A est d'ordre infini. Pour étudier le problème on choisira une base dans laquelle l'axe de rotation de A est un des axes du repère.

III. Groupes orthogonaux et unitaires

1. On pourra utiliser le cours sur les quaternions.
3. Montrer qu'un automorphisme laisse fixe globalement l'ensemble des quaternions purs.
4. On utilisera le fait que l'image d'un élément d'ordre 2 est encore un élément d'ordre 2 et la description des éléments d'ordre 2 dans le groupe $SO(3)$.
6. On pourra montrer que N est le plus petit sous-groupe qui contient T et les matrices monomiales.
7. b) On observera que l'on peut diagonaliser simultanément tous les $\alpha(z)$ et on étudiera les éléments diagonaux.

Bibliographie

- [A] ARTIN M., *Algebra*, Prentice Hall, 1991.
- [Ar] ARNAUDIES J.-M., *Les cinq polyèdres de $\mathbb{R}^3$ et leurs groupes*, C.D.U.-S.E.D.E.S, Paris, 1969.
- [Be] BERGER M., *Géométrie 1 et 2*, Nathan, 1990.
- [Co] COMTET L., *Analyse combinatoire*, tomes 1 et 2, P.U.F., 1970.
- [De] DEMAZURE M., *Primalité, divisibilité, codes*, Cassini, 1997.
- [Del] DELEZOÏDE P., *Exercices résolus d'algèbre du cours de mathématiques 1*, Dunod 1994.
- [FS] FADDEEV D.K., SOMINSKY I.S., *Recueils d'exercices d'algèbre supérieure*, Mir 1961, Ellipses.
- [Go] GODEMENT R., *Cours d'algèbre*, Hermann, 1963.
- [J] JACOBSON N., *Lectures in abstract algebra*, Springer, 1964.
- [Ko] KOBLITZ N., *A course in Number Theory and Cryptography*, Springer, GTM 114, 1187.
- [L] LANG S., *Algebra*, Addison-Wesley Publishing Company, 1965.
- [LA] LELONG-FERRAND J., M. ARNAUDIES J., *Cours de mathématiques*, tome 1, Dunod, 1974.
- [LM] LIRET F., MARTINAIS D., *Cours de DEUG, Algèbre 1^{re} et 2^e année*, Dunod, 1997.
- [LN] LIDL R., NIEDERREITER H., *Finite fields*, Encyclopedia of Mathematics and its applications, Addison-Wesley Publishing Company, 1983.
- [Mo] MONIER J.M., *Nouveau cours de mathématiques, Algèbre 1 et 2*, Dunod, 1996.
- [MS] MACWILLIAMS F.J., SLOANE N.J.A., *The theory of error correcting codes*, North-Holland, 1977.
- [Pe] PERRIN D., *Cours d'algèbre*, Ellipses, 1996.

Index

- action
 - à gauche, 51
 - transitive, 79
- algèbre de Boole, 135
- algorithme
 - d'Euclide, 3
 - de Berlekamp, 174
- anneau, 85
 - euclidien, 102
 - factoriel, 122
 - intègre, 88
 - principal, 97
- automorphisme, 14
- Berlekamp, 174
- borne
 - de Gilbert-Varshamov, 197
 - de Plotkin, 197
- caractéristique d'un corps, 150
- Cauchy, 4, 33, 61
- Cayley, 53, 211
- centralisateur, 58
- centre, 17
- classe de conjugaison, 54
- clôture algébrique, 161
- code, 180
 - BCH*, 183
 - correcteur d'erreurs, 178
 - linéaire, 181
- conjugaison, 54
- corps, 87
 - algébriquement clos, 160
 - de base, 153
 - de rupture, 159
 - des quotients, 89
 - fini, 162
 - premier, 151
- critère d'Eisenstein, 126
- cryptographie, 105
- cycle, 62
- D'Alembert, 161
- décomposition canonique en cycles, 67
- décomposition des homomorphismes, 26
- degré, 113
 - d'une extension, 154
- dérivation formelle, 129
- dilatation, 225
- dimension du code, 180
- discriminant, 132
- distance
 - de Hamming, 179
 - minimale, 180
- division euclidienne, 114
- dual, 8
- élément
 - algébrique, 155
 - irréductible, 97
 - transcendant, 157
- endomorphisme, 14
- ensemble transitif, 56, 57
- entiers de Gauss, 137
- espace vectoriel quotient, 8
- Euclide, 3
- Euler
 - formule, 142
 - indicateur, 103
- extension
 - algébrique, 158
 - d'un corps, 153
 - degré, 154
- Fermat, 152
 - petit théorème, 58, 152
- fixateur, 55

- forme
 - bilinéaire, 245
 - non dégénérée, 248
 - symétrique, 247
- hermitienne, 246
- linéaire, 8
- quadratique, 247
- sesquilinéaire, 246
- formule
 - d'Euler, 142
 - de Leibniz, 129
 - de Moivre, 257
 - de Newton, 120, 144
 - des classes, 59

- Galois, 73
- Gauss, 124, 137, 161
- générateurs, 18
- groupe, 12
 - abélien, 12
 - abélien libre, 37
 - alterné, 71
 - commutatif, 12
 - cyclique, 20
 - de torsion, 36
 - de type fini, 36
 - dérivé, 45
 - diédral, 30
 - linéaire, 8, 225
 - monogène, 20
 - orthogonal, 255
 - quaternionien, 32
 - quotient, 22, 24
 - spécial linéaire, 230
 - symétrique, 64
 - unitaire, 273

- Hamilton, 211
- homomorphisme, 14, 87
 - associé, 52
 - image, 15
 - noyau, 15

- idéal
 - bilatère, 90
 - maximal, 95
 - premier, 95
- identité de Bézout, 2, 101
- image d'un homomorphisme, 15
- indéterminée, 111
- indicateur d'Euler, 103
- inégalité de Cauchy-Schwarz, 255
- inversion, 70
- isomorphisme, 14

- k -cycle, 66

- Laplace, 161
- Leibniz, 129
- lemme
 - chinois, 35
 - d'Euclide, 100
 - de Cauchy, 33, 62
 - de Gauss, 100
 - de la base télescopique, 157
 - de Lazard, 141
- loi interne, 3, 12
 - associativité, 12
- longueur du code, 180

- matrice
 - compagnon, 213
 - de parité, 180
 - génératrice, 180
 - triangulation, 214

- Newton, 120
- noyau d'un homomorphisme, 15

- orbite, 56
- ordre, 20

- p -torsion, 33
- partition d'un entier, 69
- plan hyperbolique, 254

- polynôme, 112
 - caractéristique, 210
 - cyclotomique, 169
 - de parité, 183
 - minimal, 156
 - réducteur, 175
 - symétrique, 118
- produit semi-direct, 27
- quaternions, 264
- racine, 127
 - n -ième de l'unité, 169
 - primitive, 169
- réduction de Jordan, 214
- réflexion, 259
- relation de congruence, 16
- résultant, 130
- retournement, 259
- Rivest-Shamir-Adleman, 105
- RSA, 105
- série
 - formelle, 111
 - réciproque, 135
- signature, 70
 - d'une forme quadratique, 253
 - d'une permutation, 70
- sous-espace
 - caractéristique, 213
 - propre, 207
 - totalelement isotrope, 255
- sous-groupe, 13
 - de Sylow, 61
 - de torsion, 37
 - distingué, 24
 - engendré, 18
- stabilisateur, 55

- suite
 - de Cauchy, 4
 - de Sturm, 141
- support, 65
- Sylvester, 253
- symétrique, 245
- système RSA, 105
- théorème
 - d'isomorphismes, 26
 - de Cayley, 53
 - de Cayley-Hamilton, 211
 - de Gauss-Lucas, 143
 - de König-Rados, 196
 - de Lagrange, 23
 - de Laplace-D'Alembert-Gauss, 161
 - de Rouché-Fontené, 9
 - de Sylow, 61, 62
 - de Wedderburn, 172
 - de Wilson, 152
 - de Witt, 255
 - des noyaux, 206
- torsion, 20
- transposition, 66
- valeur propre, 207
- valuation, 113
- vecteur
 - isotrope, 254
 - propre, 207
- Wedderburn, 172
- Wilson, 152

047057-(I)-(1,5)-OSB 80°-AUT-ABS
STEDI, 1, boulevard Ney, 75018 Paris
Dépôt légal, Imprimeur, n° 7920
Dépôt légal : avril 2003
Imprimé en France



2^e édition

Lionel Schwartz

ALGÈBRE 3^e ANNÉE

Prolongement du cours de 1^{re} et 2^e année de François Liret et Dominique Martinais, ce cours de mathématiques traite en quatre volumes le programme de la troisième année de Licence.

Trois notions centrales sont abordées dans ce volume d'algèbre :

- la structure des groupes et les actions de groupes, avec une insistance particulière sur les groupes symétriques ;
- les anneaux, avec la question de la divisibilité sous ses divers aspects ;
- les extensions de corps enfin, et notamment ce qui concerne les corps finis.

L'auteur revient également sur la réduction des endomorphismes, les formes quadratiques, et sur les propriétés topologiques des groupes classiques.

À un niveau plus avancé, une introduction à deux applications de ces mathématiques est proposée : la cryptographie et les codes correcteurs d'erreurs, prolongements naturels de ce cours.

Enfin, des exercices viennent compléter chaque chapitre, et permettent d'approfondir certains points essentiels ou d'aborder des résultats plus spécifiques.

Dans cette seconde édition, des compléments en géométrie, codes correcteurs d'erreurs, cryptographie sont apportés. Quelques aspects complexes du cours ont été précisés et les exercices ont été renouvelés.

COURS DE MATHÉMATIQUES

Ce cours de mathématiques traite en quatre volumes le programme de la troisième année de Licence.

- Algèbre
- Topologie et Analyse
- Fonctions analytiques
- Calcul différentiel et Calcul intégral

LIONEL SCHWARTZ
est professeur à
l'université Paris-Nord
(Paris 13-Villetaneuse).

MATHÉMATIQUES

PHYSIQUE

CHIMIE

SCIENCES DE L'INGÉNIEUR

INFORMATIQUE

SCIENCES DE LA VIE

SCIENCES DE LA TERRE

9 782100 070572

ISBN 2 10 007057 6