

J. M. Arnaudiès
H. Fraysse

Cours de mathématiques - 1

Algèbre

Classes préparatoires
1^{er} cycle universitaire

Dunod Université

Cours de mathématiques - 1

Algèbre

Cours de mathématiques - 1

Algèbre

Jean-Marie ARNAUDIÈS Henri FRAYSSE

*Professeurs de Mathématiques Spéciales
au Lycée Pierre de Fermat à Toulouse
Anciens élèves de l'École Normale Supérieure*

DUNOD

Nouveau tirage, 1992.

© BORDAS, Paris, 1987
ISBN 2-04-016450-2

“ Toute représentation ou reproduction, intégrale ou partielle, faite sans le consentement de l’auteur, ou de ses ayants-droit, ou ayants-cause, est illicite (loi du 11 mars 1957, alinéa 1^{er} de l’article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait une contrefaçon sanctionnée par les articles 425 et suivants du Code pénal. La loi du 11 mars 1957 n’autorise, aux termes des alinéas 2 et 3 de l’article 41, que les copies ou reproductions strictement réservées à l’usage privé du copiste et non destinées à une utilisation collective d’une part, et, d’autre part, les analyses et les courtes citations dans un but d’exemple et d’illustration ”

PRÉFACE

Les changements du programme des classes préparatoires intervenus en 1984 et 1985, et l'évolution de l'enseignement, ont nécessité la refonte complète du cours de mathématiques écrit en collaboration avec Mme J. Lelong-Ferrand et réimprimé par les éditions DUNOD plusieurs fois depuis 1971.

Nous présentons ici le premier tome, consacré à l'algèbre, de ce nouveau cours de mathématiques.

Nous avons essayé de fournir à l'étudiant un ouvrage qu'il puisse utiliser depuis son entrée en classe préparatoire jusqu'au moment où il « intègre » une grande école. Face à l'abondance des brochures courtes rédigées par thèmes, et des recueils d'exercices destinés à nos étudiants, il nous a paru indispensable de continuer à leur proposer un ouvrage d'approfondissement complet, offrant une vue d'ensemble des programmes et une large gamme *d'exercices classés*.

Fidèles à l'esprit qui a guidé les précédents cours de mathématiques, nous avons voulu créer un outil de travail qui soit aussi l'occasion d'un contact avec la science. Nous espérons qu'il rendra service.

Si les mathématiques sont devenues un peu trop un instrument de « sélection », c'est plutôt à cause des qualités propres à cette science (précision, rigueur), que par un dessein explicite formé chez les mathématiciens, enseignants ou non. Pour les premiers, le souci essentiel, dans l'austère et exigeante contrainte des concours ou examens, demeure la transmission correcte de leur discipline et l'encouragement aux nouvelles vocations. Peut-être les aiderons-nous un peu dans leur tâche ; qu'ils veuillent bien accorder leur indulgence aux inévitables défauts de notre travail qui, bien évidemment, ne pourra jamais *remplacer les explications vivantes d'un professeur*.

Les auteurs tiennent à remercier les nombreuses promotions d'étudiants dont ils ont eu la charge ; le présent traité est enrichi de bon nombre de leurs idées ; ils ne sont pas tous entrés à « Polytechnique » ou à « Normale », mais quelquefois c'était le plus modeste de la classe, celui qui suivait avec peine, qui avait l'idée lumineuse.

Nous remercions vivement les éditions DUNOD du soin apporté, comme toujours, à la composition et à l'impression de ces livres, pour lesquels les éditeurs ont donné le meilleur d'eux-mêmes. Ils ont surmonté de redoutables difficultés techniques ; l'inlassable compétence de Gisèle Maïus et de Pierre Riotort a été décisive.

J. M. ARNAUDIÈS, H. FRAYSSE

AVERTISSEMENT

En Algèbre, l'idée inspiratrice des nouveaux programmes des classes préparatoires est qu'il faut mettre l'accent sur la dualité parfaite et l'interaction totale entre d'une part le *concret*, l'*observation* et l'*expérimentation* qu'ils suscitent, et d'autre part la théorisation, synthèse généralisatrice et unificatrice qui permet de démarrer un cran plus loin un nouveau cycle de recherches.

Cette philosophie, qui a guidé les rédacteurs de ces programmes, est très ambitieuse ; elle revient en effet à souhaiter que l'enseignement, dans son esprit et ses méthodes, suive d'aussi près que possible le cheminement de la science active (la dualité décrite plus haut). Or, c'est là une tâche difficile, car il faut beaucoup de temps et de patience pour allier équitablement, dans un exposé où précisément le temps est compté, le général et le particulier, la théorie et l'application ; de toutes façons, il est impossible de recréer toutes les étapes ayant abouti aux résultats que l'on enseigne, et l'on ne peut pas faire repartir chaque génération de zéro.

Mais cette philosophie, bien comprise, s'imposait tout de même : il fallait rappeler que la théorie a besoin de *matière première* pour se développer, et qu'elle va toujours très loin lorsqu'elle est issue d'un problème réel que nous pose le monde. Il fallait aussi tenir compte de ces prodigieux outils mis à notre disposition par l'électronique moderne et qui insufflent un puissant renouveau à toute la partie *algorithmique* des mathématiques, et même aux recherches sur la logique.

Nous nous sommes donc efforcés, dans ce COURS, de tenir compte au mieux de ces contraintes. *Les structures algébriques ne sont introduites que peu à peu, au fil de l'étude des nombres.* Lorsqu'elles apparaissent, elles ne sont pas étudiées systématiquement pour elles-mêmes ; on en donne juste ce qu'il faut de propriétés pour pouvoir poursuivre le texte

— La notion de groupe est définie au Chapitre II, mais n'est étudiée à fond qu'au Chapitre V, *après toute l'Arithmétique*. On constatera que cela n'a pas nui à cette dernière ; au contraire, l'Arithmétique, au cours de l'exposé, « appelle » la structure de groupe, et aussi les structures d'anneau et de corps.

— Nous n'avons développé aucune théorie suivie des *anneaux* ou des *corps*. La notion d'anneau vient juste après l'étude de \mathbb{Z} , revient à plusieurs reprises en Arithmétique, et on la retrouve encore à la fin du Chapitre VII consacré aux polynômes. Nous nous sommes bien gardés d'une théorie générale des anneaux principaux qui aurait permis, comme cas particulier, d'obtenir l'arithmétique des entiers, ou des polynômes sur un corps. Nous imposant une démarche inverse, nous avons jugé que l'étude approfondie des anneaux principaux \mathbb{Z} et $K[X]$ (si remarquables, et plus riches qu'un anneau principal abstrait) est une excellente motivation pour la *nécessité* de la structure générale d'anneau principal. C'est donc délibérément que nous n'avons même pas démontré, dans ce livre, que « tout anneau principal est factoriel », et que nous avons traité séparément \mathbb{Z} et $K[X]$.

— Les espaces vectoriels s'annoncent au Chapitre VI, d'où une plus cohérente présentation des nombres complexes, mais leur étude véritable ne commence qu'au Chapitre IX.

C'est ainsi que, dans le respect de l'esprit des nouveaux programmes, *l'introduction des grandes structures de base de l'Algèbre s'étale, très progressivement, sur les 9 premiers des 16 chapitres du livre.*

Mais nous avons aussi composé *chaque chapitre* dans ce même souci : par exemple, au Chapitre V sur les groupes, les *cycles* sont abordés *avant* la notion abstraite d'opération de groupe sur un ensemble, qui se trouve donc justifiée *a priori* par une de ses plus belles illustrations. Lorsqu'un même chapitre contient des parties enseignées en seconde année, elles sont reléguées *à la fin*, et ce qui précède dans le chapitre n'en dépend pas, comme on le vérifiera au Chapitre V, § 7, ou au Chapitre XV, §§ 5 et 6.

MODE D'EMPLOI DU LIVRE

1) Le livre est conçu pour une utilisation depuis l'entrée en classe préparatoire jusqu'au moment des concours, en fin de seconde année.

Les futurs élèves des grandes écoles, mais aussi les candidats à l'agrégation et les professeurs de lycée y trouveront, en petits caractères, tous les approfondissements désirables.

2) Les programmes d'Algèbre fondamentale et linéaire de première année sont largement couverts par les parties suivantes : Chapitres I, II, III, IV, V (moins § 7), VI, VII, VIII, IX, XI, XIII, XIV (moins § 4), et XV, §§ 1 et 3.

Ces parties peuvent s'étudier *indépendamment des autres*. Toutefois, les étudiants de 2^e année les consulteront avec profit, car depuis 1985, les programmes d'écrit et d'oral des concours portent sur *l'ensemble des matières enseignées en 1^{re} et 2^e année*.

3) Les Chapitres V (§ 7), X, XII, XIV (§ 4), XV et XVI complètent, pour les étudiants de 2^e année, le programme d'Algèbre linéaire et fondamentale. Le Chapitre X, consacré aux équations algébriques et à une présentation du prolongement des identités algébriques (qui figure explicitement au programme *M'*) a été un peu poussé, on pourra ne pas l'aborder en première lecture. Mais il nous a paru intéressant, car il illustre la notion de groupe, et donne des algorithmes.

4) Le Chapitre XVI ne doit pas être travaillé sans une connaissance sérieuse (qui implique la résolution d'exercices...) du reste de l'Algèbre linéaire, notamment du Chapitre XV.

5) Enfin, tout le temps nécessaire doit être consacré à l'étude des Chapitres I à IV, qui contiennent les bases du calcul.

TABLE DES MATIÈRES

CHAPITRE I. Vocabulaire de théorie des ensembles	1
§ 1 Un peu de logique	1
§ 2 Construction d'ensembles	7
§ 3 Correspondances et applications	14
§ 4 Familles	22
§ 5 Relations d'équivalence. Ensemble quotient	28
§ 6 Relations d'ordre	33
CHAPITRE II. Nombres entiers. Nombres rationnels	41
§ 1 Axiomes de Peano ; Récurrence	41
§ 2 Ordre naturel dans \mathbb{N}	49
§ 3 Ensembles finis, Ensembles infinis ; Ensembles dénombrables	54
§ 4 Lois de composition. Structure de groupe	62
§ 5 L'anneau des entiers relatifs, Structure d'anneau	70
§ 6 Les nombres rationnels, Structure de corps	78
CHAPITRE III. Bases du calcul algébrique et combinatoire	86
§ 1 Itération d'une loi de composition	86
§ 2 Calcul dans un anneau	91
§ 3 Composé de familles à support fini. Numération	93
§ 4 Dénombrement	95
§ 5 Formule du binôme	107
§ 6 Sous-groupes additifs de \mathbb{Z} . Application aux groupes	111
§ 7 Notion d'idéal d'un anneau commutatif	113
CHAPITRE IV. Notions d'arithmétique	119
§ 1 Congruences dans \mathbb{Z} , Anneaux $\mathbb{Z}/n\mathbb{Z}$	119
§ 2 Arithmétique dans \mathbb{Z} et \mathbb{N}	126
§ 3 Eléments inversibles des anneaux $\mathbb{Z}/n\mathbb{Z}$	135
§ 4 Nombres premiers	139
§ 5 Décomposition en facteurs premiers	:

CHAPITRE V. Groupes	153
§ 1 Génération de groupes	153
§ 2 Ordre d'un élément	159
§ 3 Classes suivant un sous-groupe. Indice	162
§ 4 Groupes de permutations	168
§ 5 Cycles dans les groupes \mathfrak{S}_E (E fini)	176
§ 6 Opération d'un groupe sur un ensemble	184
§ 7 Sous-groupes distingués. Groupe quotient	192
CHAPITRE VI. Structure d'espace vectoriel et d'algèbre ; nombres complexes	202
§ 1 Structure d'espace vectoriel	202
§ 2 Applications linéaires	206
§ 3 Combinaisons linéaires ; indépendance linéaire ; bases	211
§ 4 Structure d'algèbre sur un corps commutatif	215
§ 5 Le corps des nombres complexes	220
§ 6 Racines carrées d'un nombre complexe	224
§ 7 Nombres complexes de module 1	226
§ 8 Arguments d'un nombre complexe ; racines n -ièmes	232
§ 9 Nombres complexes et géométrie	239
§ 10 Nombres complexes et similitudes	245
§ 11 Nombres complexes, droites et cercles	252
CHAPITRE VII. Polynômes sur un corps commutatif	256
§ 1 Polynômes à une indéterminée	256
§ 2 L'anneau euclidien $K[X]$	261
§ 3 L'anneau factoriel $K[X]$	271
§ 4 Fonctions polynômes, racines	278
§ 5 Racines d'un polynôme. Formule de Taylor	291
§ 6 Factorisation dans $\mathbb{R}[X]$	303
§ 7 Congruences dans $K[X]$. Anneaux quotients	306
CHAPITRE VIII. Fractions rationnelles. Notions sur les séries formelles	315
§ 1 Le corps $K(X)$	315
§ 2 Décomposition en éléments simples	322
§ 3 Fonctions rationnelles. Dérivation	334
§ 4 Notions sur les séries formelles à une indéterminée	343
§ 5 Applications des séries formelles	351
CHAPITRE IX. Espaces vectoriels. Dimension des espaces vectoriels	365
§ 1 Sous-espaces supplémentaires, projecteurs	365
§ 2 Produits et sommes d'espaces vectoriels	369
§ 3 Espaces de dimension finie	377
§ 4 Propriétés des espaces de dimension finie	383
§ 5 Hyperplans	391

§ 6	Endomorphismes, groupe linéaire	393
§ 7	Eléments algébriques d'une extension d'un corps	397
CHAPITRE X. Fonctions polynomiales sur K^n; équations algébriques		405
§ 1	Polynômes à n lettres	405
§ 2	Dérivées partielles. Degré partiel	414
§ 3	Fonctions symétriques	418
§ 4	Formules de Newton	428
§ 5	Equations algébriques. Equations de degré 3	434
§ 6	Equations de degré 4. Equations particulières	442
CHAPITRE XI. Matrices		453
§ 1	Matrices de type (m, n)	453
§ 2	Matrices carrées	462
§ 3	Matrices et applications linéaires	470
§ 4	Rang d'une matrice	481
§ 5	Opérations élémentaires	486
§ 6	Similitude d'endomorphismes ou de matrices	491
CHAPITRE XII. Dualité. Espaces vectoriels quotients		494
§ 1	Dual ; forme bilinéaire canonique	494
§ 2	Dualité en dimension finie	497
§ 3	Quotients d'espaces vectoriels	507
§ 4	Quotients, produits et sommes directes	514
CHAPITRE XIII. Déterminants		518
§ 1	Applications multilinéaires	518
§ 2	Formes n -linéaires alternées sur E de dimension n	525
§ 3	Déterminants de n vecteurs dans une base ; déterminant d'un endomorphisme	528
§ 4	Déterminant d'une matrice carrée	533
§ 5	Exemples de déterminants	544
CHAPITRE XIV. Equations linéaires sur un corps		559
§ 1	Langage de la Géométrie affine dans un espace vectoriel	559
§ 2	Equations linéaires sur un corps ; cas d'un système de Cramer	567
§ 3	Equations linéaires sur un corps : cas général	578
§ 4	Méthodes directes de résolution ; pivot partiel	589
CHAPITRE XV. Réduction d'endomorphismes ou de matrices carrées ...		603
§ 1	Valeurs propres et polynôme caractéristique	603
§ 2	Trigonalisation	610
§ 3	Sous-espaces propres	616
§ 4	Polynômes d'endomorphismes ou de matrices	630

§ 5	Sous-espaces caractéristiques	637
§ 6	Suites définies par une relation de récurrence linéaire	654
CHAPITRE XVI. Complément : réduction de Jordan		660
§ 1	Etude des endomorphismes nilpotents	660
§ 2	Réduction de Jordan quand $\chi_u(X)$ est dissocié	666
§ 3	Sous-espaces monogènes	673
BIBLIOGRAPHIE		681
INDEX ALPHABÉTIQUE		683

INDEX DES NOTATIONS

$A \Rightarrow B ; A \Leftrightarrow B$	Implication de B par A ; équivalence des deux assertions.
$\exists x \mid \mathcal{R}(x)$	Quantificateur existentiel.
$(\forall x) \mathcal{R}(x)$	Quantificateur universel.
$x \in E, x \notin E$	Appartenance ; non-appartenance.
$\{ a_1, a_2, \dots, a_n \}$	Ensemble dont les éléments sont a_1, a_2, \dots, a_n .
$E \subset F ; E \subsetneq F$	Inclusion ; inclusion stricte.
\emptyset	Ensemble vide [certains auteurs le notent aussi $\{ \}$].
\mathbb{N}	Ensemble des entiers naturels.
\mathbb{N}^*	Ensemble des entiers naturels non nuls.
\mathbb{Z}	Ensemble des entiers relatifs (on dit aussi : entiers rationnels).
\mathbb{Q}	Ensemble des nombres rationnels.
\mathbb{Q}^*	Ensemble des nombres rationnels non nuls.
\mathbb{Q}_+	Ensemble des nombres rationnels ≥ 0 .
\mathbb{Q}_+^*	Ensemble des nombres rationnels > 0 .
\mathbb{R}	Ensemble des nombres réels.
\mathbb{R}^*	Ensemble des nombres réels $\neq 0$.
\mathbb{R}_+	Ensemble des nombres réels ≥ 0 .
\mathbb{R}_-	Ensemble des nombres réels ≤ 0 .
\mathbb{R}_+^*	Ensemble des nombres réels > 0 .
\mathbb{C}	Ensemble des nombres complexes.
\mathbb{C}^*	Ensemble des nombres complexes $\neq 0$.
$\{ x \in E \mid A(x) \}$	Ensemble des $u \in E$ qui vérifient $A(x)$.
$E \setminus F$	Ensemble des $x \in E$ tels que $x \notin F$.
$\mathbb{C}_E(F)$	Pour $F \subset E$, ensemble $E \setminus F$.
$\mathcal{P}(E)$	Ensemble des parties de l'ensemble E .
(a_1, a_2, \dots, a_p)	Suite ordonnée formée des objets a_1, a_2, \dots, a_p (les a_i distincts ou non).
$E \cup F$	Union de E et F .
$E \cap F$	Intersection de E et F .
$E_1 \times E_2 \times \dots \times E_p$	Produit cartésien des ensembles E_1, E_2, \dots, E_p .
$f : E \rightarrow F, x \mapsto f(x)$	Application f de E dans F qui envoie x sur $f(x)$.

$\mathcal{F}(E, F) = F^E$	Ensemble des applications de E dans F .
Id_E	Application $E \rightarrow E$, $x \mapsto x$.
$g \circ f$	Composée des applications f et g , qui envoie x sur $g(f(x))$.
$f _L$	Restriction de l'application f à L .
$f _H$	Corestriction de l'application f à H .
$f \parallel_L$	Application $(f _L) ^L$ pour L telle que $f(L) \subset L$, dite <i>induite</i> par f sur L .
\mathfrak{S}_E	Ensemble des bijections de E sur E .
f^{-1}	Application réciproque de la bijection f .
$(a_i)_{i \in I}$	Famille indexée par l'ensemble I .
$(u_n)_{n \in \mathbb{N}}$	Suite de terme général u_n .
$\bigcup_{i \in I} E_i$	Réunion de la famille des ensembles E_i .
$\bigcap_{i \in I} E_i$	Intersection de la famille des ensembles E_i .
$\prod_{i \in I} E_i$	Produit de la famille des ensembles E_i .
$\coprod_{i \in I} E_i$	Union disjointe des ensembles E_i .
E/\mathcal{R}	Ensemble quotient de E par la relation d'équivalence \mathcal{R} .
$x y, x \nmid y$	L'entier x divise l'entier y ; l'entier x ne divise pas l'entier y .
$\sup_E(x) = \sup_E(A)$	Borne supérieure de la partie A de l'ensemble ordonné E .
$f^{(p)}$	p -ième itérée de l'application f .
$\llbracket a, b \rrbracket$; $\llbracket a, b[$; $\llbracket a, b]$; $\llbracket a, b[$	Pour $a \in \mathbb{Z}$ et $b \in \mathbb{Z}$, ensemble des $x \in \mathbb{Z}$ tels que $a \leq x \leq b$ (resp. $a \leq x < b$, $a < x \leq b$, $a < x < b$).
$\llbracket a, \rightarrow \rrbracket$; $\llbracket a, \rightarrow [$; $\llbracket \leftarrow, a \rrbracket$; $\llbracket \leftarrow, a[$	Pour $a \in \mathbb{Z}$, ensemble des $x \in \mathbb{Z}$ tels que $x \geq a$ (resp. $x > a$, $x \leq a$, $x < a$).
$\text{card}(E)$	Cardinal de l'ensemble E .
$n!$	Factorielle de l'entier n ($0! = 1$).
A_n^p	Entier $\frac{n!}{(n-p)!}$ pour $0 \leq p \leq n$ (Arrangements).
$\binom{n}{p} = C_n^p$	Combinaisons de n objets p à p ($0 \leq p \leq n$) : c'est $\frac{n!}{p!(n-p)!}$.
$\binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n}$	Coefficient multinomial $\frac{p!}{\alpha_1! \alpha_2! \dots \alpha_n!}$ pour

$\alpha_1 + \alpha_2 + \dots + \alpha_n = p$ (les α_i et p entiers naturels).	
$x \equiv y \bmod (n)$	Congruence modulo n des entiers x et y .
$\mathbb{Z}/n\mathbb{Z}$	Anneau des classes d'entiers modulo n .
$\text{Gr}((a_i)_{i \in I})$	Groupe engendré par la famille $(a_i)_{i \in I}$.
$\prod_{i \in I} G_i$	Groupe produit des groupes G_i .
$\coprod_{i \in I} G_i$	Somme directe externe des groupes abéliens G_i .
$\Gamma^{(I)}$	Groupe (abélien) des familles à support fini d'éléments du group abélien Γ .
\mathfrak{S}_n	Groupe $\mathfrak{S}_{[1, n]}$ pour $n \in \mathbb{N}^*$. (Groupe symétrique d'ordre n).
$\varepsilon(\sigma)$	Signature de $\sigma \in \mathfrak{S}_E$ (pour E fini non vide).
\mathfrak{U}_n	Groupe alterné d'ordre $n \geq 1$.
$\langle a, c(a), \dots, c_{(a)}^{d-1} \rangle$	Cycle c de longueur d dans \mathfrak{S}_E (E fini).
$[G : H]$	Indice du sous-groupe H du groupe G dans G .
$H \triangleleft G$	H est un sous-groupe distingué dans G .
G/H	Groupe quotient de G par son sous-groupe H (pour $H \triangleleft G$).
$\text{Hom}_K(E, F)$	K -ev des applications linéaires du K -ev E dans le K -ev F .
$\text{Hom}_K(E)$	Anneau des endomorphismes du K -ev E .
$\text{Vect}_K(A)$	Sous K -ev engendré par A .
$\text{Ker}(f)$	Noyau de l'homomorphisme de groupes f (resp. de l'application K -linéaire f).
$\text{Im}(z) ; \text{Re}(z)$	Partie imaginaire (resp. réelle) de $z \in \mathbb{C}$.
\bar{z}	Conjugué de $z \in \mathbb{C}$.
$ z $	Module de $z \in \mathbb{C}$.
\mathbb{U}	Groupe multiplicatif des nombres complexes de module 1.
μ_n	Groupe des racines n -ièmes de 1 dans \mathbb{C} pour $n \in \mathbb{N}^*$.
$\widehat{(\vec{D}_1, \vec{D}_2)}$	Angle orienté des demi-droites vectorielles \vec{D}_1, \vec{D}_2 de \mathbb{C} (c'est un réel modulo $2\pi\mathbb{Z}$).
$\widehat{(D_1, D_2)}$	Angle orienté des deux droites D_1, D_2 (c'est un réel mod $\pi\mathbb{Z}$).
$\vec{\mathcal{S}}, \vec{\mathcal{S}}_+, \mathcal{S}, \mathcal{S}_+$	Groupes de similitudes dans \mathbb{C} .
$\text{O}(\mathbb{C}), \text{SO}(\mathbb{C}), \text{Is}(\mathbb{C}), \text{Is}_+(\mathbb{C})$	Groupes d'isométries dans \mathbb{C} .

$K[X]$	Algèbre des polynômes en l'indéterminée X sur le corps K .
$\deg(P)$	Degré de $P \in K[X]$ (ou de $P \in K(X)$).
$\text{val}(P)$	Valuation de $P \in K[X]$ (ou de $P \in K[[X]]$).
$\text{val}_P(F)$	P -valuation de $F \in K[X]$ ou de $F \in K(X)$ pour P irréductible dans $K[X]$.
$A \equiv B \pmod{P}$	Congruence modulo P dans $K[X]$.
$K[X]/\mathfrak{a}$	Anneau quotient de $K[X]$ par l'idéal \mathfrak{a} .
$K(X)$	Corps des fractions rationnelles en l'indéterminée X sur le corps commutatif K .
$K[[X]]$	K -algèbre des séries formelles en l'indéterminée X sur le corps commutatif K .
$\binom{\alpha}{n}$	Pour $\alpha \in \mathbb{C}$, $n \in \mathbb{N}$, coefficient binomial égal à $\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$ si $n \geq 1$, et 1 si $n = 0$.
$\prod_{i \in I} V_i$	K -ev produit des K -ev V_i .
$\coprod_{i \in I} V_i$	Somme directe externe des K -ev V_i .
$\sum_{i \in I} V_i$	Somme interne des sous K -ev V_i d'un K -ev E .
$\bigoplus_{i \in I} V_i$	Somme directe interne des sous K -ev V_i lorsqu'ils sont indépendants.
$\dim_K(E)$	Dimension du K -ev E .
E^*	Dual algébrique du K -ev E ($= \text{Hom}_K(E, K)$).
$\text{rg}((a_i)_{i \in I}); \text{rg}(f)$	Rang (de la famille de vecteurs $(a_i)_{i \in I}$; de l'application linéaire f).
$K[a_1, a_2, \dots, a_n]$	K -algèbre des polynômes en les n lettres a_1, a_2, \dots, a_n .
$\ \alpha\ $	Pour $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, entier

$$\alpha_1 + \alpha_2 + \dots + \alpha_n.$$

$\frac{\partial P}{\partial X_i}$ Dérivée partielle du polynôme P par rapport à X_i .

$\mathfrak{M}_{m,n}(A)$ Ensemble des (m, n) -matrices à coefficients dans l'anneau A .

$\mathfrak{M}_n(A)$ Ensemble $\mathfrak{M}_n(A)$.

$\mathcal{M}_{I,J}(M)$ Sous-matrices de M à lignes indexées par $I \subset \llbracket 1, m \rrbracket$ et à colonnes indexées par $J \subset \llbracket 1, n \rrbracket$.

$\mathcal{L}_i(M), \mathcal{C}_j(M)$	i -ième ligne de la matrice M ; j -ième colonne de la matrice M .
${}^tM; {}^tu$	Transposée de la matrice M ; transposé de l'endomorphisme u .
E_{ij}	Matrice élémentaire (dont les coefficients sont tous nuls sauf $a_{ij} = 1$).
$\mathcal{T}_+(n, A)$	Ensemble des matrices trigonales supérieures dans $\mathcal{M}_n(A)$.
$\mathcal{T}_-(n, A)$	Ensemble des matrices trigonales inférieures dans $\mathcal{M}_n(A)$.
$\mathcal{U}_+(n, A)$	Ensemble des matrices unipotentes supérieures dans $\mathcal{M}_n(A)$.
$\mathcal{U}_-(n, A)$	Ensemble des matrices unipotentes inférieures dans $\mathcal{M}_n(A)$.
$\text{Diag}(n, A)$	Ensemble des matrices diagonales dans $\mathcal{M}_n(A)$.
$\text{Diag}(a_1, a_2, \dots, a_n)$	La matrice diagonale de diagonale principale (a_1, a_2, \dots, a_n) .
$\text{GL}(n, A)$	Groupe des matrices inversibles dans $\mathcal{M}_n(A)$.
$\text{Tr}(M)$	Trace de la matrice $M \in \mathcal{M}_n(A)$.
$\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$	Matrice de l'endomorphisme u dans les bases \mathcal{B} (départ) et \mathcal{C} (arrivée).
$\text{rg}(M)$	Rang de la matrice M à coefficients dans un corps commutatif.
$T_{ij}(\lambda)$	Matrice de transvection.
$\text{GL}_K(E)$	Groupe linéaire de l'espace vectoriel non nul E sur le corps K .
A^0	Orthogonal de la partie A du K -ev E dans E^* .
0B	Orthogonal de la partie B de E^* dans E .
$\langle x, \varphi \rangle$	Si $x \in E$ et $\varphi \in E^*$, élément $\varphi(x)$ du corps de base K .
δ_{ij}	Symbole de Kronecker, vaut 1 si $i = j$ et 0 si $i \neq j$.
E/F	K -ev quotient de E par le sous K -ev F .
$\text{Codim}_E(F)$	Codimension du sous K -ev F du K -ev E .
$\text{ML}_n(E)$	Ensemble des formes n -linéaires sur le K -ev E .
$\text{S}_n(E); \text{A}_n(E)$	Ensemble des formes n -linéaires symétriques (resp. antisymétriques) sur E .
$\Lambda_n^*(E)$	Ensemble des formes n -linéaires alternées sur E .
$\det_{\mathcal{B}}(x_1, x_2, \dots, x_n)$	Déterminant dans la base \mathcal{B} des n vecteurs x_1, x_2, \dots, x_n .
$\det(M); \det(u)$	Déterminant de la matrice M ; de l'endomorphisme u .
$\text{SL}_K(E)$	Groupe spécial linéaire de E .

$SL(n, K)$	Groupe des matrices $M \in \mathcal{M}_n(K)$ de déterminant 1.
$\Delta_{I,J}(M)$	Mineur $\det(\mathcal{M}_{I,J}(M))$ de la matrice M pour $\text{card}(I) = \text{card}(J)$.
$(-1)^{i+j} D_{i,j}(M)$	Cofacteur d'indice (i, j) dans $M \in \mathcal{M}_n(K)$.
\check{M}	Matrice des cofacteurs de $M \in \mathcal{M}_n(K)$, appelée comatrice de M .
\tilde{M}	Matrice ${}^t(\check{M}) = ({}^t\check{M})$ pour $M \in \mathcal{M}_n(K)$.
\overrightarrow{f}	Partie linéaire de l'application affine f .
$GA(\mathcal{V})$	Ensemble des bijections affines de \mathcal{V} dans \mathcal{V} .
$\text{Aff}(E, K)$	Ensemble des fonctions affines de E dans K .
$\chi_u(X); \chi_M(X)$	Polynôme caractéristique de l'endomorphisme u ; de la matrice carrée M .
$Q_u(X); Q_M(X)$	Polynôme minimal de l'endomorphisme u ; de la matrice carrée M .
$J_\lambda(n) = J(n)$	Matrice de Jordan nilpotente d'ordre n .
$J_0(n)$	Matrice de Jordan égale à $\lambda I_n + J(n)$.
$\text{Ann}_u(A) = \text{Ann}(A)$	Idéal u -annulateur de la partie A du K -ev E .
c.-à-d. = <i>i.e.</i>	C'est-à-dire (en latin : <i>id est</i>).
cf.	Se reporter à (du latin <i>confer</i>).
C.N.S.	Condition nécessaire et suffisante.
ssi	Si et seulement si.

Chapitre I

VOCABULAIRE DE THÉORIE DES ENSEMBLES

§ I.1 UN PEU DE LOGIQUE

En Mathématiques on étudie des objets « abstraits » : nombres, vecteurs, polynômes, fonctions, courbes, surfaces, etc. Ces objets sont, en gros, des sortes de *cristallisations de concepts*, que l'on représente par des symboles d'écriture et sur lesquels on raisonne comme sur des objets du monde Physique. Il est entendu tacitement qu'un même concept peut être omniprésent grâce à sa représentation par un symbole qu'on peut reproduire à volonté.

On forme des *assertions* sur ces objets : certaines sont vraies, d'autres fausses. L'art du mathématicien consiste à mettre en évidence le plus possible d'assertions vraies, et dignes d'être énoncées du point de vue de leur utilité et de leur élégance. Le patrimoine de vérités ainsi amassé au cours du temps forme en quelque sorte un *stock de pensée concentrée*.

On renonce une fois pour toutes à la notion de *vérité* « *absolue* » : au départ de toute théorie on déclare un petit nombre d'assertions comme étant *vraies a priori* et on les appelle *axiomes* de la théorie en question. Les autres assertions de la théorie considérées comme vraies (appelées THÉO- RÈMES, PROPOSITIONS, LEMMES, SCHOLIES ou COROLLAIRES) sont celles obtenues au terme d'une *démonstration logique*, s'appuyant sur les axiomes admis au départ et sur des règles de logique qui ont, depuis des siècles, fait la preuve de leur valeur déductive.

Il va de soi que les systèmes d'axiomes des théories « utiles » et « perfor- mantes » ne doivent pas être choisis au hasard ; en effet d'une part il se pose des problèmes d'*interdépendance* et de *compatibilité entre eux* des axiomes d'un même système. D'autre part c'est l'*observation* et l'*expérimentation mathématique* qui permettent la mise au point d'un « bon » système d'axio- mes. Nous n'aborderons pas ici ces questions (voir par exemple [8], [22]).

En dépit des nombreux efforts prodigués pour tirer parti du matériel informatique moderne dans l'étude des démonstrations mathématiques, le génie créateur du mathématicien ne paraît pas près d'être :

l'ordinateur. Seule l'imagination du mathématicien est stimulée par la recherche de problèmes concrets, dont certains restent fort longtemps sans réponse et parfois sont encore non résolus de nos jours, comme les célèbres « grand théorème de Fermat » ⁽¹⁾ ou « conjecture de Goldbach » ⁽²⁾. Cependant l'informatique moderne autorise des tâches répétitives naguère hors de notre portée ; dans de nombreuses questions où il restait un très grand nombre fini de cas à examiner pour conclure, elle a permis d'achever des démonstrations là où les moyens « manuels » auraient été impuissants ; par exemple « problème des 4 couleurs », découverte de grands nombres premiers (par exemple les nombres $2^{132\,049} - 1$ et $2^{216\,091} - 1$ ont été reconnus premiers en 1985), démonstration de la conjecture de Waring ⁽³⁾, etc.

Nous nous contenterons de noter ici que le développement prodigieux de l'Informatique a replacé au premier plan l'étude de la Logique, redonnant un puissant second souffle aux travaux de Gödel ⁽⁴⁾ et de ses émules.

Dans ce qui suit, nous ne remonterons pas aux sources, et ne chercherons même pas à présenter un système d'axiomes cohérent pour la théorie des ensembles, par exemple. Nous nous contenterons de rappeler des assertions vraies qui résultent d'un tel système d'axiomes et qui sont couramment utilisées dans la pratique. Pour qu'il n'y ait aucune confusion dans l'esprit du lecteur, nous avons donné le nom de RÈGLES à ces assertions.

Connecteurs logiques, modes de raisonnement

- A chaque assertion A (vraie ou fausse) d'une théorie donnée, on associe son *contraire*, noté $(\text{non } A)$, appelé aussi *négation de A* .

Règle 1 : A est fausse ssi $(\text{non } A)$ est vraie ; A est vraie ssi $(\text{non } A)$ est fausse.

La théorie est dite *contradictoire* si elle contient une assertion à la fois vraie et fausse. On peut voir qu'alors toute assertion de la théorie est à la fois vraie et fausse. Une telle théorie est évidemment sans intérêt. Tout le monde est persuadé (bien que cela n'ait jamais été prouvé) que l'Arithmétique usuelle ainsi que les diverses théories modernes des ensembles sont *non contradictoires*.

⁽¹⁾ *Grand théorème de Fermat* : Conjecture formulée par le mathématicien français Pierre Simon de Fermat (1601-1665), selon laquelle pour x, y, z et n dans \mathbb{N}^* avec $n \geq 3$ on ne peut avoir $x^n + y^n = z^n$.

⁽²⁾ *Conjecture de Goldbach* : Conjecture selon laquelle tout nombre pair > 2 est la somme de deux nombres premiers. Énoncée en 1742 par Christian Goldbach (1690-1764), publiée en 1770 dans les *Méditationes Algebraicae* de Waring, cette assertion n'a à l'heure actuelle été ni complètement prouvée, ni réfutée, contrairement à la conjecture de Waring qui vient d'être démontrée en octobre 1985.

⁽³⁾ *Conjecture de Waring* : Dans le livre cité dans la note (2) Edward Waring (1734-1798) affirme sans preuve que tout entier positif est la somme de 4 carrés (au plus), de 9 cubes, de 19 bicarrés et ainsi de suite. C'est l'existence de la décomposition en somme de 19 bicarrés qui a résisté le plus longtemps aux efforts des mathématiciens.

⁽⁴⁾ *Kurt Gödel*, mathématicien et logicien autrichien né à Brünn en 1906

Cependant, dans une théorie non contradictoire, il peut exister des assertions A telles que ni A ni $(\text{non } A)$, ne soient insérables dans un texte démonstratif *interne à la théorie* : de telles assertions sont dites *indécidables*. Si A est indécidable, on peut à volonté rajouter A (resp. $(\text{non } A)$) aux axiomes de la théorie, et on obtient ainsi deux nouvelles théories non contradictoires. Dans les systèmes usuels d'axiomes de la théorie des ensembles on a pu prouver que l'*hypothèse du continu* (voir § II.3) est indécidable.

- Si A et B sont deux assertions, on forme une nouvelle assertion appelée *disjonction de A et B* , et notée $(A \text{ ou } B)$.

Règle 2 : $(A \text{ ou } B)$ est vraie si l'une au moins des deux assertions A , B est vraie et $(A \text{ ou } (\text{non } A))$ est toujours vraie (même si A est indécidable).

Remarque 1 : Il se peut que l'on sache que $(A \text{ ou } B)$ est vraie sans être capable de prouver A ni de prouver B . Soit par exemple l'assertion A : « $e + \pi$ est transcendant » et l'assertion B : « $e\pi$ est transcendant ».

Alors il est sûr que $(A \text{ ou } B)$ est vraie, mais on ne sait, à l'heure actuelle, prouver ni A , ni B .

- Si A et B sont deux assertions, la disjonction $((\text{non } A) \text{ ou } B)$ s'appelle *implication de B par A* et se note : $A \Rightarrow B$.

La règle 2 montre immédiatement que, pour toute assertion A , l'assertion $A \Rightarrow A$ est vraie.

Règle 3 : Si $(\text{non } A)$ est vraie, $A \Rightarrow B$ est vraie pour toute B .

Règle 4 : Si A est vraie et si $A \Rightarrow B$ est vraie, alors B est vraie.

Il ne faut surtout pas confondre la vérité de $A \Rightarrow B$ avec celle de B ! La règle 4 est appelée *principe du syllogisme*.

- Si A et B sont deux assertions, l'assertion $(\text{non } ((\text{non } A) \text{ ou } (\text{non } B)))$ s'appelle *conjonction de A et B* et se note $(A \text{ et } B)$.

Règle 5 : $(A \text{ et } B)$ est vraie ssi les deux assertions A et B sont vraies.

DÉFINITION I.1.1

$\left\{ \begin{array}{l} \text{Deux assertions } A, B \text{ sont dites } \mathbf{\textit{équivalentes}} \text{ ssi les deux implica-} \\ \text{tions } A \Rightarrow B \text{ et } B \Rightarrow A \text{ sont vraies. On note alors : } A \Leftrightarrow B. \end{array} \right.$

Exemple 1 : $A \Leftrightarrow (\text{non } (\text{non } A))$.

Exemple 2 : $A \Rightarrow B$ est équivalente à $(\text{non } B) \Rightarrow (\text{non } A)$.

Application de l'exemple 2 : pour démontrer $A \Rightarrow B$, on choisit, parmi les deux assertions $A \Rightarrow B$ et $(\text{non } B) \Rightarrow (\text{non } A)$, celle qui est la plus aisée à établir.

Tout ce qui précède pourrait se développer en un véritable « calcul propositionnel », que nous n'étudierons pas ici. Bornons-nous à signaler :

Règle 6 : Si $A \Rightarrow B$ et $B \Rightarrow C$ sont vraies, alors $A \Rightarrow C$ est vraie.

C'est la *transitivité de l'implication*, qui peut s'écrire aussi :

$$((A \Rightarrow B) \text{ et } (B \Rightarrow C)) \Rightarrow (A \Rightarrow C).$$

Règle 7 : Soit A_1, A_2, \dots, A_n et B des assertions ; si $(A_1 \text{ ou } A_2 \text{ ou } \dots \text{ ou } A_n)$ et $(A_1 \Rightarrow B), (A_2 \Rightarrow B), \dots, (A_n \Rightarrow B)$ sont vraies, alors B est vraie. (C'est la « disjonction des cas ».)

On comprendra la portée de cette règle 7 en revenant à la remarque 1 qui suit la règle 2.

Exemple 3 : assertion A : $(\sqrt{2})^{\sqrt{2}}$ est un nombre rationnel

assertion B : $(\sqrt{2})^{\sqrt{2}}$ est irrationnel

assertion C : Il existe un réel u non rationnel tel que $u^{\sqrt{2}}$ soit rationnel.

Alors $(A \text{ ou } B)$ est vraie et $A \Rightarrow C, B \Rightarrow C$ sont évidemment vraies puisqu'on sait que $\sqrt{2}$ est irrationnel (la deuxième implication résulte de l'égalité $[(\sqrt{2})^{\sqrt{2}}]^{\sqrt{2}} = (\sqrt{2})^2 = 2$). Donc C est vraie sans avoir besoin de savoir si c'est l'assertion A qui est vraie ou l'assertion B .

Règle 8 : Soit A une assertion d'une théorie ; rajoutons l'axiome $(\text{non } A)$ à la théorie considérée, et supposons trouvée dans la théorie ainsi obtenue une assertion B telle que $(\text{non } A) \Rightarrow B$ et $(\text{non } A) \Rightarrow (\text{non } B)$ soient vraies dans cette nouvelle théorie ; alors A est vraie dans la théorie de départ.

C'est le principe du « raisonnement par l'absurde ». Dans la pratique on dit qu'on *a*, à partir de $(\text{non } A)$, abouti à une contradiction.

Exemple 4 : Supposons que B soit une assertion vraie dans la théorie de départ, et que $(\text{non } A) \Rightarrow (\text{non } B)$. Alors (cf. l'exemple 2 suivant Déf. I.1.1) $B \Rightarrow A$ est vraie, donc A est vraie puisque B est vraie.

Bien souvent, un *raisonnement par l'absurde* se présente sous la forme qui vient d'être décrite dans cet exemple 4. Mais ce n'est pas toujours aussi simple, comme le montre cet exemple :

Exemple 5 : Soit à prouver l'assertion A : « l'ensemble \mathcal{P} des nombres premiers est infini ».

Assertion $(\text{non } A)$: « l'ensemble \mathcal{P} est fini ».

Rajoutons $(\text{non } A)$ aux axiomes de Peano ⁽¹⁾ qui sont les fondements de

⁽¹⁾ *Giuseppe Peano* : Logicien et mathématicien italien (1858-1932). C'est à lui qu'on doit les symboles usuels : \in, \cup, \cap, \subset (cf. § I.2) et les axiomes de l'Arithmétique (

l'Arithmétique. Nous savons que l'ensemble \mathcal{P} est non vide (par exemple 2 est premier). Soit alors N le plus grand élément de \mathcal{P} (N existe car (non A) est vraie et qu'on démontre en Arithmétique que tout ensemble fini non vide admet un plus grand élément). Le nombre $2 \times 3 \times 4 \times \cdots \times N + 1$ noté en abrégé $N! + 1$ n'est divisible par aucun nombre premier $\leq N$ (il reste 1). Soit alors q le plus petit facteur premier de $N! + 1$. Alors $q \in \mathcal{P}$ et $q > N$. Donc dans la nouvelle théorie, l'assertion B : « N est le plus grand élément de \mathcal{P} » est à la fois vraie et fausse ; donc A est vraie par application de la règle 8. C'est en substance le raisonnement qui figure déjà dans le IX^e livre des Eléments d'Euclide ⁽¹⁾ (proposition 20).

Exemple 6 : Soit à prouver l'assertion A : « le nombre e (base des logarithmes népériens) est irrationnel ».

Assertion (non A) : « le nombre e est rationnel », autrement dit il existe deux naturels p et q tels que $e = \frac{p}{q}$ ($q \geq 1$).

Rajoutons (non A) aux axiomes de l'Arithmétique et supposons construit le corps des réels. On démontre dans le cours d'Analyse que pour tout entier $n > 0$ le nombre e vérifie la double inégalité stricte

$$1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} < e < 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{nn!}.$$

Soit alors l'assertion B : « Il n'existe aucun entier qui soit compris strictement entre un naturel N et son successeur $N + 1$ ». Dans la nouvelle théorie B est à la fois vraie (résultat du début de l'Arithmétique) et fausse (car si $e = \frac{p}{q}$, en multipliant les trois membres par $q!$ on obtient pour $n = q$

$$N < p(q-1)! < N + \frac{1}{q} \leq N + 1, \text{ avec } N \in \mathbb{N}^*).$$

Donc A est vraie par application de la règle 8.

Quantificateurs

Soit $\mathcal{R}(x)$ une assertion dépendant de l'objet « variable » x ; on écrit $\exists x \mid \mathcal{R}(x)$ pour exprimer qu'il existe au moins un des objets x pour lequel $\mathcal{R}(x)$ est vraie.

Lorsque $\exists x \mid \mathcal{R}(x)$ est vraie, il ne faudrait pas croire qu'on soit toujours capable de *construire* l'un des objets x qui rendent vraie $\mathcal{R}(x)$, à l'aide d'algorithmes, ni même d'une suite de constructions autorisées dans la théorie considérée.

⁽¹⁾ Euclide : Le plus célèbre des mathématiciens de l'Antiquité ; a vécu en Grèce au III^e siècle avant notre ère. Auteur des « Eléments » qui passent depuis deux millénaires comme un modèle de rigueur difficile à dépasser.

Exemple 7 : L'assertion « il existe une famille $(a_i)_{i \in I}$ de réels qui est une base du \mathbb{Q} -espace vectoriel \mathbb{R} » est vraie, mais personne ne connaît une construction effective d'une telle famille.

La valeur des assertions vraies du type $(\exists x \mid \mathcal{R}(x))$ pour lesquelles on ne connaît pas de construction effective d'une « solution » x a donné lieu, naguère, à de vives polémiques entre mathématiciens. Aujourd'hui, ces polémiques sont bien estompées, les mathématiciens considérant ces assertions comme « aussi vraies » que les autres ⁽¹⁾.

Le symbole \exists de la relation $(\exists x \mid \mathcal{P}(x))$ est appelé *quantificateur existentiel*.

L'assertion $(\forall x, \mathcal{R}(x))$ est par définition l'assertion : $(\text{non } (\exists x \mid (\text{non } \mathcal{R}(x))))$. Le symbole \forall de cette assertion s'appelle *quantificateur universel* ; $(\forall x, \mathcal{R}(x))$ se lit :

« quel que soit x , $\mathcal{R}(x)$ » ou encore : « pour tout x , $\mathcal{R}(x)$ ».

Il faut faire très attention dans l'utilisation des quantificateurs et des règles de logique 1 à 8 ci-dessus. De manière générale, *éviter de les mêler à un texte du langage courant, et apporter le plus grand soin à leur ordre et à leur parenthésage.*

Exemple 8 : Soit f une fonction de \mathbb{R} dans \mathbb{R} . L'écriture

$$(1) \quad \forall \varepsilon > 0 \quad \exists \alpha > 0, \quad (\forall x \in \mathbb{R}) (|x - x_0| \leq \alpha \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$$

exprime la *continuité* de f en x_0 . L'écriture

$$(2) \quad \exists \alpha > 0, \quad \forall \varepsilon > 0 \quad (\forall x \in \mathbb{R}) (|x - x_0| \leq \alpha \Rightarrow |f(x) - f(x_0)| \leq \varepsilon)$$

exprime que f est *constante* au voisinage de x_0 , et pourtant entre (1) et (2) la seule différence est l'interversion de « $\forall \varepsilon > 0$ » et de « $\exists \alpha > 0$ » !

Exemple 9 : Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions de \mathbb{R} dans \mathbb{R} . L'écriture

$$(1) \quad \forall \varepsilon > 0 \quad \exists N \mid \forall n \geq N \quad \forall x \in X \quad |f_n(x) - f(x)| \leq \varepsilon$$

exprime la convergence uniforme sur X de la suite $(f_n)_{n \in \mathbb{N}}$ vers la fonction f , tandis que l'écriture

$$(2) \quad \forall x \in X \quad \forall \varepsilon > 0 \quad \exists N \mid \forall n \geq N \quad |f_n(x) - f(x)| \leq \varepsilon$$

exprime la convergence simple de la suite $(f_n)_{n \in \mathbb{N}}$ vers la fonction f , et le cours d'Analyse montrera que ce sont deux modes de convergence très différents.

⁽¹⁾ Cette attitude se justifie d'autant plus qu'on a trouvé des constructions de solutions x à des relations vraies du type $(\exists x \mid \mathcal{R}(x))$ pour lesquelles on pensait ne jamais pouvoir « exhiber » de telles solutions ; par exemple, existence de mesures de Haar sur des groupes localement compacts.

Exercice 1 : L'implication $(\exists x \mid \forall y, A(x, y)) \Rightarrow (\forall y, (\exists x \mid A(x, y)))$ est vraie. Etudier l'implication *réciproque*.

Indication : On montrera à l'aide d'un exemple que l'implication *réciproque* est fausse.

Exercice 2 : Montrer, par disjonction des cas, que l'équation $5x^3 + 11y^3 + 13z^3 = 0$ n'admet dans \mathbb{Z}^3 que la solution $(0, 0, 0)$.

Exercice 3 : Montrer que : $(A \Rightarrow B) \Rightarrow ((C \Rightarrow A) \Rightarrow (C \Rightarrow B))$.

Exercice 4 : Montrer que : $((A \Leftrightarrow B) \text{ et } (B \Leftrightarrow C)) \Rightarrow (A \Leftrightarrow C)$; que : $(A \Leftrightarrow B) \Leftrightarrow (B \Leftrightarrow A)$ et que : $A \Leftrightarrow A$.

Exercice 5 : Vérifier que :

$$((A_1 \Rightarrow A_2) \text{ et } (A_2 \Rightarrow A_3) \text{ et } \dots \text{ et } (A_{n-1} \Rightarrow A_n) \text{ et } (A_n \Rightarrow A_1)) \\ \Rightarrow ((\forall i, j), (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket) (A_i \Leftrightarrow A_j) .$$

Exercice 6 : Démontrer par l'absurde (règle 8, exemple 4) l'unicité du quotient et du reste de la division euclidienne de l'entier a par l'entier $b > 0$.

Exercice 7 : Montrer qu'il existe une infinité de nombres premiers de la forme $4n - 1$ (resp. $6n - 1$).

Exercice 8 : Expliquer pourquoi la règle : « toute règle a des exceptions » se contredit elle-même (paradoxe du menteur).

§ I.2 CONSTRUCTION D'ENSEMBLES

Ci-après, nous considérerons comme acquise une notion « intuitive » d'ensemble. Lorsque nous ferons allusion à « la Théorie des ensembles », il s'agira de l'une des théories cohérentes des ensembles obtenue avec l'un des systèmes d'axiomes qui, depuis des lustres, ont fait l'accord des mathématiciens, par exemple la théorie des ensembles de *Zermelo-Fraenkel avec axiome de fondation et avec axiome du choix* (voir [23]).

Sur la collection des êtres mathématiques d'une théorie donnée, on introduit une relation, dite *appartenance*, notée \in ; c'est une relation binaire. Si x et y sont deux « objets » mathématiques, la relation $x \in y$ se lit « x appartient à y », ou « x est élément de y ». Sa négation s'écrit $x \notin y$ et se lit « x n'appartient pas à y ».

Soit E un objet donné, que nous appellerons *ensemble*, lorsqu'on le considère dans son rapport avec la relation \in . Les objets x tels que $x \in E$ s'appellent les *éléments* de E . Intuitivement, l'ensemble E est formé des objets x pour lesquels on a : $x \in E$.

Ensemble formé avec des éléments donnés

Considérons des objets, distincts ou non, de la théorie des ensembles, que nous notons a_1, a_2, \dots, a_n , ces objets étant supposés *en nombre fini* n ;

Règle 1 : Il existe un ensemble E , et un seul, dont les éléments sont exactement les objets a_1, a_2, \dots, a_n .

Cet ensemble sera noté $\{a_1, a_2, \dots, a_n\}$. En particulier, si a est un objet, l'ensemble $\{a\}$ est appelé le *singleton* d'élément a .

Inclusion

DÉFINITION I.2.1

$\left\{ \begin{array}{l} \text{On dit que l'ensemble } E \text{ est } \mathbf{inclus} \text{ dans l'ensemble } F \text{ ssi on a :} \\ (\forall x) (x \in E) \Rightarrow (x \in F). \\ \text{Lorsqu'il en est ainsi, on écrit } E \subset F \text{ et on dit que } \mathbf{E est une partie de} \\ F. \text{ Sinon, on écrit } E \not\subset F. \end{array} \right.$

PROPOSITION I.2.1 : (transitivité de l'inclusion)

$\left\| \begin{array}{l} \text{Si } E, F \text{ et } G \text{ sont des ensembles, les relations } (E \subset F) \text{ et } (F \subset G) \\ \text{entraînent : } (E \subset G). \end{array} \right.$

C'est une conséquence de la transitivité de l'implication (règle 6 du § I.1).

DÉFINITION I.2.2

$\left\{ \begin{array}{l} \text{Deux ensembles } E \text{ et } F \text{ sont dits } \mathbf{égaux} \text{ ssi on a : } E \subset F \text{ et} \\ F \subset E. \text{ Si c'est le cas on écrit } E = F. \end{array} \right.$

Les ensembles E et F sont donc égaux ssi, pour tout x , les relations « $x \in E$ » et « $x \in F$ » sont équivalentes. Intuitivement, cela signifie que E et F sont formés exactement des mêmes éléments.

La plupart des ensembles étudiés en Mathématiques sont construits à partir d'un petit nombre d'ensembles « de base », à l'aide de *règles de construction d'ensembles*, dont certaines seront décrites dans ce qui suit. Voici une liste de ces ensembles fondamentaux :

- \emptyset : ensemble vide (voir ci-dessous la définition I.2.4)
- \mathbb{N} : ensemble des *entiers naturels* (voir Chap. II)
- \mathbb{N}^* : ensemble des entiers naturels ≥ 1
- \mathbb{Z} : ensemble de tous les *entiers relatifs* (ou *entiers rationnels*)
- \mathbb{Q} : ensemble des *nombre rationnels*
- \mathbb{Q}^* : ensemble des nombres rationnels non nuls
- \mathbb{Q}_+ : ensemble des nombres rationnels positifs ou nuls
- \mathbb{Q}_+^* : ensemble des nombres rationnels *strictement positifs*
- \mathbb{R} : ensemble des *nombres réels*
- \mathbb{R}^* : ensemble des nombres réels non nuls
- \mathbb{R}_+ : ensemble des nombres réels positifs ou nuls
- \mathbb{R}_- : ensemble des nombres réels négatifs ou nuls
- \mathbb{R}_+^* : ensemble des nombres réels *strictement positifs*
- \mathbb{R}_-^* : ensemble des nombres réels *strictement négatifs*
- \mathbb{C} : ensemble des *nombres complexes*
- \mathbb{C}^* : ensemble des nombres complexes *non nuls*

Partie d'un ensemble définie par une relation

Soit E un ensemble et $A(x)$ une assertion de la théorie des ensembles dépendant de la « variable » x ; (x représente un objet indéterminé de la théorie des ensembles). Nous admettrons la

Règle 2 : Sous ces hypothèses, il existe un et un seul ensemble F tel que les relations « $x \in F$ » et « $x \in E$ et $A(x)$ » soient équivalentes pour tout x , ce qu'on écrit : $(\forall x) (x \in F) \Leftrightarrow (x \in E \text{ et } A(x))$.

L'ensemble F ainsi défini est évidemment *une partie de E* . Nous dirons que F est l'ensemble des $x \in E$ qui possèdent (ou : « qui vérifient ») la propriété $A(x)$, et nous écrirons :

$$F = \{x \mid x \in E \text{ et } A(x)\} \quad (\text{ou aussi : } F = \{x \in E \mid A(x)\}).$$

Exemple 1 : $E = \mathbb{N}^*$, ensemble des entiers naturels non nuls

x = une « variable » dans \mathbb{N}^*

$A(x)$ est l'assertion : « $x \neq 1$ et les seuls diviseurs dans \mathbb{N} de x sont 1 et x ».

Alors l'ensemble $\mathcal{P} = \{x \mid x \in \mathbb{N}^* \text{ et } A(x)\}$ est une partie de \mathbb{N}^* , appelée *ensemble des nombres premiers*.

Exemple 2 : E est un ensemble arbitraire

F est une partie de E

x est une variable

$A(x)$ est l'assertion : « $x \notin F$ »

Alors l'ensemble $\{x \mid x \in E \text{ et } x \notin F\}$ est une partie de E .

DÉFINITION I.2.3

⎵ Avec les notations de l'exemple 2, l'ensemble $\{x \mid x \in E \text{ et } x \notin F\}$
⎵ est appelé le **complémentaire de F dans E** et noté $\complement_E F$ ou $E \setminus F$.

Il est clair que $\complement_E(\complement_E F) = F$, et que, si F et G sont deux parties de E , on a l'équivalence : $(F \subset G) \Leftrightarrow (\complement_E G \subset \complement_E F)$.

THÉORÈME I.2.1

|| Il existe un, et un seul, ensemble V tel que $(\forall x) x \notin V$.

Démonstration :

On s'assure d'abord de l'existence d'au moins un ensemble V . Soit E un ensemble, alors $V = \complement_E E$ vérifie bien : $(\forall x) x \notin V$, car à cause des définitions, $x \in V$ signifie : $x \in E$ et $x \notin E$, de sorte qu'il n'existe aucun objet x tel que $x \in V$.

Soit alors F un ensemble ; on a : $(\forall x) (x \in V) \Rightarrow (x \in F)$ puisque la relation « $x \in V$ » est toujours fausse (cf. la définition de V).

Autrement dit, $V \subset F$. Supposons alors trouvés deux ensembles V_1 et V_2 tels que $(\forall x) x \notin V_1$ et $x \notin V_2$; le raisonnement ci-dessus prouve que $V_1 \subset V_2$ et $V_2 \subset V_1$, d'où l'unicité : $V_1 = V_2$. ■

DÉFINITION I.2.4

$\left\{ \begin{array}{l} \text{L'unique ensemble } V \text{ tel que } (\forall x) x \notin V \text{ s'appelle } \textbf{ensemble vide}, \\ \text{et se note } \emptyset. \end{array} \right.$

La preuve du théorème I.2.1 montre que, pour tout ensemble F , on a : $\emptyset \subset F$. Elle montre en fait que \emptyset est le seul ensemble qui soit inclus dans tous les autres.

Remarque 1 : La règle 2 est plus « audacieuse » qu'elle ne semble au premier abord : elle permet de définir des ensembles qu'on ne peut définir de manière « plus concrète ». Par exemple cette règle permet de concevoir des parties de \mathbb{N} qu'il est impossible de « construire par application récursive d'algorithmes ». Ces questions sont du plus haut intérêt en logique et en informatique. Cf. [10].

Remarque 2 : Si l'on donne une relation $A(x)$ dépendant de l'objet indéterminé x de la théorie des ensembles, en général il n'existe pas d'ensemble E tel que

$$(1) \quad (\forall x) \quad (x \in E) \Leftrightarrow A(x).$$

On donne le nom de *relations collectivisantes* (sous-entendu : en x) aux relations $A(x)$ pour lesquelles il existe un ensemble E vérifiant (1). (Un tel ensemble E est alors unique.) Les autres relations sont dites *non collectivisantes*.

Exemple 3 : La relation $(\forall x) x \notin X$ est *collectivisante en X* ; l'ensemble des X qui la vérifient est réduit à un seul élément, qui est \emptyset . C'est donc le singleton $\{\emptyset\}$.

Exemple 4 : Si l'ensemble E est donné et si $B(x)$ est une relation dépendant de l'objet x , la relation $A(x) = \langle x \in E \text{ et } B(x) \rangle$ est collectivisante en x , et définit l'ensemble $\{x \mid x \in E \text{ et } B(x)\}$, ce qui est une autre façon d'énoncer la règle 2.

Exemple 5 : La relation $x \notin x$ n'est pas collectivisante en x , car un ensemble Ω tel que $(x \in \Omega) \Leftrightarrow (x \notin x)$ vérifierait à la fois $\Omega \in \Omega$ et $\Omega \notin \Omega$.

Remarque 3 : Dans les théories modernes des ensembles on introduit un *axiome de fondation* (cf. [23]), qui entraîne notamment le *théorème* : $(\forall x) (x \notin x)$.

Dans une théorie des ensembles avec axiome de fondation,

relation $x \in x$ n'est pas collectivisante équivaut au fait — maintenant bien accepté — qu'il n'existe aucun ensemble dont tout ensemble soit élément. Dans une telle théorie également, il n'existe aucune *suite infinie* $(u_n)_{n \in \mathbb{N}}$ telle que, pour tout n , on ait $u_{n+1} \in u_n$.

Ensemble des parties

Règle 3 : Soit E un ensemble ; il existe un, et un seul, ensemble \mathcal{F} tel que $(\forall X) (X \in \mathcal{F}) \Leftrightarrow (X \subset E)$.

DÉFINITION I.2.5

$\left\{ \begin{array}{l} \text{L'ensemble } \mathcal{F} \text{ de la règle 3 s'appelle } \mathbf{ensemble des parties} \\ \text{de } E \text{ et se note } \mathcal{P}(E). \end{array} \right.$

La règle 3 peut s'énoncer ainsi : *pour tout ensemble E , la relation « $X \subset E$ » est collectivisante en X .*

On a toujours $\emptyset \subset E$, et $E \subset E$, c'est-à-dire : $\emptyset \in \mathcal{P}(E)$ et $E \in \mathcal{P}(E)$. Donc $\mathcal{P}(E)$ est toujours un ensemble non vide. A partir d'un ensemble E (par exemple à partir de \emptyset), on peut former les ensembles itérés de parties de E :

$$\begin{aligned} \mathcal{P}_{(0)}(E) &= E; & \mathcal{P}_{(1)}(E) &= \mathcal{P}(E); \\ \mathcal{P}_{(2)}(E) &= \mathcal{P}(\mathcal{P}(E)), \dots, & \mathcal{P}_{(n)}(E) &= \mathcal{P}(\mathcal{P}_{(n-1)}(E)) \end{aligned}$$

pour $n \geq 1$. Alors, pour tout entier $n \geq 1$, on a : $\mathcal{P}_{(n-1)}(E) \in \mathcal{P}_{(n)}(E)$.

Produit cartésien

Soit p un entier naturel. Nous admettrons qu'à la donnée de tout système a_1, a_2, \dots, a_p de p objets pris dans cet ordre, est associé un élément, noté (a_1, a_2, \dots, a_p) , de sorte que la règle suivante soit satisfaite :

$$\begin{aligned} &(\forall a_1) (\forall a_2) \dots (\forall a_p) (\forall b_1) \dots (\forall b_p) \\ &(a_1, a_2, \dots, a_p) = (b_1, b_2, \dots, b_p) \Leftrightarrow (a_1 = b_1 \text{ et } a_2 = b_2 \text{ et } \dots \text{ et } a_p = b_p). \end{aligned}$$

L'élément (a_1, a_2, \dots, a_p) est appelé le **p -uple** (a_1, a_2, \dots, a_p) ; pour $p = 2$, on parle de *couple* ; pour $p = 3$, de *triplet* ; pour $p = 4$, de *quadruplet*, etc. Si $a = (a_1, a_2, \dots, a_p)$, pour tout entier i ($1 \leq i \leq p$), a_i est appelé la *i -ième projection* de a ; on notera : $a_i = \text{pr}_i(a)$; soit $a = (a_1, a_2, \dots, a_p)$ un p -uple avec $p \geq 3$ et soit q un entier de \mathbb{N}^* , alors on convient d'identifier les objets $((a_1, a_2, \dots, a_q), a_{q+1}, \dots, a_p)$ et (a_1, a_2, \dots, a_p) . Plus généralement nous identifierons le p -uple (a_1, a_2, \dots, a_p) avec tout n -uple obtenu à partir de celui-là par parenthésage arbitraire avec conservation de l'ordre. On se gardera de confondre l'objet $\{a_1, a_2, \dots, a_p\}$ avec l'objet (a_1, a_2, \dots, a_p) . Si on réécrit la liste a_1, a_2, \dots, a_p dans un ai

obtiendra un nouveau p -uplet (b_1, b_2, \dots, b_p) en général $\neq (a_1, a_2, \dots, a_p)$ tandis que les ensembles $\{b_1, b_2, \dots, b_p\}$ et $\{a_1, a_2, \dots, a_p\}$ sont égaux puisque composés des mêmes éléments. Si, par exemple, a et b sont deux objets, le couple (a, b) peut être identifié à l'ensemble $\{a, \{a, b\}\}$, qui évidemment n'a rien à voir avec $\{a, b\}$.

Règle 4 : Soit p un entier ≥ 1 et E_1, E_2, \dots, E_p des ensembles ; il existe un et un seul ensemble F possédant la propriété suivante :

$$(\forall z) \quad (z \in F) \Leftrightarrow (\exists x_1 \in E_1, \exists x_2 \in E_2, \dots, \exists x_p \in E_p \mid z = (x_1, x_2, \dots, x_p)).$$

L'ensemble F est appelé le *produit cartésien* de E_1, E_2, \dots, E_p ; on le note $E_1 \times E_2 \times \dots \times E_p$.

Par construction, $E_1 \times E_2 \times \dots \times E_p = \emptyset$ si l'un au moins des E_i est vide.

Pour tout parenthésage de E_1, E_2, \dots, E_p écrits dans cet ordre, le produit cartésien $E_1 \times E_2 \times \dots \times E_p$ sera identifié au produit cartésien des ensembles obtenus par produit cartésien des termes de chaque parenthèse.

Si $E_1 = E_2 = \dots = E_p = E$, on notera en abrégé E^p le produit cartésien $E_1 \times E_2 \times \dots \times E_p$; dans ce cas l'ensemble Δ des p -uplets (x, x, \dots, x) est appelé *diagonale* de E^p . Le lecteur pourra contrôler que Δ est bien défini comme ensemble.

Intersection et réunion de deux ensembles

Règle 5 : Etant donnés deux ensembles E et F , il existe un, et un seul, ensemble G tel que

$$(\forall x) \quad (x \in G) \Leftrightarrow (x \in E \text{ ou } x \in F).$$

DÉFINITION I.2.6

⎵ Cet ensemble G est appelé **réunion** (ou **union**) des ensembles E et F ,
⎵ et on le note $E \cup F$.

Pour tous ensembles E, F, G , on a :

$$E \cup F = F \cup E ; \emptyset \cup E = E ; (E \cup F) \cup G = E \cup (F \cup G).$$

L'associativité de la réunion permet de noter l'ensemble $(E \cup F) \cup G$ sans parenthésage : $E \cup F \cup G$.

Il est clair que $(E \subset F) \Leftrightarrow (E \cup F = F)$.

PROPOSITION I.2.2

⎵ Etant donnés deux ensembles E et F , il existe un ensemble H unique
⎵ tel que

$$(\forall x) \quad (x \in H) \Leftrightarrow (x \in E \text{ et } x \in F).$$

C'est une application immédiate de la règle 2.

DÉFINITION I.2.7

⎧ Cet ensemble H est appelé **intersection** de E et F ; on le note
 ⎨ $E \cap F$.

Pour tous ensembles E, F, G , on a :

$$E \cap F = F \cap E ; \quad \emptyset \cap E = \emptyset ; \quad (E \cap F) \cap G = E \cap (F \cap G) ,$$

noté $E \cap F \cap G$.

On dit que deux ensembles E et F sont **disjoints** ssi $E \cap F = \emptyset$. On dit que E et F **se rencontrent** pour exprimer que $E \cap F \neq \emptyset$. Il est clair que $(E \subset F) \Leftrightarrow E \cap F = E$.

L'intersection est *distributive par rapport à la réunion* :

$$E \cap (F \cup G) = (E \cap F) \cup (E \cap G) .$$

La réunion est aussi *distributive par rapport à l'intersection* :

$$E \cup (F \cap G) = (E \cup F) \cap (E \cup G) .$$

Si Ω est un ensemble, et si E et F sont des parties de Ω , on a :

$$\mathbb{C}_{\Omega}(E \cup F) = (\mathbb{C}_{\Omega} E) \cap (\mathbb{C}_{\Omega} F)$$

$$\mathbb{C}_{\Omega}(E \cap F) = (\mathbb{C}_{\Omega} E) \cup (\mathbb{C}_{\Omega} F)$$

(lois de De Morgan ⁽¹⁾).

DÉFINITION I.2.8

⎧ On appelle **différence** des ensembles E et F pris dans cet ordre
 ⎨ l'ensemble, noté $E \setminus F$, défini par : $E \setminus F = \{x | x \in E \text{ et } x \notin F\}$.

Si $F \subset E$, on a : $E \setminus F = \mathbb{C}_E F$. Pour E et F quelconques, $E \setminus F = \mathbb{C}_E(E \cap F)$.

L'union et l'intersection sont liées au produit cartésien par des relations élémentaires : soit E et F des ensembles ; pour toutes parties A, A' de E (resp. B, B' de F) on vérifie aisément que :

$$A \times B \subset E \times F ; \quad (A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B') ;$$

$$\begin{aligned} (A \times B) \cup (A' \times B') &= \\ &= [(A \cup A') \times (B \cup B')] \setminus [(A \setminus A') \times (B' \setminus B) \cup (A' \setminus A) \times (B \setminus B')] . \end{aligned}$$

Ensemble quotient

Le *passage au quotient* (d'un ensemble par une relation d'équivalence) est l'une des méthodes les plus fécondes de construction d'ensembles, surtout en Algèbre. Nous y consacrerons le § I.5.

⁽¹⁾ Augustus De Morgan : mathématicien et logicien anglais (1806-1871).

Exercice 1 : Un ensemble E quelconque est toujours « élément » d'un autre ensemble (cf. règle 3). Appliquer cette remarque et la règle 1 à la construction des nombres entiers. *Indication :* on pose $0 = \emptyset$, $1 = \{\emptyset\} = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc. Remarquer que dans cette façon d'envisager les entiers naturels la relation d'ordre naturel strict, notée habituellement $<$, est la relation d'appartenance : $(x < y) \Leftrightarrow (x \in y)$.

Exercice 2 : Soit Ω un ensemble ; pour toutes parties A, B de Ω , on définit $A * B = (\complement_{\Omega} A) \cap (\complement_{\Omega} B)$. Montrer que $\complement_{\Omega} A$, $A \cup B$, $A \cap B$ s'expriment en utilisant le seul symbole $*$.

Exercice 3 : Prouver : $(A \setminus B = A) \Leftrightarrow (B \setminus A = B)$.

Exercice 4 : Résoudre l'équation $A \cup X = B$ en l'inconnue X .

Exercice 5 : Démontrer :

$$A \cup B \cup C = (A \setminus B) \cup (B \setminus C) \cup (C \setminus A) \cup (A \cap B \cap C).$$

Exercice 6 : Montrer que pour tout ensemble E , on a $\mathcal{P}(E) \notin E$.

Exercice 7 : Soit A et B deux parties quelconques d'un ensemble E . Est-il vrai ou faux que $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$? Même question pour $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$.

Exercice 8 : On lance 3 pièces ensemble 100 fois de suite et on compte le nombre de « face » sorti : pour la pièce A , face est sorti 70 fois ; pour B , 50 fois ; pour C , 56 fois. Simultanément pour A et B : 31 fois ; simultanément pour B et C : 28 fois. Démontrer que A , B et C ont sorti simultanément « face » au moins 9 fois, et simultanément « pile » au plus 11 fois.

§ I.3 CORRESPONDANCES ET APPLICATIONS

DÉFINITION I.3.1

$\left\{ \begin{array}{l} \text{Un ensemble } G \text{ est appelé un } \mathbf{graphe} \text{ ssi tout élément de } G \text{ est un} \\ \text{couple.} \end{array} \right.$

Si G est un graphe, on définit deux ensembles E et F ainsi :

$$E = \{x \mid \exists z \in G, x = \text{pr}_1(z)\}, \quad F = \{y \mid \exists z \in G, y = \text{pr}_2(z)\}.$$

On écrit : $E = \text{pr}_1(G)$, $F = \text{pr}_2(G)$; E et F sont appelés respectivement *première* et *deuxième projection* de G .

Il est clair que G est une partie de $E \times F$. Réciproquement, si E et F sont des ensembles, toute partie de $E \times F$ peut être considérée comme un graphe : les projections d'un tel graphe sont des parties respectivement de E et F .

DÉFINITION I.3.2

$\left\{ \begin{array}{l} \text{Si } E \text{ et } F \text{ sont des ensembles, on appelle } \mathbf{correspondance} \text{ de } E \text{ vers } F \\ \text{tout triplet } \gamma = (G, E, F) \text{ où } G \text{ est une partie de } E \times F. \text{ Les} \\ \text{ensembles } G, E, F \text{ sont respectivement appelés } \mathbf{graphe}, \mathbf{ensemble de} \\ \mathbf{départ} \text{ et } \mathbf{ensemble d'arrivée} \text{ de la correspondance } \gamma \end{array} \right.$

Si $\gamma = (G, E, F)$ est une correspondance, l'ensemble

$$D_\gamma = \text{pr}_1(G) = \{x \mid x \in E \text{ et } \exists y \in F, (x, y) \in G\}$$

est appelé *ensemble de définition* de la correspondance. Il est clair que $D_\gamma = \emptyset$ ssi $G = \emptyset$. Soit alors A et $B : A \subset E$ et $B \subset F$.

On note

$$\begin{aligned}\gamma(A) &= \{y \mid y \in F \text{ et } \exists x, x \in A \text{ et } (x, y) \in G\} \\ \gamma^{-1}(B) &= \{x \mid x \in E \text{ et } \exists y, y \in B \text{ et } (x, y) \in G\}.\end{aligned}$$

Si $x \in E$ on notera $\gamma(x)$ au lieu de $\gamma(\{x\})$; et si $y \in F$ on abrégera de même $\gamma^{-1}(\{y\})$ en $\gamma^{-1}(y)$.

DÉFINITION I.3.3

*Soit E un ensemble ; on appelle **relation binaire sur E** toute correspondance de E vers E .*

Si $\mathcal{R} = (G, E, E)$ est une relation binaire sur E , nous conviendrons d'écrire $x \mathcal{R} y$ au lieu de « $x \in E$ et $y \in E$ et $(x, y) \in G$ », ce qui nous permettra de parler de « la relation binaire $x \mathcal{R} y$ sur E ».

Exemple 1 : Le graphe de la relation binaire $x = y$ sur E est la *diagonale* de $E \times E$; son domaine de définition est E .

Exemple 2 : Soit E un ensemble. La relation binaire $X \subset Y$ sur l'ensemble $\mathcal{P}(E)$ admet $\mathcal{P}(E)$ pour domaine de définition.

DÉFINITION I.3.4

*Etant donnés deux ensembles E et F , on appelle **graphe fonctionnel** de E vers F tout graphe $G \subset E \times F$ vérifiant la propriété suivante : pour tout $x \in E$, l'ensemble $\{y \mid y \in F \text{ et } (x, y) \in G\}$ est soit vide soit réduit à un élément.*

(Autrement dit, G est « coupé » en au plus un point par toute « parallèle à F ».)

DÉFINITION I.3.5

*Soit E et F deux ensembles ; on appelle **application (ou fonction) de E dans F** toute correspondance de E vers F dont le graphe est fonctionnel et dont le domaine de définition est E .*

Au lieu de : « f est une application de E dans F » on peut écrire symboliquement : $f : E \rightarrow F$, ou encore $E \xrightarrow{f} F$.

Si $f : E \rightarrow F$ est une application, et si $x \in E$, l'ensemble $f(x)$ (c'est-à-dire $f(\{x\})$) est un singleton $\{y\}$, à cause des définitions mêmes

est le *transformé* (ou *l'image*) de x par f , et on note : $y = f(x)$. Pour exprimer que f transforme *chaque* $x \in E$ en $f(x)$, on écrit :

$$f : E \rightarrow F, \quad x \mapsto f(x)$$

Pour toute partie A de E , l'ensemble $f(A)$ coïncide avec $\{f(x)\}_{x \in A}$; on l'appelle *image directe de A par f* ; en particulier, $f(E) = \text{pr}_2(G)$ où G désigne le graphe de f ; $f(E)$ est appelé simplement l'image de f et noté $\text{Im}(f)$. Pour toute partie B de F , l'ensemble $f^{-1}(B)$ est appelé *image réciproque de B par f* .

Il est essentiel de concevoir des fonctions définies par leur graphe, de façon tout à fait arbitraire, même si la transformation $x \mapsto f(x)$ ne permet pas la construction effective de $f(x)$ « par algorithmes » à partir de x . Ce concept moderne (et somme toute très intuitif), de la notion de fonction a permis de grands progrès en mathématiques.

Si $f : E \rightarrow F$ est une application, on a : $f(\emptyset) = \emptyset$; $f^{-1}(\emptyset) = \emptyset$. Si alors A est une partie *non vide* de E , on a $f(A) \neq \emptyset$. Mais attention, même si $B \subset F$ est non vide, il se peut que $f^{-1}(B)$ soit vide !

Si $E = \emptyset$, il existe *une* et une seule application de E dans F : c'est l'application de graphe vide (appelée *application vide*). En revanche si $E \neq \emptyset$ et $F = \emptyset$ il n'existe aucune application de E dans F .

Soit E et F deux ensembles, et G un *graphe fonctionnel* de E vers F : alors $(G, \text{pr}_1(G), F)$ est une application de $D = \text{pr}_1(G)$ dans F , et $D \subset E$; réciproquement, si D est une partie de E et si $f : D \rightarrow F$ est une application, le graphe de f est un graphe fonctionnel de E vers F ; il y a donc équivalence entre *graphe fonctionnel de E vers F* et *application d'une partie de E dans F* .

Pour E et F ensembles donnés, les *applications de E dans F* constituent un ensemble que nous noterons fréquemment $\mathcal{F}(E, F)$, et parfois F^E (notation particulièrement commode quand F est un ensemble fini et qui sera justifiée au § I.4).

L'ensemble $\mathcal{F}(E, F)$ peut être identifié à une partie de $\mathcal{P}(E \times F)$.

Exemple 3 : Soit E une partie d'un ensemble F ; l'application (G, E, F) est appelée *injection canonique de E dans F* quand G est la diagonale de $E \times E$: c'est l'application $j_{E,F} : E \rightarrow F, x \mapsto x$. Si $E = F$, l'application $j_{E,E}$ sera notée Id_E et appelée *application identique de E dans E* .

DÉFINITION I.3.6

Soit E, F, G des ensembles et $f : E \rightarrow F, g : F \rightarrow G$ des applications. On appelle **composée de f et de g** l'application $h : E \rightarrow G, x \mapsto g[f(x)]$. Cette composée est notée $h = g \circ f$ (attention à l'ordre des facteurs !).

La composition d'applications est associative : si E, F, G, H sont des ensembles, et si $f: E \rightarrow F, g: F \rightarrow G, h: G \rightarrow H$ sont des applications, on a : $h \circ (g \circ f) = (h \circ g) \circ f$. Le lecteur se convaincra en effet que les deux membres de cette relation ont un sens et vérifiera qu'ils sont bien égaux ; le résultat commun se note sans parenthésage : $h \circ g \circ f$ comme chaque fois qu'il y a associativité.

Soit $f: E \rightarrow F$ une application, d'image I et de graphe G_f . Alors :

- pour toute partie H de F qui contient I , (G_f, E, H) est une application.
- pour toute partie L de E , $(G_f \cap (L \times F), L, F)$ est une application.

DÉFINITION I.3.7

⎵ Avec les notations qui précèdent, l'application (G_f, E, H) est appelée **corestriction de f à H** ; l'application $(G_f \cap (L \times F), L, F)$ est appelée **restriction de f à L** .

Il faudra toujours se garder de confondre une fonction avec l'une de ses restrictions ou corestrictions sous peine de graves erreurs.

La restriction de f à L sera notée $f|_L$: c'est l'application $L \rightarrow F, x \mapsto f(x)$.

La corestriction de f à H sera notée $f|^H$: c'est l'application $E \rightarrow H, x \mapsto f(x)$.

Si L et H sont des parties respectivement de E et F telles que $f(L) \subset H$, on peut définir l'application $(f|_L)|^H$ qu'on notera pour abréger $f|_L^H$; dans le cas particulier où $E = F$, où $L \subset E$ et $f(L) \subset L$, on écrira $f||_L$ au lieu de $f|_L^L$. On dira que $f||_L$ est **induite** par f sur L .

DÉFINITION I.3.8

⎵ Soit E et F des ensembles, soit L une partie de E et $f: L \rightarrow F$ une application. On appelle **prolongement de f à E** toute application $g: E \rightarrow F$ telle que $g|_L = f$.

Pour f donnée de L dans F , lorsque $F \neq \emptyset$, il existe au moins un prolongement de f à E et en général plusieurs (prendre par exemple un $c \in F$ et considérer $g: E \rightarrow F$ telle que $g(x) = f(x)$ si $x \in L$ et $g(x) = c$ si $x \in E \setminus L$).

Pour $g: E \rightarrow F$ application donnée et pour $L \subset E$, il est clair que g est un prolongement à E de $g|_L$.

Applications injectives, surjectives ou bijectives

DÉFINITION I.3.9

⎵ Soit $f: E \rightarrow F$ une application. On dit que f est **injective** (ou que f est une **injection de E dans F**) ssi

$$(\forall x \in E), (\forall x' \in E) (f(x) = f(x')) \Rightarrow (x = x')$$

ou ce qui revient au même :

$$(\forall x \in E), (\forall x' \in E) \quad (x \neq x') \Rightarrow (f(x) \neq f(x')) .$$

On dit que f est **surjective** (ou que f est une **surjection de E sur F**) ssi

$$(\forall y \in F) \quad (\exists x \in E \mid f(x) = y) .$$

On dit que f est **bijjective** (ou que f est une **bijection de E sur F**) ssi f est à la fois injective et surjective.

Une bijection d'un ensemble E sur lui-même est appelée une **permutation** de E ; l'ensemble des permutations de E se note \mathfrak{S}_E .

Pour tout ensemble E , l'application identique $\text{Id}_E : E \rightarrow E$ est bijective.

Si F est une partie de E , l'injection canonique $j_{F,E} : F \rightarrow E$ est injective.

Si $E = F = \emptyset$, l'application vide est bijective.

Si $f : E \rightarrow F$ est une application d'image I , la corestriction $f|_I$ est surjective.

Si $f : E \rightarrow F$ est une application d'image I , f est surjective ssi $I = F$.

Si f est injective, toute restriction de f l'est encore.

Si E est un singleton, toute application de E dans F est injective.

Notons encore les deux résultats suivants dont la démonstration est évidente :

THÉORÈME I.3.1

|| La composée de deux applications injectives (resp. surjectives, bijectives) est injective (resp. surjective, bijective).

THÉORÈME I.3.2

|| Soit $f : E \rightarrow F$ et $g : F \rightarrow G$ des applications. Si $g \circ f$ est surjective, alors g est surjective. Si $g \circ f$ est injective, alors f est injective.

En voici d'autres un peu moins évidents :

THÉORÈME I.3.3

|| Soit $f : E \rightarrow F$ une bijection, de graphe G ; notons G' la partie de $F \times E$ définie par

$$(\forall x) \quad (\forall y) \quad (y, x) \in G' \Leftrightarrow (x \in E, y \in F \text{ et } (x, y) \in G) .$$

|| Alors $g = (G', F, E)$ est une bijection, et on a :

$$g \circ f = \text{Id}_E, \quad f \circ g = \text{Id}_F .$$

Pour prouver ce théorème I.3.3 on se fonde sur le fait, qui résulte des définitions, que, pour tout $y \in F$, il existe un et un seul élément $x \in E$ tel que $y = f(x)$. Par définition, avec les définitions et hypothèse:

I.3.3, l'application g est appelée **bijection réciproque de f** (ou application réciproque de f), et notée souvent f^{-1} , ou mieux : $f^{<-1>}$.

Si $(x, y) \in E \times F$, les relations $y = f(x)$ et $x = f^{<-1>}(y)$ sont équivalentes.

THÉORÈME I.3.4

|| Soit $f : E \rightarrow F$ une application, avec E et F non vides. Pour que f soit injective (resp. surjective), il faut et il suffit qu'il existe $g : F \rightarrow E$ telle que $g \circ f = \text{Id}_E$ (resp. $f \circ g = \text{Id}_F$).

Démonstration :

D'après le théorème I.3.2, les conditions énoncées ci-dessus sont bien suffisantes (car Id_E et Id_F sont bijectives). Montrons leur nécessité. Supposons f injective. Alors la corestriction $\varphi = f|_{\text{Im}(f)}$ est bijective. Notons $\psi = \varphi^{<-1>}$ la bijection réciproque ; puisque $E \neq \emptyset$, ψ est prolongeable à F ; soit g l'un de ces prolongements, alors $g \circ f = \text{Id}_E$, ce qui achève de prouver la première moitié du théorème.

Supposons maintenant f surjective. Pour chaque $y \in F$, l'ensemble $f^{-1}(y)$ qui est $\{x | x \in E \text{ et } f(x) = y\}$ est non vide. Considérons une partie Z de E telle que, pour tout $y \in F$, l'ensemble $Z \cap f^{-1}(y)$ soit un singleton (l'existence d'une telle partie Z résulte de l'axiome du choix sur lequel nous reviendrons à la fin du § I.3). Pour chaque $y \in F$, notons $g(y)$ l'unique élément de $Z \cap f^{-1}(y)$; Alors on définit ainsi une application $g : F \rightarrow E$ telle que $f \circ g = \text{Id}_F$, ce qui achève de prouver le théorème. ■

Exemple 4 : L'application $\exp : \mathbb{R} \rightarrow \mathbb{R}_+^*$ est bijective. La bijection réciproque est le logarithme népérien : $\text{Log} : \mathbb{R}_+^* \rightarrow \mathbb{R}$ (noté aussi \ln).

Exemple 5 : L'application $\text{tg} : \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[\rightarrow \mathbb{R}$ est bijective. La bijection réciproque : $\mathbb{R} \rightarrow \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ est notée Arc tg (la notation tg^{-1} se rencontre aussi).

Exemple 6 : L'application $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \rightarrow [-1, +1]$ est bijective. La bijection réciproque : $[-1, +1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ est notée Arc sin (ou \sin^{-1}).

THÉORÈME I.3.5

|| Soit $f : E \rightarrow F$ une application.
 || Pour que f soit bijective, il faut et il suffit qu'il existe $g : F \rightarrow E$ telle que $g \circ f = \text{Id}_E$ et $f \circ g = \text{Id}_F$. S'il en est ainsi, g est alors unique, c'est $g = f^{<-1>}$.

Démonstration :

Si g existe, f est bijective (d'après le théorème I.3.4). De plus on a alors

$$(g \circ f) \circ f^{\langle -1 \rangle} = \text{Id}_E \circ f^{\langle -1 \rangle} = f^{\langle -1 \rangle} = g \circ (f \circ f^{\langle -1 \rangle}) = g \circ \text{Id}_F = g ,$$

d'où $g = f^{\langle -1 \rangle}$.

Enfin, si f est bijective, l'application réciproque $g = f^{\langle -1 \rangle}$ satisfait à $g \circ f = \text{Id}_E$ et $f \circ g = \text{Id}_F$ (cf. le Théorème I.3.3). ■

On appelle **involution** d'un ensemble E toute application $f : E \rightarrow E$ telle que $f \circ f = \text{Id}_E$; le théorème I.3.5 montre immédiatement que : *toute involution de E est une bijection de E sur E .*

Exemple 7 : Soit E un ensemble ; l'application $\complement_E : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$, $F \mapsto \complement_E F$ est une involution de $\mathcal{P}(E)$.

A l'aide de l'application \complement_E , toute propriété des parties de E se transforme en une propriété des complémentaires, dite *duale* de la propriété initiale. Par exemple, notant, pour abréger, $X' = \complement_E X$ pour tout $X \in \mathcal{P}(E)$, la propriété $X \subset Y \cap Z$ (X, Y et $Z \in \mathcal{P}(E)$) admet pour propriété duale : $Y' \cup Z' \subset X'$. Toute propriété des parties de E est équivalente à sa propriété duale.

Voici, pour terminer ce rapide tour d'horizon sur les applications, quelques propriétés élémentaires des images directes ou réciproques d'ensembles par des applications, que nous livrons sans démonstration.

Soit $f : E \rightarrow F$ une application, A et B des parties de E . On a :

$$\begin{aligned} (1) \quad & f(A \cup B) = f(A) \cup f(B) \\ (2) \quad & f(A \cap B) \subset f(A) \cap f(B) , \end{aligned}$$

l'inclusion étant *stricte* en général ; cependant si f est *injective*, alors $f(A \cap B) = f(A) \cap f(B)$

$$(3) \quad A \subset f^{-1}(f(A)) .$$

Si A' et B' sont des parties de F , on a :

$$\begin{aligned} (4) \quad & f^{-1}(A' \cap B') = f^{-1}(A') \cap f^{-1}(B') \\ (5) \quad & f^{-1}(A' \cup B') = f^{-1}(A') \cup f^{-1}(B') \\ (6) \quad & f[f^{-1}(A')] \subset A' , \end{aligned}$$

l'inclusion étant *stricte* en général ; mais si f est *surjective*, alors $f[f^{-1}(A')] = A'$

$$(7) \quad f^{-1}(F \setminus A') = E \setminus f^{-1}(A') .$$

L'axiome du choix

Règle : Soit I un ensemble non vide et, pour chaque $i \in I$, soit E_i un ensemble non vide. Alors, si les E_i sont disjoints, il existe au moins un ensemble S tel que, pour tout $i \in I$, $S \cap E_i$ soit un singleton.

Cette règle est l'une des formulations d'un axiome de la théorie des ensembles appelé *axiome du choix*. Le logicien Kurt Gödel a établi, en 1940, que cet axiome n'est pas en contradiction avec les autres axiomes de la théorie usuelle des ensembles. A l'heure actuelle, cet axiome fait partie de toute théorie des ensembles utilisée par la « communauté mathématique » mondiale.

Des formes équivalentes de l'axiome du choix sont le théorème de Zermelo ou le théorème de Zorn (cf. exercice 9 du § I.6).

Les conséquences de l'axiome du choix sont nombreuses et importantes : existence d'idéaux maximaux dans un anneau commutatif ; existence de bases dans un espace vectoriel arbitraire ; existence de la clôture algébrique d'un corps commutatif arbitraire ; en Analyse, théorème de Hahn-Banach, théorème de Krein-Milman, existence d'ultrafiltres,...

Cependant nous ne développerons dans cet ouvrage aucune des conséquences « savantes » de l'axiome du choix. Nous nous contenterons de l'utiliser, sous forme de la règle ci-dessus, chaque fois que cela s'avérera nécessaire, et sans autre commentaire.

Exercice 1 : Si $f : E \rightarrow F$ et $g : F \rightarrow G$ sont des bijections, vérifier que

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

Exercice 2 : Soit $f : E \rightarrow F$, $g : F \rightarrow G$, et $h : G \rightarrow E$ des applications. On considère les trois applications $h \circ g \circ f$, $g \circ f \circ h$, $f \circ h \circ g$. Montrer que si deux d'entre elles sont injectives (resp. surjectives) et la troisième surjective (resp. injective), alors f , g et h sont bijectives.

Exercice 3 : Etudier les correspondances de \mathbb{R} vers \mathbb{R} dont les graphes sont définis par : a) $xy = 1$; b) $x^2 + y^2 \leq 1$; c) $x^3 + y^3 - x - y = 0$; d) $x^y - y^x = 0$.

Exercice 4 : Soit $f : A \rightarrow F$ et $g : B \rightarrow F$ des applications. Peut-on définir une application $h : A \cup B \rightarrow F$ telle que $h|_A = f$ et $h|_B = g$?

Exercice 5 : On considère, sur $E = \left[0, \frac{\pi}{2}\right]$, la relation \mathcal{R} définie par : $x \mathcal{R} y$ ssi $\operatorname{tg}^2 x - \cos^2 y = 0$. Est-elle fonctionnelle ? Symétrique ? Représenter son graphe.

Exercice 6 : Soit f une application de E dans F , $A \subset E$ et $B \subset F$. Montrer que $f(A \cap f^{-1}(B)) = f(A) \cap B$.

Exercice 7 : Soit A et B deux parties non vides d'un ensemble E , et $f : \mathcal{P}(E) \rightarrow \mathcal{P}(A) \times \mathcal{P}(B)$ l'application définie par $f(X) = (A \cap X, B \cap X)$. Etudier les propriétés de f , et dans le cas où f est bijective expliciter $f^{<-1>}$.

Exercice 8 : Soit $\mathcal{P}(E)$ l'ensemble des parties d'un ensemble E , non vide, A et B deux de ces parties, et $f : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ l'application définie par $f(X) = (A \cap X) \cup (B \cap \bar{X})$ où \bar{X} désigne $E \setminus X$. Résoudre et discuter l'équation $f(X) = \emptyset$.

§ I.4 FAMILLES

DÉFINITION I.4.1

$\{$ Soit I un ensemble (que nous appellerons ici **ensemble d'indices**, ou
 $\{$ **ensemble indexateur**). On appelle **famille indexée par I** tout graphe
 $\{$ fonctionnel \mathcal{G} tel que $\text{pr}_1(\mathcal{G}) = I$.

Si I est vide, toute famille indexée par I se réduit à \emptyset .

Si I est non vide, soit \mathcal{G} une famille indexée par I ; par définition, pour tout $i \in I$, il existe un élément *unique* y tel que $(i, y) \in \mathcal{G}$. Nous noterons y_i cet élément et l'appellerons *élément d'indice i* ; la famille \mathcal{G} sera alors notée $(y_i)_{i \in I}$. (Par convention, cette notation sera étendue au cas où $I = \emptyset$.)

Soit donc $\mathcal{G} = (y_i)_{i \in I}$ une famille indexée par I . Si $Y = \text{pr}_2(\mathcal{G})$, l'application $f : I \rightarrow Y, i \mapsto y_i$ admet \mathcal{G} pour graphe. Ainsi nous voyons que \mathcal{G} peut être considérée comme le graphe d'une application. C'est d'ailleurs possible d'une infinité de manières, car \mathcal{G} et aussi le graphe de toute corestriction de f .

Réciproquement, soit E et I deux ensembles, I jouant le rôle d'ensemble indexateur. Soit $f : I \rightarrow E$ une application ; alors le graphe \mathcal{G} de f est une famille indexée par I : c'est la famille $(f(i))_{i \in I}$.

DÉFINITION I.4.2

$\{$ Soit $(y_i)_{i \in I}$ une famille indexée par I , et $f : I \rightarrow Y$ l'application
 $\{$ associée à cette famille, où $Y = \text{pr}_2(\mathcal{G})$, $\mathcal{G} = (y_i)_{i \in I}$. On appelle
 $\{$ **sous-famille** de \mathcal{G} toute famille définie par une des restrictions de f .

Si $J \subset I$, il existe une et une seule sous-famille de $(y_i)_{i \in I}$ indexée par J : c'est la sous-famille $(y_i)_{i \in J}$; nous dirons que c'est la *sous-famille définie par J* .

Si E est un ensemble et si A est une partie de E , on peut considérer A comme une famille d'éléments de E en introduisant l'injection canonique $j_{A,E} : A \rightarrow E$. En effet, si nous associons à A la famille

$$(j_{A,E}(a))_{a \in A} = \mathcal{G}_A = (a)_{a \in A},$$

il est clair que l'application $A \mapsto \mathcal{G}_A$ est *injective*. Dans la suite, on identifiera souvent de la sorte une partie A de E à la famille qu'elle définit ; nous dirons que \mathcal{G}_A est la *famille canoniquement associée à A* (ou encore, que les éléments de A sont *indexés par eux-mêmes* dans \mathcal{G}_A).

Suites

DÉFINITION I.4.3

$\}$ On appelle **suite** toute famille indexée par l'ensemble \mathbb{N} des entiers naturels (ou par une partie de \mathbb{N} , ou même, parfois, par une partie de \mathbb{Z}).
 $\}$ Si E est un ensemble, on appelle **suite d'éléments de E** toute suite (u_n) telle que $(\forall n) u_n \in E$.

Ci-dessous, pour simplifier, nous ne considérerons que des suites indexées par \mathbb{N} . Une suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de E est donc définie, de manière unique, par l'application $\mathbb{N} \rightarrow E$, $n \mapsto u_n$. La notation $(u_n)_{n \in \mathbb{N}}$ d'une telle suite sera dite *indicielle*. Mais si cela s'avère indispensable on aura intérêt à revenir à la notation *fonctionnelle*, en représentant l'application $n \mapsto u_n$ par un symbole, par exemple u , et en écrivant $u_n = u(n)$ pour chaque n .

Quelle que soit la notation utilisée, on se gardera de confondre la *suite* (u_n) , qui est un *graphe*, avec l'*ensemble de ses valeurs* $\{u_n\}_{n \in \mathbb{N}}$, qui est la *deuxième projection* de ce graphe.

Une suite est dite **stationnaire** ssi il existe $n_0 \in \mathbb{N}$ tel que les valeurs de la suite pour $n \geq n_0$ soient toutes égales. On dit aussi que la suite est *constante à partir de n_0* .

DÉFINITION I.4.4

$\}$ On appelle **suite extraite** d'une suite $(u_n)_{n \in \mathbb{N}}$ toute suite $(v_n) = (u_{\varphi(n)})$, où $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est une application strictement croissante.
 $\}$

Il est clair que toute suite extraite d'une suite extraite de la suite (u_n) est extraite de la suite (u_n) : cela se ramène en effet à vérifier que si $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ et $\psi : \mathbb{N} \rightarrow \mathbb{N}$ sont deux applications strictement croissantes (cf. § I.6), alors $\psi \circ \varphi : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante, ce qui est bien clair.

Par définition, une *suite finie* est une suite indexée par une partie finie de \mathbb{N} (ou de \mathbb{Z}).

Réunion, intersection, produit

Considérons une famille d'ensembles $(E_i)_{i \in I}$ indexée par l'ensemble I supposé non vide. Il existe un, et un seul, ensemble F tel qu'on ait l'équivalence : $(\forall x) (x \in F) \Leftrightarrow ((\forall i \in I) x \in E_i)$.

C'est une conséquence de la règle 2 du § I.2.

Nous *admettrons* qu'il existe aussi un ensemble G tel qu'on ait :

$$(\forall x) (x \in G) \Leftrightarrow ((\exists i \in I), x \in E_i).$$

DÉFINITION I.4.5

Avec ces notations, on appelle **intersection des E_i** , et on note $\bigcap_{i \in I} E_i$, l'ensemble $F = \{x \mid \forall i, x \in E_i\}$; on appelle **réunion des E_i** , et on note $\bigcup_{i \in I} E_i$, l'ensemble $G = \{x \mid \exists i, i \in I \text{ et } x \in E_i\}$.

Les relations d'associativité, de commutativité et de distributivité vues au § I.2 s'étendent aux intersections et réunions générales :

• Soit $(E_i)_{i \in I}$ une famille d'ensembles ; supposons que $I = \bigcup_{\lambda \in L} I_\lambda$, où $(I_\lambda)_{\lambda \in L}$ est une famille de parties de I donnée. Alors

$$(1) \quad \bigcup_{i \in I} E_i = \bigcup_{\lambda \in L} \left(\bigcup_{i \in I_\lambda} E_i \right), \quad \text{et} \quad \bigcap_{i \in I} E_i = \bigcap_{\lambda \in L} \left(\bigcap_{i \in I_\lambda} E_i \right)$$

(associativité générale de \cap et \cup).

• Soit $(E_i)_{i \in I}$ une famille d'ensembles, et F un ensemble. Alors

$$(2) \quad \left(\bigcap_{i \in I} E_i \right) \cup F = \bigcap_{i \in I} (E_i \cup F) \quad \text{et}$$

$$(3) \quad \left(\bigcup_{i \in I} E_i \right) \cap F = \bigcup_{i \in I} (E_i \cap F)$$

(distributivité de \cap et \cup chacune par rapport à l'autre).

• En appliquant deux fois (2) et (3) on obtient : si $(E_i)_{i \in I}$ et $(F_j)_{j \in J}$ sont deux familles d'ensembles, on a :

$$(4) \quad \left(\bigcap_{i \in I} E_i \right) \cup \left(\bigcap_{j \in J} F_j \right) = \bigcap_{(i,j) \in I \times J} (E_i \cup F_j) \quad \text{et}$$

$$(5) \quad \left(\bigcup_{i \in I} E_i \right) \cap \left(\bigcup_{j \in J} F_j \right) = \bigcup_{(i,j) \in I \times J} (E_i \cap F_j).$$

Les vérifications sont élémentaires et laissées au lecteur.

DÉFINITION I.4.6

Soit $(E_i)_{i \in I}$ une famille d'ensembles et posons $E = \bigcup_{i \in I} E_i$. On appelle **produit de la famille (E_i)** l'ensemble des familles indexées par I , $(a_i)_{i \in I}$ telles que $(\forall i \in I) a_i \in E_i$. On note cet ensemble produit $\prod_{i \in I} E_i$.

Cette définition a un sens car l'ensemble produit $\prod_{i \in I} E_i$ apparaît comme un sous-ensemble de l'ensemble des applications de I dans E , défini par une relation (cf. la règle 2 du § I.2).

Comme conséquence de l'axiome du choix, on a :

THÉORÈME I.4.1

|| Si l'ensemble I est non vide, et si, pour tout $i \in I$, l'ensemble E_i est non vide, alors l'ensemble produit $\prod_{i \in I} E_i$ est non vide.

On conviendra que si $I = \emptyset$, tout produit d'ensembles indexés par I est vide. Il est clair par ailleurs que $\prod_{i \in I} E_i = \emptyset$ lorsque l'un au moins des E_i est vide, ce qui, compte tenu du théorème I.4.1, peut s'énoncer : si I est non vide, l'ensemble $\prod_{i \in I} E_i$ est vide ssi l'un au moins des E_i est vide. Ci-après, I sera supposé non vide.

Si tous les E_i sont égaux à un même ensemble E , il résulte aussitôt des définitions que $\prod_{i \in I} E_i = \mathcal{F}(I, E)$ (ensemble des applications de I dans E), tout au moins si l'on identifie une famille $(a_i)_{i \in I}$ d'éléments de E avec l'application $I \rightarrow E, i \mapsto a_i$ qu'elle définit. Pour cette raison, il est justifié dans ce cas de noter $\mathcal{F}(I, E) = E^I$.

Si $(E_i)_{i \in I}$ est une famille d'ensembles, et si, pour tout i , F_i est une partie de E_i , on a évidemment $\prod_{i \in I} F_i \subset \prod_{i \in I} E_i$.

Soit $(G_i)_{i \in I}$ une deuxième famille de parties des E_i ($G_i \subset E_i$ pour tout i). Alors :

$$(6) \quad \prod_{i \in I} (F_i \cap G_i) = \left(\prod_{i \in I} F_i \right) \cap \left(\prod_{i \in I} G_i \right).$$

On peut énoncer des règles, plus générales encore que (4) et (5), de distributivité des opérations \cup et \cap chacune par rapport à l'autre :

Soit $(I_\lambda)_{\lambda \in L}$ une famille d'ensembles ; pour chaque $\lambda \in L$, soit $(E_i)_{i \in I_\lambda}$ une famille d'ensembles indexée par I_λ . On suppose L et les I_λ non vides. Alors :

$$(7) \quad \bigcup_{\lambda \in L} \left(\bigcap_{i \in I_\lambda} E_i \right) = \bigcap_{(j_\lambda)_{\lambda \in L} = j \in \prod_{\lambda \in L} I_\lambda} \left(\bigcup_{\lambda \in L} E_{j_\lambda} \right)$$

$$(8) \quad \bigcap_{\lambda \in L} \left(\bigcup_{i \in I_\lambda} E_i \right) = \bigcup_{(j_\lambda)_{\lambda \in L} = j \in \prod_{\lambda \in L} I_\lambda} \left(\bigcap_{\lambda \in L} E_{j_\lambda} \right).$$

L'axiome du choix joue un rôle évident dans la démonstration de ces relations.

Supposons l'ensemble I fini, de cardinal $n \geq 1$, et soit $(E_i)_{i \in I}$ une famille d'ensembles. Alors si $k \mapsto i_k$ est une bijection de $\llbracket 1, n \rrbracket$ sur I , l'application $E_{i_1} \times E_{i_2} \times \cdots \times E_{i_n} \rightarrow \prod_{i \in I} E_i$, $(a_1, a_2, \dots, a_n) \mapsto (b_i)_{i \in I}$, où $b_i = a_k$ si $i = i_k$, est une *bijection*.

Parties d'un ensemble Ω

Soit Ω un ensemble non vide. En indexant chaque partie de $\mathcal{P}(\Omega)$ par elle-même, on voit que les opérations \cup et \cap s'étendent aux sous-ensembles quelconques de $\mathcal{P}(\Omega)$:

Si $\mathcal{A} \subset \mathcal{P}(\Omega)$, la **réunion de \mathcal{A}** est $\bigcup_{X \in \mathcal{A}} X$, c'est-à-dire l'ensemble des éléments de Ω qui appartiennent à au moins un élément X de l'ensemble de parties \mathcal{A} ; l'**intersection de \mathcal{A}** est $\bigcap_{X \in \mathcal{A}} X$.

Par exemple $\Omega = \bigcup_{x \in \Omega} \{x\}$, autrement dit Ω est l'union des singletons formés par ses éléments ; ou encore $\Omega = \bigcup_{X \in \mathcal{P}(\Omega) \setminus \{\emptyset\}} X$, autrement dit Ω est l'union de ses parties non vides.

Par *convention*, on pose $\Omega = \bigcap_{X \in \emptyset} X$, $\emptyset = \bigcup_{X \in \emptyset} X$, lorsque \emptyset est considérée comme la *partie vide de $\mathcal{P}(\Omega)$* , indexée par elle-même. On peut montrer que cette convention est tout à fait cohérente dans n'importe quelle théorie usuelle des ensembles. Elle peut s'avérer utile lorsqu'on a affaire à des familles de parties de Ω dépendant d'un paramètre.

Ensemble somme (ou union disjointe)

Soit $(E_i)_{i \in I}$ une famille d'ensembles, indexée par l'ensemble I non vide. Il est souvent commode de considérer un ensemble qui soit « union disjointe » des E_i , c'est-à-dire, union d'une famille d'*exemplaires* (\mathcal{E}_i) des E_i tels que $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset$ pour tous i, j tels que $i \neq j$. On y parvient aisément de la façon suivante : soit $E = \bigcup_{i \in I} E_i$; dans l'ensemble produit

$F = I \times E$, formons pour chaque $i \in I$ l'ensemble $\mathcal{E}_i = \{i\} \times E_i$. L'application $g_i : x \mapsto (i, x)$ est une bijection de E_i sur \mathcal{E}_i , si bien que \mathcal{E}_i peut être considéré comme un *exemplaire* de E_i . Il est bien clair que si $i \in I$, $j \in I$ et $i \neq j$, alors $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset$. On répond au problème posé en considérant l'ensemble $\mathcal{S} = \bigcup_{i \in I} \mathcal{E}_i$; cet ensemble est appelé **somme**, ou **union**

disjointe, des E_i et sera noté $\coprod_{i \in I} E_i$.

Lorsque cela n'entraîne pas de confusion, pour chaque i on pourra identifier E_i à son image \mathcal{E}_i par l'application g_i dans la somme $\coprod_{i \in I} E_i$.

Partages

DÉFINITION I.4.7

Soit E un ensemble, on appelle **partage** de E toute famille $(X_i)_{i \in I}$ de parties de E possédant les propriétés suivantes :

- (I) $(\forall i \in I), (\forall j \in I) (i \neq j) \Rightarrow X_i \cap X_j = \emptyset$
 (II) $\bigcup_{i \in I} X_i = E$.

Par exemple si $X_1 = \emptyset$ et $X_2 = E$, (X_1, X_2) est un partage de E .

$(\{x\})_{x \in E}$ est un autre partage de E .

Si $(E_i)_{i \in I}$ est une famille d'ensembles, et si \mathcal{E}_i est l'image canonique de E_i dans $\coprod_{i \in I} E_i$, $(\mathcal{E}_i)_{i \in I}$ constitue un partage de $\coprod_{i \in I} E_i$.

On se gardera de confondre cette notion avec celle de *partition* dont la définition sera donnée au § I.5.

Exercice 1 : Soit $(A_i)_{i \in I}$ une famille de parties de l'ensemble Ω ; comparer $\mathcal{P}\left(\bigcap_{i \in I} A_i\right)$ et $\bigcap_{i \in I} \mathcal{P}(A_i)$, puis $\mathcal{P}\left(\bigcup_{i \in I} A_i\right)$ et $\bigcup_{i \in I} \mathcal{P}(A_i)$.

Exercice 2 : On appelle *recouvrement* d'un ensemble E toute famille $(F_i)_{i \in I}$ de parties de E telles que $\bigcup_{i \in I} F_i = E$. Si $(F_i)_{i \in I}$ et $(G_j)_{j \in J}$ sont deux recouvrements de E , le second est dit *plus fin* que le premier ssi $(\forall j \in J) (\exists i \in I) \mid G_j \subset F_i$. Montrer qu'étant donnés deux recouvrements quelconques de E , il en existe un troisième plus fin que chacun des deux.

Exercice 3 : Démontrer en détail les relations (2), (3), (4), (5), (7) et (8).

Exercice 4 : Soit une famille de parties de E , indexée par les éléments d'un produit cartésien $I \times J$, X_{ij} l'élément générique de cette famille. Soit f une application de I dans J et \mathcal{F} la famille de ces applications.

- a) Comparer $\bigcap_{i \in I} \left(\bigcup_{j \in J} X_{ij} \right)$ et $\bigcup_{j \in J} \left(\bigcap_{i \in I} X_{ij} \right)$.
 b) Démontrer $\bigcap_{i \in I} \left(\bigcup_{j \in J} X_{ij} \right) = \bigcup_{f \in \mathcal{F}} \left(\bigcap_{i \in I} X_{if(i)} \right)$ et en déduire la formule duale.

Exercice 5 : Soit E l'ensemble des points d'un plan, O un point fixe de E . On considère l'application $f : E \rightarrow \delta(E)$ associant à chaque point x l'intérieur du cercle de centre O passant par x . Soit \mathcal{F} la famille des droites D passant par O . Comparer $f\left(\bigcap_{D \in \mathcal{F}} D\right)$ et $\bigcap_{D \in \mathcal{F}} f(D)$.

§ 1.5 RELATIONS D'ÉQUIVALENCE. ENSEMBLE QUOTIENT

DÉFINITION I.5.1

Soit E un ensemble non vide. On appelle **relation d'équivalence** sur E toute relation binaire \mathcal{R} à la fois **réflexive, symétrique et transitive**, c'est-à-dire telle que :

$$(I) \quad \forall x \in E, \quad x \mathcal{R} x \quad (\text{c'est la réflexivité})$$

$$(II) \quad \forall x \in E, \quad \forall y \in E, \quad x \mathcal{R} y \Rightarrow y \mathcal{R} x \quad (\text{c'est la symétrie})$$

$$(III) \quad \forall x \in E, \quad \forall y \in E, \quad \forall z \in E \quad (x \mathcal{R} y \text{ et } y \mathcal{R} z) \Rightarrow x \mathcal{R} z \\ (\text{c'est la transitivité}).$$

Si \mathcal{R} est une telle relation sur l'ensemble E , le graphe G de \mathcal{R} contient la diagonale de $E \times E$ à cause de la réflexivité.

A chaque $x \in E$, associons l'ensemble $\mathcal{C}(x) = \{ y \mid y \in E \text{ et } x \mathcal{R} y \}$. Cet ensemble $\mathcal{C}(x)$ est non vide puisque $x \in \mathcal{C}(x)$. Du fait que \mathcal{R} est symétrique, si $x \in \mathcal{C}(y)$ alors $y \in \mathcal{C}(x)$; la transitivité montre que si $y \in \mathcal{C}(x)$, alors $\mathcal{C}(y) \subset \mathcal{C}(x)$, d'où par symétrie $\mathcal{C}(x) \subset \mathcal{C}(y)$ et donc $\mathcal{C}(x) = \mathcal{C}(y)$. Par suite si $x \in E$ et $y \in E$, ou bien $\mathcal{C}(x) = \mathcal{C}(y)$, ou bien $\mathcal{C}(x) \cap \mathcal{C}(y) = \emptyset$. Si $x \in E$, l'ensemble $\mathcal{C}(x)$ s'appelle **classe d'équivalence de x modulo \mathcal{R}** ; on a ainsi défini une application $E \rightarrow \mathcal{P}(E)$, $x \mapsto \mathcal{C}(x)$.

DÉFINITION I.5.2

Avec les notations ci-dessus, la partie de $\mathcal{P}(E)$ égale à l'image de l'application $E \rightarrow \mathcal{P}(E)$, $x \mapsto \mathcal{C}(x)$ s'appelle **ensemble quotient de E par la relation d'équivalence \mathcal{R}** et se note souvent E/\mathcal{R} .

THÉORÈME I.5.1

Soit $Q = E/\mathcal{R}$ l'ensemble quotient de l'ensemble non vide donné E par la relation d'équivalence \mathcal{R} sur E . L'ensemble Q possède les propriétés suivantes :

$$(I) \quad \forall X, \quad X \in Q \Rightarrow X \neq \emptyset$$

$$(II) \quad (\forall X) \quad (\forall Y) \quad (X \in Q \text{ et } Y \in Q \text{ et } X \neq Y) \Rightarrow X \cap Y = \emptyset$$

$$(III) \quad \bigcup_{X \in Q} X = E.$$

Démonstration :

Si $X \in Q$, par définition il existe $x \in E$ tel que $X = \mathcal{C}(x)$; pour un tel x , on a : $x \in \mathcal{C}(x)$, d'où (I). Dans ce qui précède la définition I.5.2 on a vu (II) ; enfin si $x \in E$, on a $x \in \mathcal{C}(x)$ et $\mathcal{C}(x) \in Q$, c

DÉFINITION I.5.3

Si E est un ensemble non vide, on appelle **partition** de E toute partie Q de $\mathcal{P}(E)$ qui possède les propriétés (I), (II), (III) du théorème précédent. Autrement dit, une partition de E est une partie de $\mathcal{P}(E)$ dont la famille associée constitue un partage de E en ensembles non vides.

Une relation d'équivalence sur un ensemble donné E s'identifie à une partie de $E \times E$; donc les relations d'équivalence sur E constituent un ensemble, que nous noterons $\mathcal{E}qu(E)$. ($\mathcal{E}qu(E) \subset \mathcal{P}(E \times E)$.)

Une partition de E est une certaine partie de $\mathcal{P}(E)$; donc les partitions de E constituent un ensemble, que nous noterons $\mathcal{Part}(E)$. ($\mathcal{Part}(E) \subset \mathcal{P}(\mathcal{P}(E))$.)

A chaque relation d'équivalence \mathcal{R} sur E , on a associé une partition E/\mathcal{R} de E ; réciproquement, soit Q une partition de E : associons-lui la relation binaire \mathcal{R} sur E définie par $(\forall x \in E), (\forall y \in E) x \mathcal{R} y$ ssi $\exists X \in Q \mid x \in X \text{ et } y \in X$. Les propriétés (I), (II), (III) montrent immédiatement que \mathcal{R} est une relation d'équivalence, et que la partition associée à \mathcal{R} est Q . Si \mathcal{R} est une relation d'équivalence sur E et si $Q = E/\mathcal{R}$, on vérifie que la relation d'équivalence associée à la partition Q est précisément \mathcal{R} . On a donc en fin de compte :

THÉOREME I.5.2

Soit E un ensemble non vide. On obtient une bijection de l'ensemble $\mathcal{E}qu(E)$ sur l'ensemble $\mathcal{Part}(E)$ en associant, à chaque relation d'équivalence \mathcal{R} sur E , l'ensemble quotient E/\mathcal{R} . La bijection réciproque associe, à chaque partition Q de E , l'unique relation d'équivalence dont les classes sont les éléments de Q .

Exemple 1 : $x = x$ est une relation d'équivalence sur l'ensemble non vide donné E ; son graphe en est la diagonale de $E \times E$; l'ensemble quotient est $\{ \{x\} \}_{x \in E}$. On peut l'identifier avec E .

Exemple 2 : Soit E un espace euclidien de dimension $n \geq 2$; sur l'ensemble $\mathcal{P}(E)$, définissons la relation binaire $X \mathcal{R} Y$ ainsi :

$X \mathcal{R} Y$ ssi il existe un élément φ du groupe $SO(E)$ tel que $\varphi(X) = Y$.

Alors \mathcal{R} est une relation d'équivalence sur $\mathcal{P}(E)$. Deux parties X et $Y \in \mathcal{P}(E)$ seront dites « **congruentes** modulo les déplacements » (ou, dans un langage désuet, « égales »), ssi $X \mathcal{R} Y$.

Exemple 3 : Soit K un corps commutatif et E un K -espace vectoriel non nul. Sur l'ensemble $\check{E} = E \setminus \{0\}$, considérons la relation binaire \mathcal{R} définie par $x \mathcal{R} y$ ssi $\exists \lambda \in K^* \mid y = \lambda x$.

Alors \mathcal{R} est une relation d'équivalence sur \check{E} ; les classes d'équivalences sont les ensembles $(D \setminus \{0\})$, D droite vectorielle de E . L'ensemble quotient E/\mathcal{R} se note $P(E)$ et s'appelle **espace projectif** iss

Exemple 4 : Soit K un corps commutatif et \mathcal{E} un espace affine sur K , non réduit à un point, d'espace directeur $E \neq \{0\}$. Sur l'ensemble $\overrightarrow{\mathcal{E}} \times (E \setminus \{0\})$, la relation \mathcal{R} , définie par $(A, V) \mathcal{R} (A', V')$ ssi $V = V'$ et $\overrightarrow{AA'}$ est lié à V , est une relation d'équivalence. Les classes d'équivalence sont appelées les **vecteurs glissants** (ou **glisseurs**) de \mathcal{E} .

Exemple 5 : Soit K un corps commutatif et E un K -espace vectoriel non nul ; deux éléments u, v de $\text{Hom}_K(E)$ sont dits **semblables** ssi il existe $\varphi \in \text{GL}_K(E)$ tel que $v = \varphi \circ u \circ \varphi^{-1}$. La relation « u et v sont semblables » est d'équivalence sur $\text{Hom}_K(E)$. Les classes d'équivalence de cette relation sont appelées *classes de similitude* de $\text{Hom}_K(E)$. Nous les étudierons au Chapitre XVI.

Ces quelques exemples suffisent à montrer qu'il n'y a guère de circonstances, en mathématiques, où l'on ne soit amené à « passer au quotient » par une relation d'équivalence. Nous en verrons d'autres, très nombreuses, dans cet ouvrage.

Projection canonique

Soit E un ensemble non vide muni d'une relation d'équivalence \mathcal{R} . On a défini plus haut l'application $\varphi : x \mapsto \mathcal{C}(x)$, $E \rightarrow E/\mathcal{R}$.

Cette application s'appelle *application* (ou *projection*) *canonique* de E sur E/\mathcal{R} . Elle est *surjective* (cf. propriété (III)). La relation $(x \in E, y \in E \text{ et } x \mathcal{R} y)$ équivaut à $(x \in E, y \in E \text{ et } \varphi(x) = \varphi(y))$. On a

$$E/\mathcal{R} = \{\varphi^{-1}(X)\}_{X \in E/\mathcal{R}} = \{X\}_{X \in E/\mathcal{R}}$$

car $X = \varphi^{-1}(X)$ pour $X \in E/\mathcal{R}$.

THÉORÈME 1.5.3

Soit E un ensemble non vide muni d'une relation d'équivalence \mathcal{R} , $\varphi : E \rightarrow E/\mathcal{R}$ l'application canonique, et F un ensemble. Si une application $f : E \rightarrow F$ est **constante** sur chaque élément $X \in E/\mathcal{R}$, alors il existe une, et une seule, application $\bar{f} : E/\mathcal{R} \rightarrow F$ telle que $\bar{f} \circ \varphi = f$.

Démonstration :

Supposons que \bar{f} existe ; alors pour tout $X \in E/\mathcal{R}$, $\bar{f}(X)$ est la valeur constante de f sur X , d'où l'unicité.

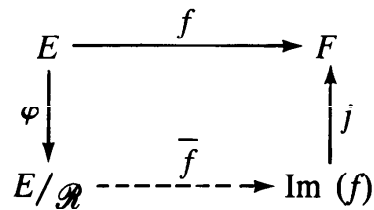
Réciproquement, pour $X \in E/\mathcal{R}$, notons $\bar{f}(X)$ la valeur constante de f sur X , c'est-à-dire l'élément $y \in F$ tel que $\forall x \in X, f(x) = y$. Alors on a défini $\bar{f} : E/\mathcal{R} \rightarrow F$ et on constate que $\bar{f} \circ \varphi = f$. ■

Décomposition canonique d'une application

THÉORÈME 1.5.4

Soit $f : E \rightarrow F$ une application, avec E et F non vides, \mathcal{R} la relation binaire définie sur E par : $x \mathcal{R} y$ ssi $f(x) = f(y)$. Alors \mathcal{R} est une relation d'équivalence ; soit $j : f(E) \rightarrow F$ l'injection canonique et $\varphi : E \rightarrow E/\mathcal{R}$ la projection canonique ; il existe une application et une seule $\bar{f} : E/\mathcal{R} \rightarrow f(E)$ telle que $j \circ \bar{f} \circ \varphi = f$. Cette application \bar{f} est une bijection.

Le schéma suivant rend compte de l'énoncé I.5.4 :



Ce schéma est un exemple de « diagramme commutatif », ce qui signifie qu'en composant (quand c'est possible) tout système de « flèches » qui joignent deux points du diagramme, on obtient un résultat qui ne dépend que des points de départ et d'arrivée et non du chemin suivi (ici il y a deux façons d'aller de E en F).

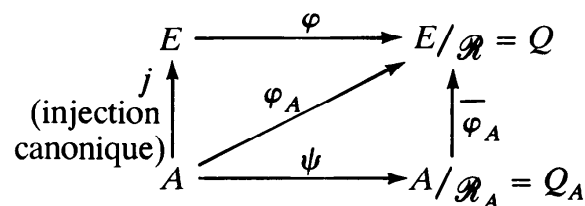
Démonstration du théorème I.5.4 :

On vérifie d'abord que \mathcal{R} est bien une relation d'équivalence. L'existence et l'unicité de \bar{f} résultent du théorème I.5.3. Comme $\bar{f} \circ \varphi$ coïncide avec la corestriction $f|_{\text{Im}(f)}$, on voit que $\bar{f} \circ \varphi$ est surjective, donc \bar{f} l'est aussi. Si $X \in E/\mathcal{R}$, $Y' \in E/\mathcal{R}$ et $\bar{f}(X) = \bar{f}(Y')$, prenons $x \in X$ et $y \in Y'$; on a : $\bar{f}(X) = f(x)$, $\bar{f}(Y') = f(y)$, donc $f(x) = f(y)$, donc $x \mathcal{R} y$, d'où $X = Y'$, ce qui prouve que \bar{f} est injective. ■

La décomposition $f = j \circ \bar{f} \circ \varphi$ s'appelle *factorisation canonique de f* .

Relation d'équivalence induite

Soit \mathcal{R} une relation d'équivalence sur l'ensemble non vide E , et soit A une partie non vide de E . La relation : « $x \in A$, $y \in A$ et $x \mathcal{R} y$ » est une relation d'équivalence \mathcal{R}_A sur A , dite *induite sur A par \mathcal{R}* . Soit Q_A la partie de $Q = E/\mathcal{R}$ formée des $X \in Q$ tels que $X \cap A \neq \emptyset$. Alors l'ensemble quotient A/\mathcal{R}_A est l'ensemble $\{X \cap A\}_{X \in Q_A}$. La restriction φ_A de la projection canonique $\varphi : E \rightarrow Q$ à l'ensemble A est constante sur chaque classe de \mathcal{R}_A . Donc φ_A se factorise en $\bar{\varphi}_A \circ \psi$ où $\psi : A \rightarrow Q_A$ est la projection canonique. On en déduit le diagramme commutatif :



dans lequel $\varphi_A = \varphi \circ j = \bar{\varphi}_A \circ \psi$. Comme toute classe mod (\mathcal{R}_A) est incluse dans une certaine classe mod (\mathcal{R}) , $\bar{\varphi}_A$ est *injective*. On dit que $\bar{\varphi}_A$ est l'*injection canonique* de A/\mathcal{R}_A dans E/\mathcal{R} .

Exercice 1 : Soit E un ensemble non vide, A une partie non vide de E et \mathcal{R} la relation d'équivalence sur $\mathcal{P}(E)$ définie par $X \mathcal{R} Y$ ssi $X \cup A = Y \cup A$; on considère $f : \mathcal{P}(E) \rightarrow \mathcal{P}(E \setminus A)$, $X \mapsto X \cap (E \setminus A)$. Montrer que f est constante sur les classes mod' (\mathcal{R}) . Ecrire la décomposition canonique de f . Si F est une partie non vide de E fixée, étudier la relation d'équivalence induite par \mathcal{R} sur $\mathcal{P}(F)$.

Exercice 2 : Soit E un espace euclidien de dimension $n \geq 2$ et S la sphère

centre O_E dans $E : S = \{x \mid x \in E \text{ et } \|x\| = 1\}$. Soit \mathcal{R} la relation binaire sur S définie par : $x \mathcal{R} y$ ssi $x = y$ ou $x + y = 0$. Montrer que \mathcal{R} est une relation d'équivalence, et qu'on a une bijection naturelle entre l'ensemble E/\mathcal{R} et l'espace projectif $P(E)$ défini dans l'exemple 3.

Exercice 3 : Soit $\mathcal{C} = \mathcal{F}(\mathbb{R}^*, \mathbb{R})$ et $n \in \mathbb{N}$. On pose (pour $f \in \mathcal{C}$ et $g \in \mathcal{C}$) $f \mathcal{R} g$ ssi $\frac{1}{x^n} (f(x) - g(x)) \rightarrow 0$ quand $x \rightarrow 0$, ce qu'on écrit encore $f(x) - g(x) \in o(x^n)$. Montrer que \mathcal{R} est une relation d'équivalence sur \mathcal{C} .

Exercice 4 : Soit E et F deux ensembles non vides, munis de relations d'équivalence \mathcal{R} et \mathcal{S} . On définit la relation binaire \mathcal{T} sur $E \times F$ ainsi : $(\forall (x, y) \in E \times F), (\forall (x', y') \in E \times F)$ $(x, y) \mathcal{T} (x', y')$ ssi $x \mathcal{R} x'$ et $y \mathcal{S} y'$. Montrer que \mathcal{T} est une relation d'équivalence. Montrer qu'il existe une bijection naturelle de $(E \times F)/\mathcal{T}$ sur $E/\mathcal{R} \times F/\mathcal{S}$.

Indication : On écrira la décomposition canonique de $\Phi : E \times F \rightarrow E/\mathcal{R} \times F/\mathcal{S}$, $(x, y) \mapsto (\varphi(x), \psi(y))$ où φ (resp. ψ) est l'application canonique de E sur E/\mathcal{R} (resp. de F sur F/\mathcal{S}). Étendre ces résultats au produit d'un nombre fini quelconque d'ensembles munis de relations d'équivalence.

Exercice 5 : Soit E un ensemble non vide muni d'une relation d'équivalence \mathcal{R} , A une partie non vide de E , et \mathcal{R}_A la relation d'équivalence induite par \mathcal{R} sur A . Condition nécessaire et suffisante pour que l'injection canonique $A/\mathcal{R}_A \rightarrow E/\mathcal{R}$ soit bijective ? Appliquer à l'exercice 2.

Exercice 6 : Soit E un ensemble non vide. Si \mathcal{R} et \mathcal{S} sont deux relations d'équivalence sur E , on dit que \mathcal{S} est *plus fine* que \mathcal{R} ssi $(\forall x \in E), (\forall y \in E) x \mathcal{S} y \Rightarrow x \mathcal{R} y$. On considère deux telles relations, avec \mathcal{S} plus fine que \mathcal{R} .

- Vérifier que toute classe mod (\mathcal{S}) est contenue dans une, et une seule, classe mod (\mathcal{R}) .
- Soit $f : E/\mathcal{S} \rightarrow E/\mathcal{R}$ l'application qui, à tout $X \in E/\mathcal{S}$, associe l'unique classe Z mod (\mathcal{R}) telle que $X \subset Z$. Vérifier que f est surjective. Décomposition canonique de f ?
- Soit \mathcal{T} la relation d'équivalence définie par f sur E/\mathcal{S} (si $X \in E/\mathcal{S}$ et si $Y \in E/\mathcal{S}$, $X \mathcal{T} Y$ ssi $f(X) = f(Y)$). Dédurre de b) une bijection naturelle entre les ensembles $(E/\mathcal{S})/\mathcal{T}$ et E/\mathcal{R} .

Exercice 7 : Chercher, sous les hypothèses du théorème I.5.3, une condition nécessaire et suffisante pour que f soit injective (resp. surjective).

Exercice 8 : Dans $\mathcal{P}(\Omega)$, ensemble des parties d'un ensemble non vide Ω , on dit que A rencontre B ssi $A \cap B \neq \emptyset$. Montrer que cette relation binaire n'est ni réflexive, ni transitive.

Exercice 9 : Soit \mathcal{S} une relation d'équivalence sur F et f une application de E dans F . On définit sur E la relation \mathcal{R} par : $x \mathcal{R} y$ ssi $f(x) \mathcal{S} f(y)$. Montrer que \mathcal{R} est une relation d'équivalence.

Exercice 10 : Soit \mathcal{S} une relation binaire sur l'ensemble E qui est réflexive et transitive, mais pas symétrique. Considérons la relation \mathcal{R} définie par $x \mathcal{R} y$ ssi $x \mathcal{S} y$ et $y \mathcal{S} x$. Montrer que \mathcal{R} est d'équivalence.

Exercice 11 : Soit E un ensemble non vide et \mathcal{A} un ensemble de parties de E . On définit la relation binaire \mathcal{R} dans E par : $x \mathcal{R} y$ ssi $\forall A \in \mathcal{A}, \{x, y\} \subset A$ ou $\{x, y\} \subset E \setminus A$. Montrer que \mathcal{R} est une relation d'équivalence. Décrire les classes d'équivalence.

§ I.6 RELATIONS D'ORDRE

DÉFINITION I.6.1

Une relation binaire \mathcal{R} sur un ensemble E est appelée **relation d'ordre** ssi elle est **réflexive, antisymétrique, et transitive** c'est-à-dire ssi elle vérifie :

$$(I) \quad (\forall x \in E) \quad x \mathcal{R} x \quad \text{(c'est la réflexivité)}$$

$$(II) \quad (\forall x \in E) \quad (\forall y \in E) \quad (x \mathcal{R} y \text{ et } y \mathcal{R} x) \Rightarrow x = y \quad \text{(c'est l'antisymétrie)}$$

$$(III) \quad (\forall x \in E, \forall y \in E, \forall z \in E), \quad (x \mathcal{R} y \text{ et } y \mathcal{R} z) \Rightarrow x \mathcal{R} z \quad \text{(c'est la transitivité)}.$$

Un ensemble muni d'une relation d'ordre est appelé un *ensemble ordonné*.

Pour désigner une relation d'ordre, on emploie souvent le symbole \leq , ou le symbole \preceq ; si c'est le cas, la relation « $x \leq y$ et $x \neq y$ » s'écrira $x < y$ (resp. $x \prec y$). Cependant certaines relations d'ordre spécifique sont notées avec des symboles adaptés (par exemple, l'inclusion d'ensembles qui se note \subset , et \subsetneq pour l'inclusion stricte, ou la divisibilité dans \mathbb{N}^* qui se note « $x \mid y$ »). $x \leq y$ se note aussi $y \geq x$.

Un ordre \leq sur un ensemble E est dit **total** (et E est alors dit *totalelement ordonné*) ssi deux éléments quelconques de E sont *comparables*, c'est-à-dire ssi $(x \in E \text{ et } y \in E) \Rightarrow (x \leq y \text{ ou } y \leq x)$, ou encore, de manière équivalente, ssi $(x \in E, y \in E \text{ et } x \neq y) \Rightarrow (x < y \text{ ou } y < x)$. Un ordre est dit **partiel** ssi il est non total.

Soit (E, \leq) un ensemble ordonné ; fixons une partie F de E ; soit \mathcal{R} la relation binaire définie sur F par : $x \mathcal{R} y$ ssi $x \in F, y \in F$ et $x \leq y$. Alors \mathcal{R} est une relation d'ordre sur F , dite **induite sur F** par l'ordre de E . On notera en général \mathcal{R} par le même symbole que celui (ici, \leq) qui représente l'ordre de E . L'ensemble (F, \leq) est alors appelé le *sous-ensemble ordonné F de E* . Tout sous-ensemble ordonné d'un sous-ensemble ordonné de E est évidemment un sous-ensemble ordonné de E . Tout sous-ensemble ordonné d'un ensemble totalement ordonné est totalement ordonné.

Exemple 1 : Sur chacun des ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ et \mathbb{R} on définit l'ordre dit naturel ou *ordre usuel* (voir le chapitre II de cet ouvrage pour les trois premiers, le tome d'Analyse pour le quatrième). Cet ordre usuel se note \leq (qu'on lit « inférieur ou égal à » ou « au plus égal à » ou plus brièvement « inférieur à »). Il est total.

L'ordre usuel de \mathbb{R} induit l'ordre usuel sur \mathbb{N}, \mathbb{Z} et \mathbb{Q} .

Exemple 2 : Soit A un ensemble non vide, (B, \leq) un ensemble non vide ordonné, et $E = \mathcal{F}(A, B)$. Si $f \in E$ et $g \in E$, convenons que : $f \leq g$ ssi on a :

$$(\forall x) \quad (x \in A) \Rightarrow f(x) \leq g(x).$$

Alors $f \leq g$ est une relation d'ordre sur E ; mais même si l'ordre \leq est total, en général l'ordre \leq n'est que partiel.

Exemple 3 : Soit \mathbb{N}^* ; si $x \in \mathbb{N}^*$ et $y \in \mathbb{N}^*$, on dit que x *divise* y , et on écrit $x \mid y$, ssi il existe $z \in \mathbb{N}^*$ tel que $y = xz$. Alors la relation $x \mid y$ est un ordre partiel sur \mathbb{N}^* .

Exemple 4 : Soit E un ensemble. Sur l'ensemble $\mathcal{P}(E)$, la relation $X \subset Y$ est une relation d'ordre partiel (sauf si $E = \emptyset$ ou si E est un singleton).

Cet ordre est l'un des plus intéressants parmi les ordres partiels sur des ensembles.

Exemple 5 : Soit Ω un ensemble non vide et $\mathcal{E}qu(\Omega)$ l'ensemble des relations d'équivalence sur Ω ; sur $\mathcal{E}qu(\Omega)$, la relation « \mathcal{S} est plus fine que \mathcal{R} » (cf. exercice 6 du § I.5) est une relation d'ordre.

Applications monotones

Soit (E, \leq) et (F, \leq) deux ensembles ordonnés. Une application $f : E \rightarrow F$ est dite **croissante** ssi elle « respecte l'ordre », c'est-à-dire, vérifie :

$$(\forall x \in E, \forall y \in E) \quad (x \leq y) \Rightarrow (f(x) \leq f(y));$$

f est dite **strictement croissante** ssi

$$(\forall x \in E, \forall y \in E) \quad (x < y) \Rightarrow (f(x) < f(y));$$

f est dite **décroissante** ssi

$$(\forall x \in E, \forall y \in E) \quad (x \leq y) \Rightarrow (f(x) \geq f(y));$$

f est dite **strictement décroissante** ssi

$$(\forall x \in E, \forall y \in E) \quad (x < y) \Rightarrow (f(x) > f(y));$$

f est dite **monotone** ssi f est croissante ou décroissante.

f est dite **strictement monotone** ssi f est strictement croissante ou strictement décroissante.

Lorsque (E, \leq) est *totale*ment ordonné, si f est injective et monotone, alors elle est strictement monotone ; si f est strictement monotone, alors elle est injective.

Dans le cas général il est clair que si f est strictement croissante (resp. strictement décroissante), f est croissante (resp. décroissante). De même la composée d'applications croissantes (resp. strictement croissantes) est croissante (resp. strictement croissante). Mais la composée de deux applications décroissantes (resp. strictement décroissantes) est croissante (resp. strictement croissante).

Une bijection f de (E, \leq) sur (F, \leq) telle que f et f^{-1} soient croissantes est appelée un *isomorphisme d'ensembles ordonnés*. Alors $f^{-1}: F \rightarrow E$ est aussi un tel isomorphisme. La composée de deux isomorphismes d'ensembles ordonnés en est un. Id_E est un isomorphisme d'ensembles ordonnés : $E \rightarrow E$.

Exemple 6 : Soit E un ensemble, l'application $\mathcal{P}(E) \rightarrow \mathcal{P}(E)$, $X \mapsto E \setminus X$ est une bijection (involutive) strictement décroissante pour l'inclusion :

$$\text{si } X \in \mathcal{P}(E) \text{ et } Y \in \mathcal{P}(E), \quad X \subsetneq Y \Rightarrow E \setminus Y \subsetneq E \setminus X.$$

Exemple 7 : Soit $f: E \rightarrow F$ une application de l'ensemble E dans l'ensemble F . Ordonnons $\mathcal{P}(E)$ et $\mathcal{P}(F)$ par inclusion. L'application $\hat{f}: \mathcal{P}(E) \rightarrow \mathcal{P}(F)$, $X \mapsto f(X)$ est croissante. Si f est *injective*, alors \hat{f} est strictement croissante. Si f est bijective, \hat{f} est un isomorphisme d'ensembles ordonnés.

L'application $\tilde{f}: \mathcal{P}(F) \rightarrow \mathcal{P}(E)$, $Y \mapsto f^{-1}(Y)$ est croissante ; si f est *surjective*, \tilde{f} est strictement croissante. Si f est bijective, \tilde{f} est un isomorphisme d'ensembles ordonnés : c'est l'isomorphisme réciproque de \hat{f} .

Exemple 8 : Ordonnons \mathbb{N}^* par la divisibilité : $x \mid y$. Soit $k \in \mathbb{N}^*$. Alors l'application : $\mathbb{N}^* \rightarrow \mathbb{N}^*$, $x \mapsto x^k$ est strictement croissante ; en effet $(x \mid y \text{ et } x \neq y) \Rightarrow (x^k \mid y^k \text{ et } x^k \neq y^k)$.

Maximum et minimum

DÉFINITION I.6.2

Un élément μ d'un ensemble ordonné (E, \leq) est dit **plus petit élément** de E (resp. **plus grand élément** de E) ssi $(\forall x \in E) \mu \leq x$ (resp. $(\forall x \in E) x \leq \mu$).

Si un plus petit élément existe, il est unique (à cause de l'antisymétrie de \leq) ; on l'appelle l'**élément minimum** de E . De même, si un plus grand élément existe, il est unique, on l'appelle l'**élément maximum** de E .

Si A est une partie d'un ensemble ordonné (E, \leq) , un plus petit élément de A (resp. plus grand élément de A) est, par définition, un plus petit (resp. plus grand) élément du sous-ensemble ordonné (A, \leq) de E . En cas d'existence, il est unique, et on l'appelle *minimum* de A (resp. *maximum* de A) ; on le note $\text{Min}(A)$ (resp. $\text{Max}(A)$).

DÉFINITION I.6.3

Un ensemble ordonné (E, \leq) est dit **bien ordonné** ssi toute partie non vide de E admet un plus petit élément. On dit encore que \leq est un **bon ordre** sur E .

Un ensemble bien ordonné non vide possède, en particulier, un plus petit élément, appelé son **premier élément**. Un bon ordre est touje

Exemple 9 : \mathbb{N} muni de l'ordre usuel \leq est bien ordonné ; il n'a pas d'élément maximum.

Exemple 10 : Si E est un ensemble, $\mathcal{P}(E)$ muni de l'inclusion admet \emptyset pour plus petit élément et E pour plus grand élément.

Exemple 11 : Dans \mathbb{N}^* ordonné par $x \mid y$, le plus petit élément existe : c'est 1 ; il n'y a pas de plus grand élément. Dans \mathbb{N} ordonné par $x \mid y$ (relation toujours définie par : $\exists z \in \mathbb{N}$ tel que $y = xz$), il y a toujours un plus petit élément, mais il y a aussi un plus grand élément qui est 0. Dans le sous-ensemble ordonné $\mathbb{N} \setminus \{0, 1\}$ de \mathbb{N} ordonné par $x \mid y$, il n'y a ni plus grand ni plus petit élément.

Exemple 12 : Pour l'ordre usuel, ni \mathbb{Z} , ni \mathbb{Q} , ni \mathbb{R} ne sont bien ordonnés.

Élément minimal, élément maximal

DÉFINITION 1.6.4

Soit (E, \leq) un ensemble ordonné. Un élément $\mu \in E$ est dit **minimal** (resp. **maximal**) ssi :

$$(\forall x) \quad (x \in E \text{ et } x \leq \mu) \Rightarrow (x = \mu)$$

$$(\text{resp. } (\forall x) \quad (x \in E \text{ et } x \geq \mu) \Rightarrow (x = \mu)) .$$

Tout élément maximum est maximal, mais la réciproque est fausse ; de même, tout élément minimum est minimal, mais la réciproque est fausse. Dans un ensemble *totalelement ordonné*, les notions d'élément minimum et minimal (resp. maximum et maximal) coïncident. Mais dans un ensemble partiellement ordonné, il peut y avoir de nombreux éléments maximaux (resp. minimaux) distincts. On dit qu'un élément est **extrémal** ssi il est, ou maximal, ou minimal. C'est surtout dans les ensembles partiellement ordonnés que la notion d'élément extrémal est intéressante ; presque toujours, les éléments extrémaux d'un tel ensemble possèdent des propriétés remarquables. Lorsqu'ils sont « suffisamment nombreux » ils permettent (en un sens qu'il faut préciser à chaque fois) de « reconstituer » l'ensemble ordonné.

Exemple 13 : Soit $E = \mathbb{N} \setminus \{0, 1\}$ ordonné par $x \mid y$; alors les éléments minimaux de E ne sont autres que les **nombre premiers**.

Exemple 14 : Soit K un corps commutatif et E un K -espace vectoriel non nul. Notons $\mathcal{G}^*(E)$ l'ensemble des sous-espaces vectoriels *stricts* de E , et ordonnons $\mathcal{G}^*(E)$ par inclusion : les éléments maximaux de $\mathcal{G}^*(E)$ sont les **hyperplans** de E (voir Chap. IX).

Exemple 15 : Soit A un anneau commutatif $\neq \{0\}$ et $\mathcal{I}^*(A)$ l'ensemble des idéaux de A autres que A , ordonné par inclusion. On peu

$\mathcal{J}^*(A)$ admet toujours des éléments maximaux (cf. [27]) : ces éléments sont appelés les **idéaux maximaux** de A .

Majorant, minorant. Borne supérieure, borne inférieure

DÉFINITION I.6.5

Soit (E, \leq) un ensemble ordonné et A une partie de E . On dit que M est un **majorant de A** ssi $M \in E$ et $(\forall x) (x \in A) \Rightarrow (x \leq M)$. La partie A est dite **majorée** ssi elle admet au moins un majorant. De même, m est appelé un **minorant de A** ssi $m \in E$ et $(\forall x) (x \in A) \Rightarrow (x \geq m)$. La partie A est dite **minorée** ssi elle admet au moins un minorant.

Par exemple, dans \mathbb{R} muni de l'ordre usuel, \mathbb{N} est une partie minorée, mais non majorée ; 0 est minorant de \mathbb{N} .

Dans \mathbb{N} muni de l'ordre usuel, toute partie *finie* est majorée.

DÉFINITION I.6.6

Soit (E, \leq) un ensemble ordonné et soit A une partie de E qui soit majorée ; on dit que A **admet une borne supérieure dans E** ssi l'ensemble $\mathcal{Maj}(A)$ des majorants de A dans E admet un plus petit élément ; dans ce cas, ce plus petit élément est appelé la **borne supérieure de A dans E** . On définit de manière analogue la notion de borne inférieure de A dans E .

Si elle existe, la borne supérieure se note $\sup_{x \in A} x$, ou, si cela n'entraîne aucune confusion, $\sup_E(A)$, ou $\sup(x)$. Il faut bien remarquer que l'élément $\sup_E(A)$ dépend du couple (A, E) . Si F est un sous-ensemble de E qui contient A , il se peut que les deux éléments $\sup_F(A)$ et $\sup_E(A)$ existent et soient distincts ; ou bien, qu'un seul des deux éléments $\sup_F(A)$ et $\sup_E(A)$ existe.

Exemple 16 : Soit Ω un ensemble non vide. Ordonnons $\mathcal{P}(\Omega)$ par inclusion ; si \mathcal{S} est une partie non vide de $\mathcal{P}(\Omega)$, \mathcal{S} possède une borne inférieure, qui est $\bigcap_{X \in \mathcal{S}} X$, et une borne supérieure, qui est $\bigcup_{X \in \mathcal{S}} X$.

Exemple 17 : Dans \mathbb{N}^* ordonné par $x \mid y$, toute partie *finie* non vide A admet une borne inférieure, qui est le p.g.c.d. des $a \in A$, et une borne supérieure, qui est le p.p.c.m. des $a \in A$.

Exemple 18 : Munissons \mathbb{R} de l'ordre usuel et soit l'ensemble $A = \{x \mid x \in \mathbb{Q} \text{ et } x^2 < 2\}$. Alors $\sup_{\mathbb{R}}(A)$ existe et vaut $\sqrt{2}$; $\sup_{\mathbb{Q}}(A)$ n'existe pas. Si $E = A \cup \mathbb{Z}$, alors $E \subset \mathbb{Q}$ et $\sup_E(A)$ existe.

Exemple 19 : Dans un ensemble *bien ordonné* non vide (E, \leq) , toute partie non vide et majorée admet une borne supérieure.

Exemple 20 : Soit K un corps commutatif et E un K -espace vectoriel. Dans l'ensemble, ordonné par inclusion, des sous-espaces vectoriels de E , toute partie non vide \mathcal{S} admet une borne inférieure, qui est $\bigcap_{V \in \mathcal{S}} V$, et une borne supérieure, qui est $\sum_{V \in \mathcal{S}} V$, comme on le verra au chapitre X.

La définition I.6.6 peut s'exprimer ainsi : pour qu'un élément $b \in E$ soit la borne supérieure de A dans E , il faut et il suffit qu'il vérifie les deux propriétés (I) et (II) :

(I) pour tout $a \in A$, on a : $a \leq b$;

(II) si $c \in E$ et si on n'a pas $b \leq c$, il existe $a \in A$ tel qu'on n'ait pas $a \leq c$.

Dans le cas d'un ensemble (E, \leq) *totalelement ordonné*, la condition (II) se simplifie en la suivante :

(II') pour tout $c \in E$ tel que $c < b$, il existe $a \in A$ tel que $a > c$, d'où :

THÉOREME I.6.1

Soit A une partie d'un ensemble *totalelement ordonné* (E, \leq) . Pour qu'un élément $b \in E$ soit la borne supérieure de A dans E , il faut et il suffit qu'il vérifie simultanément les deux conditions :

(I) $\forall a \in A \quad a \leq b \quad (b \text{ est un majorant})$

(II) $\forall c \in E, \quad (c < b) \Rightarrow \exists a \in A, \quad a > c$
(b est le premier majorant) .

Soit enfin f une application d'un ensemble X dans un ensemble ordonné (F, \leq) . Si l'ensemble $f(X)$ possède une borne supérieure dans F , on dit que c'est la borne supérieure de f dans F et on la note $\sup_{x \in X} f(x)$ (ou $\sup f$).

Les propriétés des bornes inférieures s'étudieraient de même.

Ordre produit

Soit $(E_i, \leq)_{i \in I}$ une famille d'ensembles ordonnés non vides avec I non vide ; sur l'ensemble produit $E = \prod_{i \in I} E_i$, on définit une *relation d'ordre* en convenant, si $x \in E$ et $y \in E$, $x = (x_i)_{i \in I}$, $y = (y_i)_{i \in I}$, que $x \leq y$ ssi $(\forall i \in I) x_i \leq y_i$.

La relation d'ordre ainsi obtenue s'appelle **ordre produit** des ordres des E_i , et (E, \leq) est appelé l'**ensemble ordonné produit** des (E_i, \leq) . Même si les (E_i, \leq) sont *totalelement ordonnés*, l'ordre \leq n'est en général que *partiel* ; l'ordre défini dans l'exemple 2 est le cas particulier où tous les E_i sont égaux.

Ordre lexicographique

Soit $(E_i, \leqslant)_{i \in I}$ une famille non vide d'ensembles *totale*ment ordonnés non vides, et supposons l'ensemble I *bien ordonné* par une relation d'ordre \mathcal{R} ; considérons, sur l'ensemble produit $E = \prod_{i \in I} E_i$, la relation \leqslant ainsi définie : pour $x \in E, y \in E, x = (x_i)_{i \in I}, y = (y_i)_{i \in I}$, si $x = y$ alors $x \leqslant y$; si $x \neq y$, soit i_0 le plus petit des $i \in I$ tels que $x_i \neq y_i$ pour l'ordre \mathcal{R} (qui existe, puisque \mathcal{R} est un bon ordre et que l'ensemble $\{i \in I \mid x_i \neq y_i\}$ est non vide) ; alors $x \leqslant y$ ssi $x_{i_0} \leqslant y_{i_0}$.

On constate immédiatement que la relation $x \leqslant y$ sur E est une *relation d'ordre total* : par définition, cet ordre est le **produit lexicographique** (des ordres des E_i relatifs à l'ordre \mathcal{R} de I), et l'ensemble ordonné (E, \leqslant) est le **produit lexicographique des ensembles ordonnés** (E_i, \leqslant) selon l'ordre \mathcal{R} de I .

Lorsque tous les E_i sont égaux à un même ensemble totalement ordonné \mathcal{E} , et que $I = \llbracket 1, n \rrbracket$, avec $n \in \mathbb{N}^*$ (resp. $I = \mathbb{N}^*$), muni de l'ordre usuel, l'ensemble ordonné $E = \mathcal{E}^n$ (resp. $\mathcal{E}^{\mathbb{N}^*}$), ordonné comme ci-dessus, est dit **ordonné lexicographiquement** (à partir de l'ordre de \mathcal{E}). Par exemple, si $\mathcal{E} = \mathbb{N}$ muni de l'ordre usuel, dans \mathbb{N}^2 ordonné lexicographiquement, on a :

$$(0, 0) \leqslant (0, 1) \leqslant (1, 0) \leqslant (1, 1) \leqslant (2, 0) \leqslant (2, 1) \leqslant (2, 2) \leqslant (3, 0) \dots$$

Remarque 1 : Soit \mathcal{E} un ensemble bien ordonné non vide ayant au moins 2 éléments et \square son plus petit élément, et posons $\mathcal{A} = \mathcal{E} \setminus \{\square\}$: on peut considérer \mathcal{A} comme un *alphabet* et \square comme un *symbole d'espacement* ; munissons l'ensemble produit $E = \mathcal{E}^{\mathbb{N}^*}$ de l'ordre lexicographique construit sur l'ordre naturel de \mathbb{N}^* ; l'élément $(x_p)_{p \in \mathbb{N}^*}$ de E tel que $x_p = \square$ pour tout p sera appelé le *mot vide* ; pour chaque $n \in \mathbb{N}^*$, appelons *mot de longueur n* tout élément $(x_p)_{p \geqslant 1}$ de E tel que $(x_1, \dots, x_n) \in \mathcal{A}^n$ et $x_p = \square$ pour $p \geqslant n+1$, et notons $\mathcal{M}_n(\mathcal{A})$ l'ensemble de ces mots ; le sous-ensemble $\mathcal{M}(\mathcal{A}) = \bigcup_{n \geqslant 1} \mathcal{M}_n(\mathcal{A})$ est un sous-ensemble ordonné de E , appelé *ensemble des mots construits sur l'alphabet \mathcal{A}* : l'ordre induit par E sur $\mathcal{M}(\mathcal{A})$ est évidemment l'ordre de classement des mots d'un *dictionnaire des mots de l'alphabet \mathcal{A}* . (On retrouve l'ordre des dictionnaires usuels en prenant $\mathcal{A} = \{A, B, C, \dots, Z\}$ avec l'ordre $A < B < C < \dots < Z$). Pour chaque $n \in \mathbb{N}^*$, on peut identifier $\mathcal{M}_n(\mathcal{A})$ à \mathcal{A}^n , alors l'ordre induit par E sur $\mathcal{M}_n(\mathcal{A})$ est l'ordre lexicographique de \mathcal{A}^n .

Exercice 1 : Soit \mathcal{A} un ensemble d'applications d'un ensemble E dans lui-même et soit $\mathcal{F} = \{X \in \mathcal{P}(E) \mid \forall f \in \mathcal{A}, f(X) \subset X\}$. Montrer que dans \mathcal{F} ordonné pour inclusion, toute partie non vide de \mathcal{F} admet une borne supérieure et une borne inférieure.

Exercice 2 : Soit E et F deux ensembles ; on considère l'ensemble $\mathcal{F}_{E, F}$ des couples (X, f) où $X \subset E$ et où $f : X \rightarrow F$ est une application. Si $(X, f) \in \mathcal{F}_{E, F}$ et $(Y, g) \in \mathcal{F}_{E, F}$ et si

écrit $(X, f) \leq (Y, g)$ ssi on a à la fois $X \subset Y$ et $f = g|_X$. Vérifier que $(\mathcal{F}_{E,F}, \leq)$ est un ensemble ordonné. Quels sont ses éléments maximaux ?

Exercice 3 : Dans l'exercice précédent on prend $E = F = \mathbb{C}$. On note Φ le sous-ensemble de $\mathcal{F}_{\mathbb{C},\mathbb{C}}$ formé des (X, f) tels que f soit *rationnelle*, c'est-à-dire telle qu'il existe $\varphi \in \mathbb{C}(X)$ pour laquelle X ne contient aucun pôle de φ et $\varphi|_X = f$. Trouver les éléments maximaux de l'ensemble ordonné (Φ, \leq) .

Exercice 4 : Soit E un ensemble non vide ; on ordonne $\mathcal{P}(E)$ par inclusion, et on considère une application croissante $f: \mathcal{P}(E) \rightarrow \mathcal{P}(E)$. On note

$$\mathcal{Q}(E) = \{ X \in \mathcal{P}(E) \mid f(X) \subset X \}.$$

Montrer que si $X \in \mathcal{Q}(E)$, alors $f(X) \in \mathcal{Q}(E)$. Montrer que si $X_0 = \bigcap_{X \in \mathcal{Q}(E)} X$, alors $f(X_0) = X_0$.

Exercice 5 : Des hussards de tailles variées sont disposés en rectangle. On repère le plus grand de chaque rangée et l'on retient le plus petit de ces plus grands, soit X . Puis on repère le plus petit de chaque colonne et on retient le plus grand de ces plus petits, soit Y . Comparer les tailles de X et de Y .

Exercice 6 : a) Soit E_1, \dots, E_n des ensembles bien ordonnés non vides ($n \geq 1$). Démontrer que l'ensemble $E = E_1 \times E_2 \times \dots \times E_n$ muni de l'ordre *produit lexicographique*, est bien ordonné ; pour $n = 2$ et chaque E_i égal à \mathbb{N} , montrer que cet ordre n'est pas isomorphe à celui de \mathbb{N} .

b) Montrer que $[0, 1]^{\mathbb{N}}$ ordonné lexicographiquement, n'est pas un ensemble bien ordonné.

c) Soit \mathcal{E} un ensemble bien ordonné de cardinal ≥ 3 , de plus petit élément \square , et soit $\mathcal{A} = \mathcal{E} \setminus \{ \square \}$. Montrer que l'ensemble des mots construits sur l'alphabet \mathcal{A} , ordonné lexicographiquement (cf. remarque 1) n'est pas bien ordonné.

Exercice 7 : On munit \mathbb{R}^n (où $n \geq 1$) de l'ordre *produit* relatif à l'ordre usuel de \mathbb{R} . (Par définition, si $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ appartiennent à \mathbb{R}^n , on a donc $x \leq y$ ssi $(\forall i \in [1, n]) x_i \leq y_i$.)

Soit B le sous-ensemble ordonné de (\mathbb{R}^n, \leq) défini par :

$$B = \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + x_2^2 + \dots + x_n^2 \leq 1 \}.$$

Etudier les éléments maximaux de B .

Exercice 8 : Un ensemble (E, \leq) s'appelle un *treillis* (ou *lattice*) si toute paire d'éléments a et b distincts de E admet une borne supérieure notée $\sup(a, b)$ et une borne inférieure notée $\inf(a, b)$. Montrer que $(\mathcal{P}(\Omega), \subset)$ est un treillis. Représenter graphiquement le treillis des diviseurs de 90 dans $(\mathbb{N}^*, |)$.

Exercice 9 : On montre que, moyennant l'axiome du choix, tout ensemble E peut être bien ordonné (*théorème de Zermelo*) (voir par exemple [12]). Prouver que, réciproquement, tout ensemble bien ordonné satisfait à l'axiome du choix (se reporter à la fin du § 1.3 et définir l'ensemble S).

Soit (E, \leq) un ensemble partiellement ordonné. Une partie non vide C de E est appelée *chaîne* si la relation d'ordre induite sur C par \leq est un ordre total. Si toute chaîne de E possède une borne supérieure, on dit que E est *inductif*. Le *théorème de Zorn* s'énonce alors : dans tout ensemble inductif il existe un élément maximal. On démontre que ce théorème équivaut à l'axiome du choix (cf. [12]).

Chapitre II

NOMBRES ENTIERS, NOMBRES RATIONNELS

§ II.1 AXIOMES DE PEANO ; RÉCURRENCE

La notion de *nombre entier naturel* se présente à l'intuition sous des aspects variés, dont les principaux sont :

— l'aspect *ordinal* : il s'agit de « numéroté » (les pages d'un livre, les abonnés au téléphone, les jours, etc.)

— l'aspect *cardinal* : il s'agit de « compter » le nombre d'éléments de divers « ensembles finis »

— l'aspect *d'ensemble infini* « le plus simple possible » : après n'importe quel entier, il en existe toujours au moins un autre.

En privilégiant certaines de ces propriétés intuitives, et à partir du langage de la théorie des ensembles, on a construit de nombreuses *théories axiomatiques des entiers naturels*. Ci-après, nous allons esquisser les grandes lignes de la construction de ces entiers à partir des *axiomes de Peano*, et rappeler les principales propriétés qui en résultent.

Axiomes de Peano

Nous postulons l'existence d'un triplet $(0, \mathbb{N}, S)$, où \mathbb{N} est un ensemble, 0 un élément particulier de \mathbb{N} , et $S : \mathbb{N} \rightarrow \mathbb{N}$ une application qui vérifient les propriétés suivantes :

(PI) **S est injective**

(PII) **L'image de S est $\mathbb{N} \setminus \{0\}$ (ensemble noté \mathbb{N}^*)**

(PIII) **Si A est une partie de \mathbb{N} telle que $0 \in A$ et**

$$(\forall n \in \mathbb{N}) \quad (n \in A) \Rightarrow (S(n) \in A), \quad \text{alors} \quad A = \mathbb{N}.$$

L'ensemble \mathbb{N} s'appelle *ensemble des entiers naturels* ; l'élément 0 s'appelle **zéro** ; l'application S s'appelle *application successeur* : si $n \in \mathbb{N}$, $S(n)$ est le *successeur* de n . La propriété (PIII) s'appelle **axiome de récurrence**.

Si $n \in \mathbb{N}^*$, d'après (PII) et (PI), n est l'image par S d'un unique élément $n' \in \mathbb{N}$: cet élément n' est appelé le *prédécesseur* de n . A

l'unique élément n'ayant pas de prédécesseur. Le successeur de zéro s'appelle **un** et se note 1. On note encore

$$\begin{aligned} S(1) &= 2, & S(2) &= 3, & S(3) &= 4, & S(4) &= 5, \\ S(5) &= 6, & S(6) &= 7, & S(7) &= 8, & S(8) &= 9. \end{aligned}$$

Remarque 1 : Sur les écrans de calculatrice et les imprimantes d'ordinateur, zéro est noté \emptyset pour éviter des confusions avec la lettre O .

Remarque 2 : Il est aisé de prouver que les axiomes (PI), (PII), (PIII) caractérisent à un isomorphisme près le triplet $(0, \mathbb{N}, S)$. De manière précise, soit $(0', \mathbb{N}', S')$ un autre triplet vérifiant (PI), (PII) et (PIII) ; alors il existe une et une seule application $\varphi : \mathbb{N} \rightarrow \mathbb{N}'$ telle que : $\varphi(0) = 0'$ et $\varphi \circ S = S' \circ \varphi$ (cf. exercice 3), et cette unique application est bijective.

THÉORÈME II.1.1 (dit « de récurrence »)

|| Soit $\mathcal{P}(n)$ une assertion dépendant d'une variable $n \in \mathbb{N}$. Alors les relations $[\mathcal{P}(0)]$ et $[(\forall n \in \mathbb{N}) \mathcal{P}(n) \Rightarrow \mathcal{P}(S(n))]$ impliquent : $(\forall n \in \mathbb{N}) \mathcal{P}(n)$.

Démonstration :

Soit A l'ensemble $\{n, n \in \mathbb{N} \text{ et } \mathcal{P}(n)\}$; A est non vide puisque $0 \in A$ par hypothèse ; de plus, toujours par l'hypothèse, $n \in A \Rightarrow S(n) \in A$; d'après l'axiome (PIII), il s'ensuit que $A = \mathbb{N}$. ■

L'hypothèse $[(\forall n \in \mathbb{N}) \mathcal{P}(n) \Rightarrow \mathcal{P}(S(n))]$ s'exprime en disant que la propriété $\mathcal{P}(n)$ est héréditaire. Ainsi le théorème II.1.1 signifie que si la propriété héréditaire $\mathcal{P}(n)$ est vraie pour $n = 0$, elle est vraie pour tout $n \in \mathbb{N}$: il s'agit donc là de bien plus qu'un simple syllogisme !

Exemple 1 : Montrons que : $(\forall n \in \mathbb{N}) S(n) \neq n$. D'après l'axiome (PII), on a : $S(0) \neq 0$; d'autre part, si $S(n) \neq n$, alors $S(S(n)) \neq S(n)$ d'après l'axiome (PI), donc la propriété étudiée est héréditaire. Par application du théorème II.1.1, elle est donc vraie pour tout $n \in \mathbb{N}$. Pour reconstituer à partir de ce qui précède toute l'Arithmétique usuelle, qui est le fondement de presque toutes les mathématiques, on a besoin de *construire des applications par récurrence*. Dans la pratique, cette construction s'opère implicitement, sans justification, mais il importe d'autant plus de vérifier une fois pour toutes son bien-fondé :

THÉORÈME II.1.2

|| Soit \mathcal{E} un ensemble non vide, $b \in \mathcal{E}$, et $g : \mathcal{E} \rightarrow \mathcal{E}$ une application. Il existe une, et une seule, application $\varphi : \mathbb{N} \rightarrow \mathcal{E}$ telle que : $\varphi(0) = b$ et $\varphi \circ S = g \circ \varphi$.

Démonstration.

1) *Unicité* : Soit $\varphi_1 : \mathbb{N} \rightarrow \mathcal{E}$ et $\varphi_2 : \mathbb{N} \rightarrow \mathcal{E}$ telles que $\varphi_1(0) = \varphi_2(0) = b$ et $\varphi_i \circ S = g \circ \varphi_i$ ($i = 1, 2$). Une application immédiate du théorème II.1.1 montre que $\varphi_1(n) = \varphi_2(n)$ pour tout $n \in \mathbb{N}$, d'où $\varphi_1 = \varphi_2$.

2) *Existence* : On considère l'application

$$h : \mathbb{N} \times \mathcal{E} \rightarrow \mathbb{N} \times \mathcal{E}, (x, y) \mapsto (S(x), g(y)),$$

et l'ensemble \mathcal{F} des $A \subset \mathbb{N} \times \mathcal{E}$ tels que $h(A) \subset A$ et $(0, b) \in A$. Soit $G = \bigcap_{A \in \mathcal{F}} A$.

Nous allons montrer que \dot{G} est le graphe d'une application $\varphi : \mathbb{N} \rightarrow \mathcal{E}$ qui répond à la question. G est bien une partie de $\mathbb{N} \times \mathcal{E}$; \mathcal{F} est non vide, car $\mathbb{N} \times \mathcal{E} \in \mathcal{F}$.

— Montrons d'abord que $\text{pr}_1(G) = \mathbb{N}$: on a d'abord $0 \in \text{pr}_1(G)$ car $(0, b) \in G$. Soit $x \in \text{pr}_1(G)$; choisissons $y \in \mathcal{E}$ tel que $(x, y) \in G$; pour tout $A \in \mathcal{F}$, on a : $(x, y) \in A$, d'où $h(x, y) = (S(x), g(y)) \in A$; d'où $(S(x), g(y)) \in G$ et par suite : $S(x) \in \text{pr}_1(G)$. Donc, par application directe de l'axiome (PIII), $\text{pr}_1(G) = \mathbb{N}$.

— Montrons que G est un graphe fonctionnel ; pour cela notons \mathcal{U} l'ensemble des $x \in \mathbb{N}$ tels que l'ensemble $G_x = \{y \in \mathcal{E} \mid (x, y) \in G\}$ soit un singleton. On voit d'abord que $0 \in \mathcal{U}$; en effet, soit $c \in \mathcal{E}$ tel que $(0, c) \in G$; posons $B = \{(0, b)\} \cup h(G)$, il est clair que $B \in \mathcal{F}$, donc $(0, c) \in B$. Comme $0 \notin \text{Im}(S)$, nécessairement $(0, c) \notin h(G)$, donc $(0, c) = (0, b)$, d'où $c = b$, c'est-à-dire $G_0 = \{b\}$ et $0 \in \mathcal{U}$.

On a déjà dit que $B \in \mathcal{F}$, d'où $G \subset B$. Mais d'autre part il est bien clair que : $h(G) \subset G$, d'où $B \subset G$, et finalement $B = G$. Cela va nous permettre de voir que : $x \in \mathcal{U} \Rightarrow S(x) \in \mathcal{U}$. En effet si $x \in \mathcal{U}$ nous savons que $G_{S(x)} \neq \emptyset$; considérons y_1 et y_2 éléments de $G_{S(x)}$; alors $S(x) \neq 0$, $(S(x), y_i) \in G$ pour $i = 1, 2$ et $G = \{(0, b)\} \cup h(G)$ d'où l'on déduit : $(S(x), y_i) \in h(G)$ pour $i = 1, 2$. Donc $(S(x), y_i) = h(x'_i, y'_i)$ avec $y_i = g(y'_i)$, $S(x'_i) = S(x)$ et $(x'_i, y'_i) \in G$ pour $i = 1, 2$. L'injectivité de S montre que $x'_i = x$ pour chaque i . Alors $(x, y'_i) \in G$, $y'_i \in G_x$, donc $y'_1 = y'_2$ puisque $x \in \mathcal{U}$, et il s'ensuit que $y_1 = y_2$. Donc $G_{S(x)}$ est bien un singleton. Il suffit maintenant d'appliquer (PIII) à \mathcal{U} pour conclure que $\mathcal{U} = \mathbb{N}$ et par conséquent G est bien un graphe fonctionnel.

— Soit alors $\varphi : \mathbb{N} \rightarrow \mathcal{E}$ l'application dont G est le graphe ; puisque $G_0 = \{b\}$, on a $\varphi(0) = b$. Si $x \in \mathbb{N}$, d'après ce qui précède, on a $G_{S(x)} = \{g(\varphi(x))\}$ d'où $\varphi(S(x)) = g(\varphi(x))$, donc φ satisfait aux conditions de l'énoncé II.1.2. ■

Du théorème II.1.2 le lecteur attentif pourra déduire le corollaire suivant dont la démonstration est proposée à l'exercice 1.

COROLLAIRE

Soit \mathcal{E} et \mathcal{F} deux ensembles non vides, $g : \mathcal{F} \rightarrow \mathcal{F}$ et $\psi : \mathcal{E} \rightarrow \mathcal{F}$ deux applications. Il existe une, et une seule, application $\varphi : \mathcal{E} \times \mathbb{N} \rightarrow \mathcal{F}$ vérifiant :

$$(I) \quad (\forall e \in \mathcal{E}) \quad \varphi(e, 0) = \psi(e)$$

$$(II) \quad (\forall e \in \mathcal{E}, \forall n \in \mathbb{N}) \quad \varphi(e, S(n)) = g(\varphi(e, n))$$

Le théorème II.1.2 et son corollaire permettent de définir les opérations de base dans \mathbb{N} , qui sont l'*addition* et la *multiplication* :

Addition dans \mathbb{N}

THÉORÈME II.1.3

Il existe une et une seule application, notée $+$, de $\mathbb{N} \times \mathbb{N}$ dans \mathbb{N} vérifiant :

$$\begin{aligned} \text{(I)} \quad & (\forall n \in \mathbb{N}) \quad n + 0 = n \\ \text{(II)} \quad & (\forall p \in \mathbb{N}) \quad (\forall q \in \mathbb{N}) \quad p + S(q) = S(p + q). \end{aligned}$$

Démonstration :

On applique le corollaire du théorème II.1.2, en prenant $\mathcal{E} = \mathcal{F} = \mathbb{N}$; $g : \mathcal{F} \rightarrow \mathcal{F}$ est l'application S , et $\psi : \mathcal{E} \rightarrow \mathcal{F}$ est l'application $\text{Id}_{\mathbb{N}} : n \mapsto n$. ■

L'application $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, $(p, q) \mapsto p + q$ est une *loi interne* sur \mathbb{N} appelé **addition** ; si $(p, q) \in \mathbb{N} \times \mathbb{N}$, $p + q$ est appelé la **somme** de p et de q . Cette loi interne possède les propriétés suivantes :

(A₁) Associativité :

$$(\forall (p, q, r) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (p + q) + r = p + (q + r)$$

(A₂) Commutativité :

$$(\forall (p, q) \in \mathbb{N} \times \mathbb{N}) \quad p + q = q + p$$

(A₃) Régularité :

$$(\forall (p, q, r) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (p + r = q + r) \Rightarrow p = q$$

ce que l'on exprime en disant que *tout élément $r \in \mathbb{N}$ est régulier pour l'addition.*

$$\text{(A}_4\text{)} \quad (\forall (p, q) \in \mathbb{N} \times \mathbb{N}) \quad (p + q = 0) \Leftrightarrow (p = 0 \text{ et } q = 0)$$

$$\text{(A}_5\text{)} \quad (\forall n \in \mathbb{N}) \quad S(n) = n + 1 (= 1 + n).$$

Toutes ces propriétés se démontrent aisément par récurrence. A titre d'exemple, démontrons (A₅).

On a d'abord $S(0) = 1 = 0 + S(0) = 0 + 1$;
puis $S(1) = S(1 + 0) = 1 + S(0) = 1 + 1$.

Supposons ensuite que $n \in \mathbb{N}$ et que $S(n) = n + 1$; alors

$$S(S(n)) = S(n + 1) = n + S(1) = n + (1 + 1) = (n + 1)$$

(en utilisant l'associativité) $= S(n) + 1$, ce qui prouve que la propriété (A_5) est héréditaire, et comme elle est vraie pour $n = 0$, elle est vraie pour tout $n \in \mathbb{N}$ par récurrence.

On exprime la propriété $0 + n = n + 0 = n$ en disant que 0 est *élément neutre pour l'addition*. Il est évidemment unique.

L'utilisation de la notation additive, et en particulier de la propriété (A_5) , permet de donner leur forme définitive aux théorèmes de récurrence, déjà rencontrés.

THÉORÈME II.1.4

|| Soit $\mathcal{P}(n)$ une assertion dépendant d'une variable $n \in \mathbb{N}$. Alors les relations $[\mathcal{P}(0)]$ et $[(\forall n \in \mathbb{N}) \mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)]$ impliquent :
|| $(\forall n \in \mathbb{N}) \mathcal{P}(n)$.

L'hypothèse $[(\forall n \in \mathbb{N}) \mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)]$ se traduit par l'expression : la propriété $\mathcal{P}(n)$ est héréditaire.

THÉORÈME II.1.5

|| Soit \mathcal{E} un ensemble non vide, $b \in \mathcal{E}$, et $g : \mathcal{E} \rightarrow \mathcal{E}$ une application. Il existe une, et une seule, application $\varphi : \mathbb{N} \rightarrow \mathcal{E}$ telle que : $\varphi(0) = b$ et $(\forall n \in \mathbb{N}) \varphi(n+1) = g(\varphi(n))$.

En d'autres termes une suite d'éléments de \mathcal{E} est parfaitement définie par la donnée de son terme initial u_0 et de la relation de récurrence $u_{n+1} = g(u_n)$.

COROLLAIRE

|| Soit \mathcal{E} et \mathcal{F} deux ensembles non vides, $g : \mathcal{F} \rightarrow \mathcal{F}$ et $\psi : \mathcal{E} \rightarrow \mathcal{F}$ deux applications. Il existe une, et une seule, application $\varphi : \mathcal{E} \times \mathbb{N} \rightarrow \mathcal{F}$ vérifiant :
|| (I) $(\forall e \in \mathcal{E}) \varphi(e, 0) = \psi(e)$
|| (II) $(\forall e \in \mathcal{E}, \forall n \in \mathbb{N}) \varphi(e, n+1) = g(\varphi(e, n))$.

L'usage du théorème II.1.5 permet de définir des suites par récurrence dans des cas où les questions de domaine de définition sont délicates.

Exemple 2 : Soit a, b, c des nombres complexes tels que $a \neq c$ et $b \neq 0$. A partir de $u_0 \in \mathbb{C}$ donné, peut-on définir une suite (u_n) de nombres complexes telle que $(\forall n \in \mathbb{N}) u_{n+1} = a + \frac{b}{u_n - c}$?

Solution : On introduit l'ensemble $\tilde{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ (où ∞ est un élément tel que $\infty \notin \mathbb{C}$) et la bijection $\tilde{g} : \tilde{\mathbb{C}} \rightarrow \tilde{\mathbb{C}}$ telle que $\tilde{g}(z) = a + \frac{b}{z - c}$ pour $z \neq c$, $\tilde{g}(c) = \infty$ et $\tilde{g}(\infty) = a$. On note $g = \tilde{g} |_{\mathbb{C} \setminus \{c\}}$, \tilde{h} la bijection réciproque de \tilde{g} , $h = \tilde{h} |_{\mathbb{C} \setminus \{a\}}$, de sorte que $h(z) = c + \frac{b}{z - a}$ pour $z \neq a$.

Pour pouvoir exploiter le théorème II.1.5 il s'agit de trouver les ensembles $A \subset \mathbb{C} \setminus \{c\}$ tels que $g(A) \subset A$; la réunion \mathcal{E} de tous ces ensembles vérifie encore $g(\mathcal{E}) \subset \mathcal{E}$, et c'est donc le plus grand, pour l'inclusion, de ces ensembles. Nous allons trouver \mathcal{E} .

Si $z \in \mathcal{E}$, on a : $g(z) \in \mathcal{E}$, donc $g(z) \neq c$, donc $z \neq \tilde{h}(c) = h(c)$ en utilisant le fait que g est une bijection de $\mathbb{C} \setminus \{c\}$ sur $\mathbb{C} \setminus \{a\}$ dont la bijection réciproque est h . Par récurrence, on voit que \mathcal{E} ne contient aucun des points de $H = \{\tilde{h}^{\langle k \rangle}(c)\}_{k \in \mathbb{N}}$.

Donc $\mathcal{E} \subset \mathbb{C} \setminus H$; réciproquement, si $z \in \mathbb{C} \setminus H$, montrons que $g(z) \in \mathbb{C} \setminus H$: d'abord $\tilde{g}(z) = g(z)$ car $z \neq c$ (puisque $c \in H$); ensuite, si on avait, pour un certain $k \in \mathbb{N}$, $g(z) = \tilde{h}^{\langle k \rangle}(c)$, on en déduirait

$$\tilde{h} \circ \tilde{g}(z) = z = \tilde{h} \circ g(z) = \tilde{h}^{\langle k+1 \rangle}(z) \in H,$$

contrairement à l'hypothèse, donc $g(z) \notin H$. On a donc prouvé : $\mathcal{E} = \mathbb{C} \setminus H$.

Appliquant le théorème I.2.5, on voit donc : si $b_0 \in \mathbb{C} \setminus H$ est donné, il existe une et une seule suite (u_n) de nombres complexes telle que $u_0 = b_0$, et $(\forall n \in \mathbb{N})$

$$u_{n+1} = a + \frac{b}{u_n - c}.$$

En revanche si b_0 est donné dans $H \cap \mathbb{C}$, l'étude directe de la suite des itérées $\tilde{g}^{\langle k \rangle}(b_0)$ est immédiate et montre qu'il existe un $n_0 \in \mathbb{N}^*$ tel que $\tilde{g}^{\langle n_0 \rangle}(b_0) = c$, mais que $\tilde{g}^{\langle k \rangle}(b_0) = g^{\langle k \rangle}(b_0) \neq c$ pour $k \in \llbracket 0, n_0 - 1 \rrbracket$, et donc la définition d'une suite (u_n) répondant à la question posée avec $u_0 = b_0$ est impossible. On a donc résolu entièrement le problème posé : les u_0 qui conviennent sont les éléments de $\mathbb{C} \setminus H$.

Le problème serait évidemment très différent si la suite recherchée était une application de \mathbb{N} dans \mathbb{C} .

Exemple 3 : Itérées successives d'une application.

Soit \mathcal{E} un ensemble non vide, et $g : \mathcal{E} \rightarrow \mathcal{E}$ une application. Pour tout $x \in \mathcal{E}$, notons $\varphi_x : \mathbb{N} \rightarrow \mathcal{E}$ l'unique application telle que :

$$\varphi_x(0) = x \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad \varphi_x(n+1) = g[\varphi_x(n)].$$

Alors pour chaque entier $k \in \mathbb{N}$, on peut définir l'application $f_k : \mathcal{E} \rightarrow \mathcal{E}$, $x \mapsto \varphi_x(k)$. On a évidemment $f_0 = \text{Id}_{\mathcal{E}}$, et, pour tout $x \in \mathcal{E}$ et $k \in \mathbb{N}^*$, $\varphi_x(k) = f_k(x) = g(g(\dots g(x))\dots)$, expression où figure k fois la lettre g . L'application f_k se note $g^{\langle k \rangle}$, notation que nous avons déjà utilisée dans l'exemple 2, ou même plus simplement g^k , à condition que cela ne risque pas d'entraîner de confusion, et s'appelle *k-ième itérée de g* :

c'est $g \circ g \circ \dots \circ g$ (où figure k fois la lettre g).

On peut définir « globalement » les $g^{\langle k \rangle}$ de la manière suivante : soit \mathcal{F} l'ensemble des applications de \mathcal{E} dans \mathcal{E} ; considérons l'application $G : \mathcal{F} \rightarrow \mathcal{F}$, $f \mapsto g \circ f$. Alors il existe une unique application $\Phi : \mathbb{N} \rightarrow \mathcal{F}$ telle que $\Phi(0) = \text{Id}_{\mathcal{E}}$ et $(\forall n \in \mathbb{N}) \quad \Phi(n+1) = G(\Phi(n))$. On voit que $\Phi(k) = g^{\langle k \rangle}$ pour tout $k \in \mathbb{N}$. A titre d'exercice, on vérifiera que :

$$(\forall (n, p) \in \mathbb{N} \times \mathbb{N}) \quad g^{\langle n+p \rangle} = g^{\langle n \rangle} \circ g^{\langle p \rangle}.$$

Dans la pratique, au lieu de reproduire le raisonnement ci-dessus, qui fait appel au théorème II.1.5, on se contente de la phrase suivante pour introduire les $g^{<k>}$: « soit $g^{<k>}$ ($k \in \mathbb{N}$) la suite d'applications définie par récurrence par : $g^{<0>} = \text{Id}_{\mathcal{E}}$ et $(\forall k \in \mathbb{N}) g^{<k+1>} = g \circ g^{<k>}$ ».

Les applications φ obtenues par application du théorème II.1.5 (resp. de son corollaire) sont dites *définies par récurrence par la relation* $\varphi(n+1) = g(\varphi(n))$ *et la condition initiale* $\varphi(0) = b$ (resp. *par la relation* $\varphi(e, n+1) = g(\varphi(e, n))$ *et la condition initiale* $\varphi(e, 0) = \psi(e)$ *pour tout* $e \in \mathcal{E}$).

Multiplication dans \mathbb{N}

THÉORÈME II.1.6

Il existe une unique application (notée \times , ou $.$, ou sans symbole si cela n'entraîne pas de confusion) de $\mathbb{N} \times \mathbb{N}$ dans \mathbb{N} vérifiant :

(I) $(\forall n \in \mathbb{N}) \quad n \times 0 = 0$

(II) $(\forall (p, q) \in \mathbb{N} \times \mathbb{N}) \quad p \times (q + 1) = (p \times q) + p .$

Démonstration :

Application immédiate du corollaire du théorème II.1.5. ■

L'application \times définie dans le théorème II.1.6 est une loi interne sur \mathbb{N} appelée **multiplication**. Si $(p, q) \in \mathbb{N} \times \mathbb{N}$, l'élément $p \times q$ est appelé *produit* de p par q .

Par récurrence, on démontre aisément les propriétés suivantes :

(M₁) **Distributivité** (du produit par rapport à la somme) :

$$(\forall (p, q, r) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad p(q + r) = pq + pr \quad \text{et} \quad (p + q)r = pr + qr .$$

(M₂) **Associativité** :

$$(\forall (p, q, r) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (pq)r = p(qr) .$$

(M₃) **Commutativité** :

$$(\forall (p, q) \in \mathbb{N} \times \mathbb{N}) \quad pq = qp .$$

(M₄) **Régularité des éléments de \mathbb{N}^*** pour la multiplication :

$$(\forall (p, q, r) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^*) \quad (pr = qr) \Rightarrow (p = q) .$$

Notons enfin que :

$$(\forall p \in \mathbb{N}) \quad p \times 1 = p \times (0 + 1) = (p \times 0) + p = 0 + p$$

d'où (à cause de (M_3)) $p \times 1 = 1 \times p = p$, ce que l'on traduit en disant que *1 est élément neutre pour la multiplication*.

Outre l'addition et la multiplication on peut définir dans \mathbb{N} beaucoup d'autres opérations. Bornons-nous aux plus importantes :

DÉFINITION II.1.1

On appelle **exponentiation** l'unique application de $\mathbb{N} \times \mathbb{N}$ dans \mathbb{N} , (notée $(a, b) \mapsto a^b$) définie par récurrence à l'aide des relations :

$$\begin{aligned} \text{(I)} \quad & (\forall n \in \mathbb{N}) \quad n^0 = 1 \\ \text{(II)} \quad & (\forall (n, p) \in \mathbb{N} \times \mathbb{N}) \quad n^{p+1} = n^p \times n. \end{aligned}$$

C'est toujours le corollaire du théorème II.1.5 qui permet de justifier cette définition. L'exponentiation possède les propriétés suivantes qui se démontrent par récurrence :

$$(E_1) \quad (\forall (n, p, q) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad n^p \times n^q = n^{p+q}$$

$$(E_2) \quad (\forall (n, p, q) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (n^p)^q = n^{pq}$$

$$(E_3) \quad (\forall (n, p, q) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (np)^q = n^q \times p^q.$$

De plus, on voit que $n^1 = n$, que $1^n = 1$ pour tout $n \in \mathbb{N}$, et que $0^n = 0$ pour tout $n \in \mathbb{N}^*$ tandis que $0^0 = 1$.

Soustraction

Si $a \in \mathbb{N}$ et $b \in \mathbb{N}$, en général il n'existe pas d'entier $c \in \mathbb{N}$ tel que $a = b + c$. Mais si un tel c existe, il est nécessairement unique à cause de la propriété (A_3) de l'addition : on l'appelle *différence* de a et de b et on le note $a - b$.

Soit \mathcal{D} la partie de $\mathbb{N} \times \mathbb{N}$ formée des couples (a, b) pour lesquels $a - b$ existe. L'application $\mathcal{D} \rightarrow \mathbb{N}$, $(a, b) \mapsto a - b$ s'appelle soustraction.

Exemple 4 : Soit $n \in \mathbb{N}^*$. Son prédécesseur est $n - 1$.

Quotient exact

Si $(a, b) \in \mathbb{N} \times \mathbb{N}^*$, en général il n'existe pas d'entier c tel que $a = bc$. Mais si un tel entier c existe, il est unique à cause de la propriété (M_4) de la multiplication. On l'appelle dans ce cas le *quotient* de a par b , et on le note $\frac{a}{b}$ ou a/b et on dit alors que *a est divisible par b*. Soit Δ la partie de $\mathbb{N} \times \mathbb{N}^*$ formée des couples (a, b) tels que a/b existe. L'application $\Delta \rightarrow \mathbb{N}$, $(a, b) \mapsto a/b$ s'appelle *division exacte*.

Exercice 1 : Démontrer en détail le corollaire du théorème II.1.2.

Exercice 2 : Prouver en détail les propriétés (A_1) à (A_4) de l'addition,
les propriétés (M_1) à (M_4) de la multiplication,
les propriétés (E_1) à (E_3) de l'exponentiation.

Exercice 3 : On appelle *chaîne simple* tout triplet $\mathcal{C} = (a, \mathcal{E}, f)$ où \mathcal{E} est un ensemble, a un élément de \mathcal{E} , et $f: \mathcal{E} \rightarrow \mathcal{E}$ une application vérifiant (C_1) $a \notin f(\mathcal{E})$; (C_2) f est injective; (C_3) la seule partie E de \mathcal{E} telle que $a \in E$ et $(\forall x \in \mathcal{E}) (x \in E) \Rightarrow (f(x) \in E)$ est $E = \mathcal{E}$.

a) Montrer que si $\mathcal{C} = (a, \mathcal{E}, f)$ est une chaîne simple, on a :

$$\mathcal{E} = \{a\} \cup f(\mathcal{E}).$$

b) Montrer que si $\mathcal{C} = (a, \mathcal{E}, f)$ est une chaîne simple, si \mathcal{F} est un ensemble non vide, b un élément de \mathcal{F} , et $g: \mathcal{F} \rightarrow \mathcal{F}$ une application, alors il existe une, et une seule, application $\varphi: \mathcal{E} \rightarrow \mathcal{F}$ telle que $\varphi(a) = b$ et $\varphi \circ f = g \circ \varphi$ [s'inspirer de la preuve du théorème II.1.2].

c) Montrer que si $\mathcal{C} = (a, \mathcal{E}, f)$ est une chaîne simple, il existe une, et une seule, application $\varphi: \mathbb{N} \rightarrow \mathcal{E}$, telle que $\varphi(0) = a$ et $\varphi(n+1) = f(\varphi(n))$ pour tout $n \in \mathbb{N}$, et montrer que cette application φ est bijective.

Exercice 4 : Montrer que pour tout $n \in \mathbb{N}^*$, le nombre $3^{2n+1} + 2^{n+2}$ est divisible par 7.

Exercice 5 : Montrer que si $x \in \mathbb{N}$ et $n \in \mathbb{N}^*$, $x \neq 0$, $x \neq 1$, si $p \in \mathbb{N}^*$ et si p divise $x^2 - x$, alors p divise $x^n - x$.

Exercice 6 : Résoudre dans $\mathbb{N}^* \times \mathbb{N}^*$ l'équation $x^y = y^x$ avec $x \neq y$.

Exercice 7 : Montrer que $\forall m \in \mathbb{N}^*, \forall n \in \mathbb{N}^*, \forall r \in \mathbb{N}, m^{2r+1} + n^{2r+1}$ est divisible par $m + n$.

Exercice 8 : Etudier pour quels nombres réels a on peut définir une suite (u_n) de réels telle que $u_0 = a$ et $(\forall n \in \mathbb{N}) u_{n+1} = \frac{1}{2} \left(u_n - \frac{1}{u_n} \right)$.

Même question avec $u_0 = a$ et $(\forall n \in \mathbb{N}) u_{n+1} = \frac{1}{|\sqrt{u_n} - 1|}$.

Exercice 9 : Pour quels nombres complexes a peut-on définir une suite (u_n) de nombres complexes telle que $u_0 = a$ et $(\forall n \in \mathbb{N}) u_{n+1} = 1 + \frac{1}{u_n}$?

Exercice 10 : Montrer qu'il existe une suite (u_n) de réels et une seule telle que $u_0 = 1$ et $(\forall n \in \mathbb{N}) u_{n+1} = \frac{-1}{1 + u_n}$.

§ II.2 ORDRE NATUREL DANS \mathbb{N}

THÉORÈME II.2.1

|| Sur \mathbb{N} , la relation binaire (notée \leq) définie pour tous $a, b \in \mathbb{N}$ par
|| $(a \leq b) \Leftrightarrow (\exists d \in \mathbb{N} \mid b = a + d)$ est une relation d'ordre total.

Démonstration (abrégée) :

Pour la relation \leq :

La *réflexivité* découle du fait que $a = a + 0$ pour tout $a \in \mathbb{N}$.

La *transitivité* découle de l'associativité de l'addition.

L'antisymétrie résulte des propriétés (A_3) et (A_4) de l'addition.

Par suite, la relation \leq est une relation d'ordre sur \mathbb{N} . Montrons que cet ordre est *total* ; soit $p \in \mathbb{N}$, il s'agit de voir que : $(\forall q \in \mathbb{N}) p \leq q$ ou $q \leq p$. On le vérifie par récurrence sur p . En premier lieu, si $p = 0$ on a évidemment, $q = 0 + q$ pour tout $q \in \mathbb{N}$, d'où $0 \leq q$. Il reste à prouver que la propriété « p est comparable à tout naturel » est héréditaire. Supposons cette propriété vraie à l'ordre p . Il est certain que $0 \leq p + 1$.

Si $q \neq 0$ désignons par q' son prédécesseur. L'hypothèse de récurrence montre que $q' \leq p$ ou $q' \geq p$. Examinons séparément les 3 cas possibles : si $q' = p$, alors $q = p + 1$; si $q' > p$, c'est qu'il existe $d \in \mathbb{N}^*$ tel que $q' = p + d$, alors $q = (p + d) + 1 = (p + 1) + d$ et $q > p + 1$. De même si $q' < p$, alors $q < p + 1$. La propriété est donc vraie à l'ordre $p + 1$. Etant vraie à l'ordre 0, elle est établie par récurrence. ■

La relation d'ordre ainsi définie s'appelle **ordre naturel** sur \mathbb{N} ; $n \leq p$ se lit « n est inférieur ou égal à p ». La relation ($n \leq p$ et $n \neq p$) s'écrit $n < p$ (notation déjà utilisée dans la démonstration précédente) et se lit « n est *strictement* inférieur à p ». On vérifie aisément que

$$(n < p) \Leftrightarrow (n + 1 \leq p).$$

Par des raisonnements par récurrence immédiats on montre que :

PROPOSITION II.2.1

L'ordre naturel de \mathbb{N} :

(I) est **compatible avec l'addition**, c'est-à-dire vérifie

$$(\forall (m, n, p) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad (m \leq n) \Leftrightarrow (m + p \leq n + p)$$

(II) est **compatible avec la multiplication par un naturel non nul**, c'est-à-dire :

$$(\forall (m, n, p) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^*) \quad (m \leq n) \Leftrightarrow (mp \leq np).$$

On déduit notamment de cette proposition que si $(m, n) \in \mathbb{N} \times \mathbb{N}$, alors

$$(mn = 1) \Leftrightarrow (m = 1 \text{ et } n = 1).$$

THÉORÈME II.2.2

L'ensemble ordonné (\mathbb{N}, \leq) est **bien ordonné**, c'est-à-dire : toute partie non vide de \mathbb{N} admet un plus petit élément.

Démonstration :

Soit A une partie non vide de \mathbb{N} . L'ensemble M des minorants de A est non vide puisque $0 \in M$. Montrons que $M \neq \mathbb{N}$. En effet soit a un élément de A , alors $a + 1 \notin M$, donc $M \neq \mathbb{N}$. Puisque $0 \in M$, il s'ensuit que l'implication $(\forall n \in \mathbb{N}) (n \in M) \Rightarrow (n + 1 \in M)$ e

donc $n_0 \in M$ tel que $n_0 + 1 \notin M$. Si l'on avait $n_0 \notin A$, il s'ensuivrait $n_0 < a$ pour tout $a \notin A$, d'où $n_0 + 1 \leq a$ pour tout $a \notin A$, d'où $n_0 + 1 \in M$, ce qui est absurde. Donc $n_0 \in A$.

Finalement, $n_0 \in A$ et $n_0 \in M$, donc n_0 est le plus petit élément de A . ■

On notera que le plus petit élément de \mathbb{N} lui-même est 0.

COROLLAIRE (récurrence transfinie dans \mathbb{N})

|| Soit A une partie de \mathbb{N} telle que
 || (RT) $(\forall n \in \mathbb{N}) \quad [(\forall p \in \mathbb{N}) (0 \leq p < n) \Rightarrow (p \in A)] \Rightarrow [n \in A]$.
 || Alors $A = \mathbb{N}$.

Ce corollaire se prouve par l'absurde, en considérant $B = \mathbb{N} \setminus A$. Si B était non vide, son plus petit élément contredirait (RT).

Remarquons que (RT) $\Rightarrow 0 \in A$ (cependant dans la pratique, on vérifie que $0 \in A$).

Remarque 1 : Il existe des ensembles bien ordonnés non isomorphes à l'ensemble bien ordonné (\mathbb{N}, \leq) (cf. l'exercice 2). Or, dans tout ensemble bien ordonné E , une partie A qui vérifie (RT) est égale à E (même preuve que pour le corollaire ci-dessus). Donc ce théorème de récurrence transfinie n'est pas particulier à (\mathbb{N}, \leq) .

Exemple 1 (principe de « descente infinie » de Fermat).

Il n'existe aucune suite $(u_k)_{k \in \mathbb{N}}$ strictement décroissante dans \mathbb{N} .

En effet, si $(u_k)_{k \in \mathbb{N}}$ était une telle suite, l'ensemble $U = \{u_k\}_{k \in \mathbb{N}}$ de ses valeurs serait non vide et n'aurait pas de plus petit élément, en contradiction avec le théorème II.2.2. Ce principe de descente infinie est évidemment valable dans tout ensemble bien ordonné, mais il semble que Fermat soit le premier à l'avoir utilisé pour démontrer des théorèmes d'Arithmétique, à partir de 1638. Nous en donnerons plus loin quelques exemples.

THÉORÈME II.2.3

|| Toute partie non vide et majorée de \mathbb{N} admet un plus grand élément.

Démonstration :

Soit A une partie non vide et majorée de \mathbb{N} . L'ensemble M des majorants de A est non vide par hypothèse. Cet ensemble M admet donc un plus petit élément, notons-le a : si $a = 0$, alors nécessairement $A = \{0\}$ et le théorème est prouvé. Sinon $a - 1$ existe et $a - 1 \notin M$: le seul élément b de A tel que $b > a - 1$ est alors $b = a$, d'où $a \in A$, et par suite a est bien le plus grand élément de A . ■

On peut montrer que les théorèmes II.2.2 et II.2.3 caractérisent à eux deux l'ensemble ordonné (\mathbb{N}, \leq) à un isomorphisme près (cf

On définit les **intervalles de \mathbb{N}** : par définition, un intervalle de \mathbb{N} est une partie I de \mathbb{N} telle que : $(\forall a \in I, \forall b \in I, \forall n \in \mathbb{N}) (a \leq n \leq b) \Rightarrow (n \in I)$. A l'aide des théorèmes II.2.2 et II.2.3, on voit que tout intervalle non vide I est de l'un des deux types suivants :

$$[a, b] = \{n \in \mathbb{N} \mid a \leq n \leq b\} \quad \text{avec } a \in \mathbb{N}, b \in \mathbb{N} \text{ et } a \leq b$$

ou
$$[a, \rightarrow[= \{n \in \mathbb{N} \mid n \geq a\} \quad \text{avec } a \in \mathbb{N}.$$

On remarque que si $a > b$, $[a, b]$ est l'ensemble vide ; si $a = b$, $[a, b]$ est le singleton $\{a\}$. On convient, pour $b \in \mathbb{N}^*$, de noter $[a, b[$ l'intervalle $[a, b - 1]$ et $]a, b[$ l'intervalle $[a + 1, b - 1]$.

On note également $]a, b]$ l'intervalle $[a + 1, b]$ et $]a, \rightarrow[$ l'intervalle $[a + 1, \rightarrow[$. Il convient de remarquer que \mathbb{N} n'a pas de plus grand élément : en effet, pour tout $n \in \mathbb{N}$, on a $n + 1 > n$. Donc aucun entier ne majore \mathbb{N} .

Nous allons maintenant donner quelques variantes utiles du théorème de récurrence. Rappelons qu'une assertion $\mathcal{P}(n)$ dépendant de l'entier $n \in \mathbb{N}$, est dite *héréditaire* ssi $(\forall n \in \mathbb{N}) \mathcal{P}(n) \Rightarrow \mathcal{P}(n + 1)$. Cela dit :

THÉORÈME II.2.4

|| Soit $a \in \mathbb{N}$, et A une partie de \mathbb{N} telle que $a \in A$ et que $(\forall n \in \mathbb{N}) n \in A \Rightarrow n + 1 \in A$. Alors $[a, \rightarrow[\subset A$. En conséquence, si une assertion $\mathcal{P}(n)$ héréditaire est vraie pour $n = a$, elle est vraie pour tout $n \in [a, \rightarrow[$.

Démonstration :

La deuxième assertion découle de la première, en considérant l'ensemble $A = \{n \in \mathbb{N} \mid \mathcal{P}(n)\}$. Pour montrer la première assertion, on observe que l'application $p \mapsto p + a$ est une bijection de \mathbb{N} sur $[a, \rightarrow[$; la bijection réciproque envoie $A \cap [a, \rightarrow[$ sur une partie E de \mathbb{N} telle que $0 \in E$ et $(\forall k \in \mathbb{N}) (k \in E) \Rightarrow (k + 1 \in E)$, d'où $E = \mathbb{N}$, et par suite $[a, \rightarrow[\subset A$. ■

THÉORÈME II.2.5

|| Soit $a, b \in \mathbb{N}$ avec $a \leq b$; si une partie A de \mathbb{N} vérifie $a \in A$ et $(\forall n \in \mathbb{N}) (n \in A \text{ et } n < b) \Rightarrow (n + 1 \in A)$, alors $[a, b] \subset A$. En conséquence si une assertion $\mathcal{P}(n)$ dépendant de $n \in \mathbb{N}$ est vraie pour $n = a$ et si $(\mathcal{P}(n) \text{ et } n < b) \Rightarrow \mathcal{P}(n + 1)$, alors $\mathcal{P}(n)$ est vraie pour tout $n \in [a, b]$ (récurrence finie).

Démonstration :

Comme pour le théorème II.2.4, il suffit de vérifier la première assertion. On le fait par l'absurde : supposons $[a, b] \not\subset A$, alors l'ensemble $B = [a, b] \setminus A$ serait non vide, donc admettrait un plus petit élément, noté b' ; on aurait $b > a$, car $a \in A$; d'où b'

$b' - 1 < b' \leq b$, par suite $b' - 1 \in A \cap \llbracket a, b \rrbracket$, d'où $b' = (b' - 1) + 1 \in A$ à cause de l'hypothèse. Contradiction. Donc $\llbracket a, b \rrbracket \subset A$. ■

COROLLAIRE (récurrence descendante)

Soit $a \in \mathbb{N}$, $b \in \mathbb{N}^*$, avec $a \leq b$. Si une partie A de \mathbb{N} vérifie $b \in A$ et $(\forall n \in \mathbb{N}) (n \in A \text{ et } n > a) \Rightarrow n - 1 \in A$, alors $\llbracket a, b \rrbracket \subset A$.
En conséquence, si une assertion $\mathcal{P}(n)$ est vraie pour $n = b$ et vérifie $(\forall n \in \mathbb{N}) (n > a \text{ et } \mathcal{P}(n)) \Rightarrow \mathcal{P}(n - 1)$, alors $\mathcal{P}(n)$ est vraie pour tout $n \in \llbracket a, b \rrbracket$.

Ce corollaire se démontre à partir du théorème II.2.5 en considérant la bijection strictement décroissante $k \mapsto b - k$ de $\llbracket 0, b - a \rrbracket$ sur $\llbracket a, b \rrbracket$.

Division euclidienne

Observons d'abord que (\mathbb{N}, \leq) vérifie la *propriété d'Archimède* ⁽¹⁾ : si $(a, b) \in \mathbb{N} \times \mathbb{N}^*$, il existe $n \in \mathbb{N}$ tel que $nb \geq a$ (en effet si $a = 0$, $n = 1$ convient ; si $a \geq 1$, $n = a$ convient puisque $b \geq 1 \Rightarrow ab \geq a$).

THÉORÈME II.2.6

Soit $(a, b) \in \mathbb{N} \times \mathbb{N}^*$. Il existe un couple (q, r) et un seul d'entiers naturels tels que :

(I) $a = bq + r$;
(II) $0 \leq r < b$.

Démonstration :

a) *Unicité.* Supposons trouvés deux couples (q_1, r_1) , (q_2, r_2) qui conviennent, d'où $a = bq_1 + r_1 = bq_2 + r_2$, $0 \leq r_1 < b$, $0 \leq r_2 < b$. Puisque $q_1 = q_2 \Rightarrow r_1 = r_2$, si ces couples étaient distincts, on aurait $q_1 \neq q_2$, donc par exemple $q_1 > q_2$; d'où $b(q_1 - q_2) + r_1 = r_2$; mais on aurait alors $q_1 - q_2 \geq 1$, donc $b(q_1 - q_2) + r_1 \geq b + r_1 \geq b$ alors que $r_2 < b$. On voit la contradiction ; par suite $q_1 = q_2$ et $r_1 = r_2$.

b) *Existence.* Par la propriété d'Archimède, l'ensemble M des $c \in \mathbb{N}$ tels que $cb > a$ est non vide ; soit q' le plus petit élément de M : comme $q'b > a$, nécessairement $q' \geq 1$. Posons alors $q = q' - 1$. On a : $q \notin M$, d'où $bq \leq a < b(q + 1) = bq + b$. Donc l'élément $r = a - bq$ vérifie $0 \leq r < b$ et on a bien $a = bq + r$; le couple (q, r) convient. ■

Pour achever ce paragraphe, notons que le *théorème II.1.5* et son *corollaire* s'étendent sans difficulté en remplaçant \mathbb{N} par un intervalle non vide quelconque I de \mathbb{N} . Dans la suite, nous emploierons ces extensions du théorème II.1.5 sous le nom général « *d'objets définis par récurrence* » (finie ou non).

⁽¹⁾ *Archimède* (287-212 av. J.-C.) est l'un des plus grands mathématiciens de l'Antiquité. On lui doit aussi le principe fondamental de l'hydrostatique.

Exercice 1 : Soit (E, \leq) un ensemble totalement ordonné non vide tel que : (I) (E, \leq) est bien ordonné ; (II) E n'a pas de plus grand élément ; (III) toute partie de E non vide et majorée admet un plus grand élément. Montrer qu'il existe une, et une seule, bijection croissante ψ de (\mathbb{N}, \leq) sur (E, \leq) .

Indication pour l'existence de ψ : définir $\psi(0) = \min_{x \in E} (x)$, puis définir par récurrence $\psi(n+1) = \min (E \setminus \{\psi(0), \psi(1), \dots, \psi(n)\})$ en justifiant.

Exercice 2 : a) Soit p un entier, $p > 1$; on considère l'ensemble $E^{[p]}$, somme de p ensembles E_i ($1 \leq i \leq p$), chaque E_i étant égal à \mathbb{N} . On munit $E^{[p]}$ de la relation d'ordre suivante, notée \leq : si x et $y \in E_i$ pour un certain i , alors $x \leq y$ ssi $x \leq y$ dans \mathbb{N} ; si $x \in E_i$ et $y \in E_j$ avec $i \neq j$, alors $x \leq y$ ssi $i < j$. Montrer que $(E^{[p]}, \leq)$ est un ensemble bien ordonné, dans lequel la partie E_1 est non vide, majorée mais sans plus grand élément, et que par suite $(E^{[p]}, \leq)$ n'est pas isomorphe à (\mathbb{N}, \leq) .

b) Etendre la construction précédente en y remplaçant $\llbracket 1, p \rrbracket$ par \mathbb{N}^* .

c) Plus généralement, soit (I, \mathcal{R}) un ensemble bien ordonné non vide, et pour chaque $i \in I$, (E_i, \mathcal{R}_i) un ensemble bien ordonné non vide. Sur l'ensemble somme $E^{[I]} = \coprod_{i \in I} E_i$, on

définit la relation binaire \leq ainsi : si $x, y \in E^{[I]}$, et si x et y sont dans un même E_i , alors $x \leq y$ ssi $x \mathcal{R}_i y$; si $x \in E_i$ et $y \in E_j$ avec $i \neq j$, alors $x \leq y$ ssi $i \mathcal{R} j$. Montrer que $(E^{[I]}, \leq)$ est un ensemble bien ordonné.

Exercice 3 : a) Soit (E, \leq) un ensemble bien ordonné non vide sans plus grand élément. Montrer que l'application $S : E \rightarrow E$, $x \mapsto \min \{y \in E \mid y > x\}$ est bien définie (c'est l'application « successeur »). Si 0 désigne le plus petit élément de E , montrer qu'en général $\text{Im}(S) \neq E \setminus \{0\}$ (cf. exercice 2).

b) On note $E' = (E - \{0\}) \setminus \text{Im}(S)$. En remarquant que tout sous-ensemble ordonné d'un ensemble bien ordonné est bien ordonné, étudier si on peut définir $E'', \dots, E^{(k)} = (E^{(k-1)})'$, pour $k \in \mathbb{N}^*$.

c) Donner un exemple d'ensemble bien ordonné E tel que tous les $E^{(k)}$ soient définis pour $k \in \mathbb{N}^*$ (on pourra utiliser l'exercice 2b).

Exercice 4 : Soit (E, \leq) un ensemble bien ordonné non vide sans plus grand élément. On définit l'ensemble E' comme dans l'exercice 3 au b).

a) Si $E' = \emptyset$, (E, \leq) est isomorphe à (\mathbb{N}, \leq) ;

b) Si E' est fini de cardinal $p \geq 1$, alors E est isomorphe à $E^{[p+1]}$ construit dans l'exercice 2 au a). c) On suppose $E' = I$ non vide et non majoré ; pour $x \in E$, soit $i(x) = \min \{y \in E' \mid y > x\}$. Si $i \in I$, soit $E_i = \{x \in E \mid i(x) = i\}$. Montrer que chaque sous-ensemble ordonné E_i de E est isomorphe à (\mathbb{N}, \leq) , et montrer que E est isomorphe à l'ensemble ordonné $E^{[I]}$ construit dans l'exercice 2, à la question c).

Exercice 5 : Soit $f : \mathbb{N}^* \rightarrow \mathbb{N}^*$ une application qui vérifie la propriété $(\forall n \in \mathbb{N}^*) f(n+1) > f(f(n))$. Montrer que $f(n) = n$ ($\forall n \in \mathbb{N}^*$). [Olympiades 1977.]

§ II.3 ENSEMBLES FINIS, ENSEMBLES INFINIS ; ENSEMBLES DÉNOMBRABLES

DÉFINITION II.3.1

$\{$ Un ensemble non vide E est dit **fini** ssi il existe un entier $p \in \mathbb{N}^*$ et une bijection de E sur $\llbracket 1, p \rrbracket$. Par convention on dit aussi que l'ensemble vide est fini. Un ensemble non fini e

Commençons par développer certaines propriétés des intervalles de \mathbb{N} puisque ce sont eux qui servent de référence.

THÉORÈME II.3.1

|| Soit p et $q \in \mathbb{N}^*$. Pour qu'il existe une application injective de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, q \rrbracket$, il faut et il suffit que $p \leq q$.

Démonstration :

a) Si $p \leq q$ on a : $\llbracket 1, p \rrbracket \subset \llbracket 1, q \rrbracket$, et alors il est clair que l'injection canonique $x \mapsto x$ de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, q \rrbracket$ est injective.

b) Montrons par récurrence sur p la propriété $\mathcal{P}(p)$: pour $q \in \mathbb{N}^*$, s'il existe une application injective de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, q \rrbracket$, alors $p \leq q$. Pour $p = 1$ il n'y a rien à prouver, car alors $\llbracket 1, p \rrbracket$ est un singleton. Supposons $p \geq 1$ et $\mathcal{P}(p)$ vraie. Soit $q \in \mathbb{N}^*$ et $f : \llbracket 1, p+1 \rrbracket \rightarrow \llbracket 1, q \rrbracket$ une application injective. Notons que $1 \neq p+1$, donc $\llbracket 1, q \rrbracket$ n'est pas un singleton, puisque f est injective, donc $q \geq 2$.

1^{er} cas : $f(p+1) = q$. Alors, par injectivité de f , $f(\llbracket 1, p \rrbracket) \subset \llbracket 1, q-1 \rrbracket$. $f|_{\llbracket 1, p \rrbracket}$ est une injection de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, q-1 \rrbracket$, donc $q-1 \geq p$ à cause de l'hypothèse de récurrence, d'où $q \geq p$.

2^e cas : $f(p+1) < q$. Soit alors $g : \llbracket 1, q \rrbracket \rightarrow \llbracket 1, q \rrbracket$ la bijection telle que

$$g(q) = f(p+1), \quad g[f(p+1)] = q$$

et $g(i) = i$ pour $i \neq q, i \neq f(p+1)$. Alors $g \circ f : \llbracket 1, p+1 \rrbracket \rightarrow \llbracket 1, q \rrbracket$ est une injection, et $g \circ f(p+1) = q$ d'où $q \geq p+1$ par le 1^{er} cas.

On a donc prouvé $\mathcal{P}(p)$ pour tout $p \in \mathbb{N}^*$ par récurrence sur p . ■

COROLLAIRE 1

|| Soit p et q dans \mathbb{N}^* . S'il existe une **surjection** f de $\llbracket 1, p \rrbracket$ sur $\llbracket 1, q \rrbracket$, on a : $p \geq q$.

Démonstration :

Pour chaque $k \in \llbracket 1, q \rrbracket$, l'ensemble $E_k = f^{-1}(k)$ est non vide, donc admet un plus petit élément, que nous notons $g(k)$. On a ainsi construit $g : \llbracket 1, q \rrbracket \rightarrow \llbracket 1, p \rrbracket$ et il est clair que g est injective, d'où $p \geq q$. ■

COROLLAIRE 2

|| Soit p et q dans \mathbb{N}^* . Pour qu'il existe une bijection de $\llbracket 1, p \rrbracket$ sur $\llbracket 1, q \rrbracket$, il faut et il suffit que $p = q$.

C'est immédiat, compte tenu du théorème II.3.1, de son corollaire 1, et du fait que si $p = q$, l'application $x \mapsto x$ de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, p \rrbracket$ est bijective.

THÉORÈME II.3.2

|| Soit $p \in \mathbb{N}^*$. Toute application injective (resp. surjective) de $\llbracket 1, p \rrbracket$ dans lui-même est bijective.

Démonstration :

a) Soit $f : \llbracket 1, p \rrbracket \rightarrow \llbracket 1, p \rrbracket$ une injection. Montrons par l'absurde que f est surjective. Si ce n'était pas le cas, on pourrait choisir $a \in \llbracket 1, p \rrbracket \setminus \text{Im}(f)$, d'où $p > 1$. On définirait alors

$$g : \llbracket 1, p \rrbracket \rightarrow \llbracket 1, p - 1 \rrbracket$$

ainsi : $g(n) = f(n)$ si $f(n) < a$, $g(n) = f(n) - 1$ si $f(n) > a$, et g serait une injection de $\llbracket 1, p \rrbracket$ dans $\llbracket 1, p - 1 \rrbracket$, ce qui contredirait le théorème II.3.1.

b) Soit $f : \llbracket 1, p \rrbracket \rightarrow \llbracket 1, p \rrbracket$ une surjection. On peut alors définir $g : \llbracket 1, p \rrbracket \rightarrow \llbracket 1, p \rrbracket$ par $g(k) = \text{Min} \{x \in \llbracket 1, p \rrbracket, f(x) = k\}$ (cf. preuve du corollaire 1 du théorème II.3.1), et g est injective, donc bijective d'après le a) ; le fait que g est bijective et que $f \circ g = \text{Id}_{\llbracket 1, p \rrbracket}$ montre que f est bijective. ■

THÉORÈME II.3.3

|| Si E est un ensemble fini non vide, il existe un **unique** $p \in \mathbb{N}^*$ pour lequel on puisse trouver une bijection de E sur $\llbracket 1, p \rrbracket$.

Démonstration :

En effet, soit p et q deux tels entiers, et $f : E \rightarrow \llbracket 1, p \rrbracket$, $g : E \rightarrow \llbracket 1, q \rrbracket$ des bijections. Alors $g \circ f^{-1}$ est une bijection de $\llbracket 1, p \rrbracket$ sur $\llbracket 1, q \rrbracket$, d'où $p = q$. ■

DÉFINITION II.3.2

⎵ Si E est un ensemble fini non vide, on appelle **cardinal** de E , et on note $\text{card}(E)$, l'unique $p \in \mathbb{N}^*$ tel qu'il existe une bijection de E sur $\llbracket 1, p \rrbracket$. Si $E = \emptyset$, on convient que $\text{card}(E) = 0$.

De l'étude précédente on déduit facilement des propriétés des ensembles finis, par exemple :

THÉORÈME II.3.4

|| Soit E et F deux ensembles finis, de cardinaux p et q :

(I) $p \leq q$ ssi il existe une injection de E dans F ;

(II) $p \leq q$ ssi il existe une surjection de F sur E ;

(III) $p = q$ ssi il existe une bijection de E sur F

Comme conséquence, si l'ensemble E est fini, et si $f : E \rightarrow E$ est une application, il y a équivalence entre les 3 propriétés suivantes : (I) f est injective ; (II) f est surjective ; (III) f est bijective.

Notons dès maintenant que l'ensemble \mathbb{N} est infini, puisque l'application successeur : $\mathbb{N} \rightarrow \mathbb{N}$, $n \mapsto n + 1$ est injective et non surjective.

THÉORÈME II.3.5

|| Si E est un ensemble fini et si $F \subset E$, alors F est fini, et $\text{card}(F) \leq \text{card}(E)$; de plus on a $\text{card}(E) = \text{card}(F)$ ssi $F = E$.

Les parties finies de \mathbb{N} sont les sous-ensembles majorés dans \mathbb{N} .

THÉORÈME II.3.6

|| Si E et F sont deux ensembles finis, alors $E \cup F$ est fini et on a : $\text{card}(E \cup F) = \text{card}(E) + \text{card}(F) - \text{card}(E \cap F)$.

Démonstration (abrégée) :

Il suffit de prouver le théorème pour des ensembles disjoints car $E \cup F$ peut toujours s'écrire comme réunion des ensembles disjoints $E \setminus (E \cap F)$, $F \setminus (E \cap F)$ et $E \cap F$. Supposons donc E et F disjoints, de cardinaux respectifs p et q , que l'on peut supposer ≥ 1 , le cas où E serait vide étant évident. Soit alors $f : E \rightarrow \llbracket 1, p \rrbracket$ et $g : F \rightarrow \llbracket 1, q \rrbracket$ des bijections. Alors l'application $h : E \cup F \rightarrow \llbracket 1, p + q \rrbracket$, $x \mapsto f(x)$ si $x \in E$, $x \mapsto p + g(x)$ si $x \in F$, est une bijection, d'où la conclusion que $E \cup F$ est fini, de cardinal $p + q$. ■

Si E et F sont des ensembles et si $f : E \rightarrow F$ est surjective, alors E fini $\Rightarrow F$ fini, et $\text{card}(F) \leq \text{card}(E)$, l'égalité ayant lieu ssi f est bijective. Si f est injective, alors $(F \text{ fini}) \Rightarrow E$ fini, l'égalité ayant lieu ssi f est bijective.

THÉORÈME II.3.7

|| Si E et F sont des ensembles finis, alors $E \times F$ est fini, et on a : $\text{card}(E \times F) = \text{card}(E) \times \text{card}(F)$.

Démonstration (abrégée) :

Soit $n = \text{card}(E)$, $p = \text{card}(F)$; si $n = 0$ ou $p = 0$, $E \times F = \emptyset$ et il n'y a rien à prouver ; si $n \geq 1$ et $p \geq 1$, on raisonne par récurrence sur p . D'abord si $p = 1$, E et $E \times F$ sont en bijection, d'où le résultat ; ensuite en supposant le théorème vrai à l'ordre p , soit $p + 1$ le cardinal de F : on choisit $a \in F$, et alors $E \times F$ est union disjointe de $E \times (F \setminus \{a\})$ et de $E \times \{a\}$, d'où par le théorème II.3.6 :

$$\text{card}(E \times F) = \text{card}(E \times (F \setminus \{a\})) + \text{card}(E \times \{a\})$$

en appliquant l'hypothèse de récurrence et le fait que $E \times \{a\}$ est en bijection avec E , cela fournit

$$\text{card}(E \times F) = np + n = n(p + 1) = \text{card}(E) \times \text{card}(F)$$

et le théorème est démontré. ■

Ensembles infinis

DÉFINITION II.3.3

Deux ensembles E et F sont dits **équipotents** ssi il existe au moins une bijection de E sur F .

Tout ensemble est équipotent à lui-même ; E et F sont équipotents ssi F et E le sont ; si E est équipotent à F et F équipotent à G , alors E est équipotent à G (voir Chap. I, § 3).

Si E et F sont équipotents, alors E est fini ssi F l'est, et dans ce cas on a : $\text{card}(E) = \text{card}(F)$.

THÉORÈME II.3.8

Pour qu'un ensemble E soit infini, il faut et il suffit qu'il existe une injection de \mathbb{N} dans E .

Démonstration :

Puisque \mathbb{N} est infini, s'il existe une injection de \mathbb{N} dans E (c'est-à-dire une bijection de \mathbb{N} sur une partie de E), E est infini d'après tout ce qui précède. Réciproquement, supposons E infini. Nous nous contenterons de la preuve suivante : choisissons $a \in E$ et construisons par récurrence une application $f : \mathbb{N} \rightarrow E$ obéissant à la loi suivante : $f(0) = a$, et pour tout $n \in \mathbb{N}$, $f(n+1)$ est un élément « arbitraire » de $E \setminus \{f(0), f(1), \dots, f(n)\} = E_n$ (la possibilité de construire f provient de ce que E est infini, d'où, à tout rang, $E_n \neq \emptyset$, et de l'axiome du choix). Alors f est une injection de \mathbb{N} dans E . ■

THÉORÈME II.3.9

Si F est une partie infinie de \mathbb{N}^* , il existe une, et une seule, bijection croissante de F sur \mathbb{N}^* .

Démonstration (abrégée) :

L'unicité vient de ce que l'unique bijection croissante de \mathbb{N}^* sur \mathbb{N}^* est $\text{Id}_{\mathbb{N}^*}$. On construit une bijection croissante $f : \mathbb{N}^* \rightarrow E$ par récurrence de la manière suivante : on pose $f(1) = \text{Min}(F)$. Supposant construits $f(1), f(2), \dots, f(n)$ pour $n \geq 1$, l'ensemble $F \setminus \{f(1), f(2), \dots, f(n)\}$ est non vide puisque F est infini, et on pose :

$$f(n+1) = \text{Min}(F \setminus \{f(1), f(2), \dots, f(n)\}).$$

On vérifie successivement que f est strictement croissante (en raisonnant par l'absurde), donc injective, puis que f est surjective (toujours en raisonnant par l'absurde). ■

En 1904, Zermelo a démontré qu'étant donnés deux enser

quelconques, l'une au moins des deux assertions suivantes est vraie : « A est équipotent à une partie de B » ou « B est équipotent à une partie de A ». Bornons-nous à établir :

THÉORÈME II.3.10 (Bernstein ⁽¹⁾, 1897)

|| Soit A et B deux ensembles. S'il existe une injection $f : A \rightarrow B$ et une injection $g : B \rightarrow A$, alors A et B sont équipotents.

Démonstration :

Soit $B' = f(A)$ et $A' = g(B)$. On pose (en convenant que $(g \circ f)^{<0>} = \text{Id}_A$, $(f \circ g)^{<0>} = \text{Id}_B$) :

$$X = \bigcup_{n \in \mathbb{N}} (g \circ f)^{<n>} (A \setminus A'), Y = \bigcup_{n \in \mathbb{N}} (f \circ g)^{<n>} (B \setminus B').$$

Alors $X \subset A$ et $Y \subset B$; si $x \in A \setminus X$, a fortiori $x \in A \setminus (A \setminus A') = A'$, donc x est l'image par g d'un unique $z \in B$. Posons : si $x \in X$, $\varphi(x) = f(x)$; si $x \in A \setminus X$, $\varphi(x) =$ l'unique $z \in B$ tel que $g(z) = x$. On a ainsi défini une application $\varphi : A \rightarrow B$. De même on définit $\psi : B \rightarrow A$ par $\psi(x) = g(x)$ si $x \in Y$ et $\psi(x) =$ l'unique $z \in A$ tel que $f(z) = x$ si $x \in B \setminus Y$. On contrôle facilement que $\psi \circ \varphi = \text{Id}_A$ et que $\varphi \circ \psi = \text{Id}_B$, donc φ et ψ sont des bijections réciproques l'une de l'autre, et A est bien équipotent à B . ■

Le lecteur pourra faire un dessin pour bien comprendre la dynamique de la démonstration précédente.

Parmi les ensembles infinis, les ensembles *dénombrables* sont les plus simples :

DÉFINITION II.3.4

Un ensemble E est dit **dénombrable** ssi il est équipotent à \mathbb{N} ; il est dit **au plus dénombrable** ssi il est fini ou dénombrable.

Les propriétés suivantes sont immédiates, compte tenu des théorèmes II.3.7 et II.3.8 : si E est au plus dénombrable, toute partie de E est au plus dénombrable ; si E est dénombrable, toute partie infinie de E l'est aussi ; si E est dénombrable, et si $f : E \rightarrow F$ est surjective, alors F est au plus dénombrable, donc dénombrable s'il est infini. On en déduit aisément que si E et F sont dénombrables, alors $E \cup F$ l'est aussi : car d'abord, $E \cup F$ est infini ; ensuite si $f : \mathbb{N} \rightarrow E$ et $g : \mathbb{N} \rightarrow F$ sont des bijections, l'application $h : \mathbb{N} \rightarrow E \cup F$, $2n \mapsto f(n)$, $2n+1 \mapsto g(n)$ pour tout $n \in \mathbb{N}$, est surjective, donc $E \cup F$ est au plus dénombrable, donc dénombrable puisque infini. Par une récurrence immédiate on en déduit que toute union finie d'ensembles dénombrables (resp. au plus dénombrables) l'est encore.

THÉORÈME II.3.11

|| Si E et F sont des ensembles dénombrables, $E \times F$ l'est aussi.

Démonstration :

Il suffit d'exhiber une bijection de $\mathbb{N} \times \mathbb{N}$ sur \mathbb{N}^* . Admettons provisoirement la factorisation des entiers en facteurs premiers, qui sera étudiée plus

⁽¹⁾ Bernstein (Serge) : Mathématicien russe (1880-1968), né à Odessa, auteur notamment de travaux sur l'approximation des fonctions continues.

loin (et, cela va de soi, indépendamment de la notion d'ensemble dénombrable). Il est alors immédiat que $\varphi : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}^*$, $(m, n) \mapsto 2^m(2n+1)$ est bijective. ■

COROLLAIRE 1

|| Si E_1, E_2, \dots, E_p ($p \in \mathbb{N}^*$) sont des ensembles dénombrables, l'ensemble produit $E_1 \times E_2 \times \dots \times E_p$ l'est aussi.

COROLLAIRE 2

|| Soit I un ensemble dénombrable, et pour tout $i \in I$, soit E_i un ensemble au plus dénombrable ; alors $E = \bigcup_{i \in I} E_i$ est au plus dénombrable.

Démonstration :

Il suffit d'envisager le cas où $I = \mathbb{N}$. Soit alors pour chaque $i \in \mathbb{N}$, $f_i : E_i \rightarrow \mathbb{N}$ une injection. Dans l'ensemble produit $\mathbb{N} \times \mathbb{N}$, posons : $F_i = \{i\} \times f_i(E_i)$, enfin soit $F = \bigcup_{i \in \mathbb{N}} F_i$. On a une surjection évidente de F sur E

en associant, à tout $(i, x) \in F$ ($i \in I, x \in F_i$), l'élément $t \in E_i$ tel que $f_i(t) = x$. Or F , partie de $\mathbb{N} \times \mathbb{N}$, est au plus dénombrable par le théorème II.3.11. Donc E est au plus dénombrable. ■

Pour « mesurer la taille » des ensembles infinis (indépendamment de toute idée d'ordre interne) a été élaborée une *théorie des cardinaux* que nous n'aborderons pas dans cet ouvrage. Signalons cependant que le cardinal d'un ensemble fini n'est autre que celui que nous avons défini plus haut (cf. définition II.3.2), c'est-à-dire un entier naturel.

Le cardinal de \mathbb{N} s'appelle *aleph zéro* et se note \aleph_0 ; c'est le « premier » cardinal non fini. Comme on va le voir ci-dessous, l'ensemble $\mathcal{P}(\mathbb{N})$ des parties de \mathbb{N} est *infini et non dénombrable*, ce qui prouve qu'il existe de tels ensembles ; son cardinal s'appelle *aleph un* et se note \aleph_1 , ou encore 2^{\aleph_0} : on dit aussi que c'est la *puissance du continu* (cf. exercice 8). Il est facile de démontrer que cette puissance du continu est « strictement supérieure » à la puissance du dénombrable, conséquence du résultat suivant, dû à Cantor ⁽¹⁾ : *Il n'existe aucune surjection d'un ensemble E sur l'ensemble $\mathcal{P}(E)$ de ses parties* ⁽²⁾.

Démonstration :

Soit $f : E \rightarrow \mathcal{P}(E)$ une application ; on considère

$$A = \{x \in E \mid x \notin f(x)\},$$

d'où $A \subset E$, i.e. $A \in \mathcal{P}(E)$. En raisonnant par l'absurde on voit que $A \notin \text{Im}(f)$. En effet si a était un élément de E tel que $f(a) = A$, chacun des deux seuls cas possibles $a \in A$ ou $a \notin A$ conduirait à une contradiction. ■

Hypothèse du continu

Disons de façon imagée qu'un ensemble F est « *plus peuplé* » qu'un ensemble E ssi

⁽¹⁾ Cantor (Georg) (1845-1918) : Mathématicien russe, un des principaux fondateurs de la théorie des ensembles, qui fut victime de l'incompréhension de ses contemporains.

⁽²⁾ Si E est non vide, l'application $E \rightarrow \mathcal{P}(E)$, $x \mapsto \{x\}$, est *injective*. Donc, lorsque E est infini, $\mathcal{P}(E)$ l'est aussi.

il existe une injection de E dans F , mais pas de F dans E . Compte tenu du théorème II.3.10, cela revient à dire qu'il existe une injection de E dans F , mais que E et F ne sont pas équipotents.

Primitivement, l'hypothèse du continu se proposait de répondre à la question suivante : existe-t-il un ensemble E plus peuplé que \mathbb{N} , mais moins peuplé que $\mathcal{P}(\mathbb{N})$? L'hypothèse du continu proprement dite était la conjecture selon laquelle il n'existerait pas de tel ensemble.

Plus généralement, étant donné un ensemble infini E , on peut se demander s'il existe un ensemble F , plus peuplé que E , mais moins peuplé que $\mathcal{P}(E)$ et proposer comme réponse l'hypothèse du continu généralisée (c'est-à-dire qu'il n'existe pas de tel ensemble F). Mais à la suite des travaux de Gödel, un mathématicien américain, Paul Cohen, a réussi à démontrer en 1962, que l'hypothèse du continu généralisée est une proposition **indécidable**.

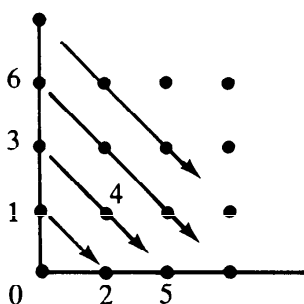
Le corps \mathbb{R} et chacun de ses intervalles non réduits à un point, et non vides, sont équipotents à $\mathcal{P}(\mathbb{N})$ (cf. exercice 8 ci-dessous) : ce sont donc des ensembles infinis non dénombrables. (Voir aussi, tome 2, Analyse).

Exercice 1 : Montrer que, pour qu'un ensemble E soit fini, il faut et il suffit que toute partie non vide de $\mathcal{P}(E)$ possède un élément maximal pour l'inclusion. *Indication :* Si E n'est pas fini, utiliser une injection de \mathbb{N} dans E .

Exercice 2 : Montrer que l'ensemble $\mathcal{F}(\mathbb{N}, \mathbb{N})$ n'est pas dénombrable. *Indication :* On pourra utiliser le fait que $\mathcal{F}(\mathbb{N}, \{0, 1\})$ est équipotent à $\mathcal{P}(\mathbb{N})$ ou alors utiliser la suite $g(n) = f_n(n) + 1$ (où (f_n) est une suite d'applications $\mathbb{N} \rightarrow \mathbb{N}$).

Exercice 3 : S'inspirer de la figure ci-contre pour exhiber une bijection explicite de $\mathbb{N} \times \mathbb{N}$ sur \mathbb{N} (*Réponse :* $\varphi(i, j) = \frac{(i+j)(i+j+1)}{2} + i$). Quelles sont les coordonnées du point qui porte le n° 100 ?

Démontrer que l'ensemble \mathbb{Q}_+ des rationnels positifs est dénombrable. *Indication :* L'application $f : \mathbb{N} \times \mathbb{N}^* \rightarrow \mathbb{Q}_+$, $(p, q) \mapsto \frac{p}{q}$ est surjective.



$\psi : J \rightarrow \mathbb{Q}_+^*$, $(\alpha_n) \mapsto \prod_{n \geq 1} p_n^{\alpha_n}$ est bijective. En déduire une bijection *explicite* de \mathbb{N}^* sur \mathbb{Q}_+^* , à l'aide du a).

c) Construire une bijection explicite de \mathbb{N}^* sur \mathbb{Q} .

Exercice 5 : Utiliser les théorèmes II.3.8 et II.3.10 pour prouver : si E est un ensemble infini et si $a \in E$, alors $E \setminus \{a\}$ est équipotent à E . Construire une bijection explicite de $[0, 1]$ sur $[0, 1[$.

Exercice 6 : On considère un ensemble E infini non dénombrable. Soit I une partie dénombrable de E . Utiliser les théorèmes II.3.8, II.3.10 et II.3.11 pour montrer que $E \setminus I$ et E sont équipotents.

Exercice 7 : Dans \mathbb{C} , on appelle **nombres algébriques** les nombres $z \in \mathbb{C}$ tels qu'il existe un polynôme non constant F à coefficients dans \mathbb{Z} pour lequel $F(z) = 0$. Montrer que l'ensemble des nombres algébriques dans \mathbb{C} est dénombrable. *Indication :* Si $F(z) = a_0 + a_1 z + \dots + a_n z^n$ on pourra utiliser la « hauteur » du polynôme : $h = |a_0| + |a_1| + \dots + |a_n| + n$.

Exercice 8 (cet exercice utilise la numération binaire des réels étudiée dans le cours d'Analyse).

a) Si E est un ensemble, l'ensemble $\mathcal{P}(E)$ des parties de E est en bijection naturelle avec l'ensemble $\mathcal{F}(E, \llbracket 0, 1 \rrbracket)$ des applications de E dans $\llbracket 0, 1 \rrbracket$.

b) En utilisant le développement dyadique d'un réel, démontrer qu'il existe une bijection de $\mathcal{F}(\mathbb{N}^*, \llbracket 0, 1 \rrbracket)$ sur l'intervalle $[0, 1]$ de \mathbb{R} . *Indication :* L'exercice 6 permet de se débarrasser des développements « impropres ».

c) Construire une bijection de $\mathcal{E} = \mathcal{F}(\mathbb{N}^*, \llbracket 0, 1 \rrbracket)$ sur $\mathcal{E} \times \mathcal{E}$. *Indication :* Séparer les termes de rang pair et impair.

d) En déduire que les ensembles suivants sont équipotents : $\mathcal{P}(\mathbb{N})$, $[0, 1]$, $[0, 1] \times [0, 1]$, et plus généralement $[0, 1]^p$ pour $p \in \mathbb{N}^*$.

e) En utilisant le théorème II.3.11, construire une bijection de \mathcal{E} sur $\mathcal{E}^{\mathbb{N}^*} = \mathcal{F}(\mathbb{N}^*, \mathcal{E})$. En déduire que $[0, 1]$ est équipotent à $[0, 1]^{\mathbb{N}} = \mathcal{F}(\mathbb{N}, [0, 1])$.

Exercice 9 : Démontrer que $\mathcal{F}(\mathbb{N}, \mathbb{N})$ est équipotent à $\mathcal{F}(\mathbb{N}, \llbracket 0, 1 \rrbracket)$.

Exercice 10 : Démontrer que si E est un ensemble infini, $E \times E$ est équipotent à E .

§ II.4 LOIS DE COMPOSITION. STRUCTURE DE GROUPE

Pour mieux tirer parti des ensembles de nombres que nous allons présenter, il est commode d'introduire dès maintenant la notion de groupe qui sera reprise de façon plus détaillée dans un chapitre ultérieur.

Lois de composition.

DÉFINITION II.4.1

Soit M un ensemble ; on appelle **loi de composition** sur M toute application de $M \times M$ dans M . La valeur de cette application en un couple (x, y) est appelée le **composé de x et y pris dans l'ordre énoncé**.

Une loi interne sur l'ensemble M peut être notée diversement, d'autant que plusieurs lois peuvent être définies sur le même ensemble. On peut utiliser un *symbole fonctionnel*, par exemple $f : (x, y) \mapsto f(x, y)$. Mais le plus souvent on utilise un *symbole de composition*, par exemple \top , que l'on intercale entre les deux composants : $(x, y) \mapsto x \top y$; on emploie aussi \perp , \vee , \wedge , $+$, \times , $*$, \oplus , \otimes , \cdot , \odot , \cup , \cap , \circ , etc.

Si le symbole choisi est $+$, la loi est dite notée *additivement* et l'on compose deux *termes* pour donner leur *somme*. Si le symbole choisi est \times , $*$, \otimes , ou \cdot , la loi est dite notée *multiplicativement* et l'on compose deux *facteurs* pour donner leur *produit* : dans ce dernier cas, on omettra souvent d'écrire le symbole intercalaire (notation sans symbole, le composé étant simplement écrit xy). Dans tous les cas, si l'on « calcule » des éléments de M par itération de la loi interne, il est indispensable d'utiliser des *parenthèses* pour définir l'ordre de priorité des diverses compositions effectuées. C'est ainsi par exemple que, pour $x, y, z, t \in M$ et une loi notée sans symbole, les éléments $(xy)(zt)$, $(x(yz)t)$, $x(y(zt))$, etc. sont en général distincts.

Dans ce qui suit, nous supposons l'ensemble M muni d'une loi interne notée \top . Les notions de base sont celles de la définition suivante :

DÉFINITION II.4.2

- ~ La loi \top est dite **associative** ssi $(x \top y) \top z = x \top (y \top z)$ pour tous $x, y, z \in M$.
- ~ La loi \top est dite **commutative** ssi $x \top y = y \top x$ pour tous $x, y \in M$.
- ~ Un élément $e \in M$ est dit **neutre** ssi $(\forall x \in M) e \top x = x \top e = x$.
- ~ Un élément $a \in M$ est dit **régulier à gauche** (resp. **à droite**) ssi l'application $M \rightarrow M, x \mapsto a \top x$ est **injective** (resp. l'application $x \mapsto x \top a$ est **injective**).
- ~ Un élément $a \in M$ est dit **régulier** ssi il est régulier à droite **et** à gauche.

On voit que dire que a est régulier à gauche (resp. à droite) signifie que $a \top x = a \top y \Rightarrow x = y$ (resp. $x \top a = y \top a \Rightarrow x = y$), ce qui s'énonce en disant que a est *simplifiable* à gauche (resp. à droite).

Si M admet un élément neutre pour la loi \top , il est unique : car si e et e' sont neutres, alors $e \top e' = e$ (neutralité de e') $= e'$ (neutralité de e).

Le gros intérêt d'une loi associative est de permettre l'écriture sans parenthèses du composé itéré d'éléments *pris dans un ordre déterminé*.

Par exemple si l'on a à composer les 4 éléments de M : x, y, z, t pris dans cet ordre, par une loi notée sans symbole, la notation $xyzt$ représente par convention le seul élément de M défini par $((xy)z)t$ tandis que si la loi est associative, $xyzt$ pourra représenter indifféremment $((xy)z)t$ ou $(xy)(zt)$ ou $(x(yz))t$ ou $x((yz)t)$ ou $x(y(zt))$ car ces 5 éléments de M sont égaux.

Nous verrons que si la loi est associative et commutative on peut en outre lever la contrainte sur l'ordre des éléments.

Supposons maintenant que la loi interne \top sur M admette un élément neutre e : on dit que $x \in M$ est **symétrisable à droite** pour \top (resp. à **gauche**) ssi : $\exists x' \in M$ tel que $x \top x' = e$ (resp. $\exists x'' \in M$ tel que $x'' \top x = e$) ; x est dit **symétrisable** ssi il est symétrisable à droite et à gauche. Lorsque x est symétrisable à droite (resp. à gauche), tout $y \in M$ tel que $x \top y = e$ s'appelle **un symétrique à droite** de x (resp. à gauche si $y \top x = e$). Il va de soi que pour une loi commutative tout symétrique à droite est aussi symétrique à gauche. Pour une loi notée multiplicativement les mots « symétrisable » et « symétrique » sont souvent remplacés par « inversible » et « inverse ».

PROPOSITION II.4.1

|| Si la loi interne \top sur l'ensemble M est **associative** et admet un **élément neutre** e , pour tout élément $x \in M$ symétrisable, le symétrique à droite est **unique**, le symétrique à gauche est unique, et ils sont égaux.

Démonstration :

Soit x' un symétrique à droite et x'' un symétrique à gauche de x ; alors

$$(x'' \top x) \top x' = x'' \top (x \top x') = e \top x' = x' = x'' \top e = x'' ;$$

bloquant x'' , on voit que x' est unique et $x' = x''$; de même x'' est unique et $x'' = x'$. ■

En raison de cette proposition, le symétrique à droite et à gauche d'un élément symétrisable x s'appelle **le symétrique** de x (on dit aussi **l'opposé** de x et on l'écrit $(-x)$ en notation additive, ou **l'inverse** de x en notation multiplicative, et on l'écrit x^{-1}).

Il importe de remarquer que *lorsque la loi \top est associative et admet un élément neutre e , tout élément symétrisable est régulier*. En effet si x' est le symétrique de x , la relation $x \top u = x \top v$ entraîne

$$x' \top (x \top u) = x' \top (x \top v)$$

d'où
$$(x' \top x) \top u = (x' \top x) \top v ,$$

d'où $e \top u = e \top v$ et enfin $u = v$ et de même $u \top x = v \top x \Rightarrow u = v$.

On réserve l'usage de la notation additive à certaines lois commutatives.

DÉFINITION II.4.3

⎵ On appelle **monoïde** tout couple (M, \top) où M est un ensemble et
⎵ \top une loi de composition sur M , **associative** et

élément neutre ; le monoïde est dit **abélien** ssi, en outre, la loi \top est **commutative**.

On observera qu'ayant un élément neutre un monoïde est donc **non vide**.

Exemple 1 : \mathbb{N} , muni de l'addition, est un monoïde abélien, dont l'élément neutre est 0.

\mathbb{N}^* , muni de la multiplication, est un monoïde abélien, dont l'élément neutre est 1.

Dans ces deux monoïdes, tout élément est régulier (cf. § II.1) : on dit que ce sont des monoïdes *réguliers*.

Exemple 2 : Soit X un ensemble ; alors les couples $(\mathcal{P}(X), \cup)$ et $(\mathcal{P}(X), \cap)$ sont des monoïdes abéliens ; l'élément neutre du premier est \emptyset , celui du deuxième est X . Ces monoïdes ne sont en général pas réguliers.

Exemple 3 : Soit $n \in \mathbb{N}^*$; l'ensemble $\mathbb{N}^n = \mathbb{N} \times \mathbb{N} \times \dots \times \mathbb{N}$, muni de la loi interne $+$ définie par

$$(\alpha_1, \alpha_2, \dots, \alpha_n) + (\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n),$$

est un monoïde abélien, d'élément neutre $\mathbf{0} = (0, 0, \dots, 0)$, et régulier.

DÉFINITION II.4.4

Soit (M, \top) et (M', \top') deux ensembles munis des lois de composition respectives \top et \top' . On dit qu'une application $f : M \rightarrow M'$ est un **morphisme** pour \top et \top' ssi :

$$(\forall (x, y) \in M \times M') f(x \top y) = f(x) \top' f(y).$$

On dit que f est un **isomorphisme** de (M, \top) sur (M', \top') ssi : f est **bijjective**, et f et $f^{\langle -1 \rangle}$ sont toutes deux des morphismes.

Un isomorphisme de (M, \top) sur lui-même est appelé un **automorphisme** de (M, \top) .

Il est clair qu'un morphisme bijectif est un isomorphisme, l'application Id_M est un automorphisme de (M, \top) . Le composé de deux morphismes (resp. isomorphismes) en est un.

Si f est un isomorphisme, $f^{\langle -1 \rangle}$ en est aussi un.

On dit que (M, \top) et (M', \top') sont **isomorphes** ssi il existe au moins un isomorphisme de (M, \top) sur (M', \top') . Lorsqu'il en est ainsi, si f est un tel isomorphisme, on voit que f établit une rigoureuse équivalence entre les propriétés de la loi \top et celles de \top' , par exemple \top et \top' sont associatives (ou non) en même temps, etc...

Un morphisme de (M, \top) dans lui-même est appelé un **endomorphisme** de (M, \top) .

Exemple 4 : Les monoïdes $(\mathbb{N}, +)$ et (\mathbb{N}^*, \times) de l'exemple 1 ne sont pas isomorphes (la démonstration est demandée dans l'exercice n° 3).

Groupes

DÉFINITION II.4.5

On appelle **groupe** tout monoïde dans lequel chaque élément est symétrisable. La loi interne est alors appelée la **loi de groupe** de ce groupe. Le groupe est dit **abélien** ⁽¹⁾ lorsque sa loi de groupe est commutative.

En d'autres termes, un groupe est un ensemble G muni d'une loi interne (que nous noterons sans symbole pour plus de commodité), vérifiant les axiomes suivants :

(G₁) G admet un élément neutre.

(G₂) Associativité : $(\forall (x, y, z) \in G \times G \times G) \quad (xy)z = x(yz)$.

(G₃) Tout élément de G est symétrisable.

(G₁) montre qu'un groupe n'est jamais vide.

Lorsque la loi du groupe G est notée additivement (rappelons qu'on réserve cette notation exclusivement à des groupes abéliens), l'élément neutre de G se note 0_G (ou simplement 0 si aucune confusion n'est possible). Pour les questions théoriques sur les groupes on préfère utiliser la notation multiplicative sans symbole, le neutre de G étant alors noté e_G (ou simplement e) ; en tout cas, sauf mention contraire, c'est ce que nous ferons dans la suite.

Exemple 5 : Les monoïdes présentés dans les exemples 1, 2, 3 ne sont pas des groupes car ils ne vérifient pas (G₃).

Exemple 6 : Soit Ω un ensemble ; munissons l'ensemble $\mathcal{P}(\Omega)$ de la loi interne : $(A, B) \mapsto A \triangle B = (A \cup B) \setminus (A \cap B)$. Cette loi interne est appelée *différence symétrique* (car $A \triangle B = (A \setminus B) \cup (B \setminus A)$) et nous verrons dans un prochain chapitre que $(\mathcal{P}(\Omega), \triangle)$ est un groupe abélien (ce qui autorise une notation additive). Ce groupe est fondamental en Calcul des Probabilités. Son élément neutre est \emptyset et chaque élément est son propre symétrique.

Dans un groupe, tout élément est régulier (voir la remarque qui suit la proposition II.4.1).

Exemple 7 : Soit G un groupe et I un ensemble non vide ; sur l'ensemble $\mathcal{F}(I, G)$ noté aussi G^I , on définit une structure de groupe en posant, pour

⁽¹⁾ Niels Henrik *Abel*, mathématicien norvégien (1802-1829), auteur de travaux fondamentaux sur les équations algébriques, les fonctions elliptiques et les intégrales. mort de tuberculose en pleine jeunesse sans avoir connu la gloire qu'il méritait.

$f, g \in G^I$: $fg : I \rightarrow G, i \mapsto f(i)g(i)$; cette structure est dite *naturelle*, le produit de deux applications étant défini par sa valeur au point i obtenue comme « produit » des valeurs de f et de g . L'élément neutre est l'application $\chi : I \rightarrow G, x \mapsto e_G$. Si G est abélien, le groupe G^I l'est aussi, et en notation additive son élément neutre est *l'application nulle*.

Pour $I = \llbracket 1, n \rrbracket$, avec $n \in \mathbb{N}^*$, on écrit G^n au lieu de $G^{\llbracket 1, n \rrbracket}$.

DÉFINITION II.4.6

Soit G et G' deux groupes dont les lois sont notées sans symbole. On appelle **homomorphisme de G dans G'** toute application $f : G \rightarrow G'$ telle que $(\forall (x, y) \in G \times G) f(xy) = f(x)f(y)$.
On appelle **isomorphisme de G sur G'** toute bijection $f : G \rightarrow G'$ telle que f et $f^{\langle -1 \rangle}$ soient des homomorphismes.
On appelle **automorphisme de G** tout isomorphisme de G sur G et **endomorphisme de G** tout homomorphisme de G dans G .
Les groupes G et G' sont dits **isomorphes** ssi il existe au moins un isomorphisme de G sur G' .

Par exemple $\text{Id}_G : G \rightarrow G$ est un automorphisme de G . Le composé de deux homomorphismes (resp. isomorphismes) en est un.

Si f est un isomorphisme, $f^{\langle -1 \rangle}$ en est un.

Si $f : G \rightarrow G'$ est un homomorphisme bijectif, alors c'est un isomorphisme.

Si $f : G \rightarrow G'$ est un homomorphisme, alors l'image de e_G est $e_{G'}$ car :

$$f(e_G) = f(e_G e_G) = f(e_G) f(e_G) = e_{G'} f(e_G)$$

puisque $e_{G'}$ est neutre, d'où en simplifiant par $f(e_G)$: $f(e_G) = e_{G'}$.

Exemple 8 : Reprenons l'exemple 7 et fixons une valeur $i \in I$. L'application $p_i : G^I \rightarrow G, f \mapsto f(i)$ est un homomorphisme de groupes.

Soit G un groupe. Une partie H de G est dite **stable pour la loi de groupe** de G ssi : $(\forall (x, y) \in H \times H) xy \in H$.

DÉFINITION II.4.7

On appelle **sous-groupe** d'un groupe G toute partie H de G **stable pour la loi de groupe** de G et telle que la loi \top définie sur H par $(\forall x \in H, \forall y \in H) x \top y = xy$ soit une **loi de groupe**.
La loi \top est appelée **loi induite** sur H par la loi de groupe de G .

On convient de noter la loi induite sur une partie stable de G par le même symbole que celui de la loi de G .

Une étude détaillée des sous-groupes d'un groupe sera faite dans un chapitre ultérieur. Bornons-nous ici à quelques remarques él

— Si H est un sous-groupe de G , alors $e_H = e_G$. En effet, $e_H e_H = e_H e_G$, d'où le résultat en simplifiant par e_H .

— Si H est un sous-groupe de G et si $x \in H$, alors le symétrique x'_H de x dans H n'est autre que le symétrique x' de x dans G . En effet, $xx'_H = e = xx'$, d'où $x'_H = x'$ en simplifiant par x .

Les propriétés ci-dessus sont caractéristiques :

THÉORÈME II.4.1

Soit G un groupe et H une partie de G ; pour que H soit un sous-groupe, il faut et il suffit qu'on ait :

- (I) $e_G \in H$
- (II) H est stable pour la loi de groupe de G
- (III) $(\forall x \in H)$ le symétrique x' de x appartient à H .

La démonstration est immédiate et sera laissée au lecteur.

Pratiquement on peut remplacer les conditions (II) et (III) par l'unique condition suivante (dans laquelle, pour tout $x \in G$, on désigne le symétrique de x par x^{-1})

$$(IV) \quad (\forall (x, y) \in H \times H) \quad xy^{-1} \in H.$$

Notons que toute intersection de sous-groupes d'un groupe G en est encore un.

Exemple 9 : Dans tout groupe G , les ensembles $\{e_G\}$ et G lui-même sont des sous-groupes.

Exemple 10 : Dans le groupe $(\mathcal{P}(\Omega), \Delta)$ de l'exemple 6, le sous-ensemble $\{\emptyset, \Omega\}$ est un sous-groupe.

Exemple 11 : L'ensemble $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ est un groupe pour la multiplication (voir § II.6) ; l'ensemble \mathbb{R}_+^* des réels > 0 est un sous-groupe de ce groupe.

De même, l'ensemble $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ est un groupe pour la multiplication (cf. *ibid.*) ; l'ensemble \mathbb{Q}_+^* des rationnels > 0 est un sous-groupe du groupe \mathbb{Q}^* ; c'est aussi un sous-groupe du groupe \mathbb{R}_+^* . On a le système d'inclusion suivant entre ces divers groupes :

$$\begin{array}{ccc} \mathbb{Q}_+^* & \hookrightarrow & \mathbb{Q}^* \\ \downarrow & & \downarrow \\ \mathbb{R}_+^* & \hookrightarrow & \mathbb{R}^* \end{array}.$$

THÉORÈME II.4.2

Soit deux groupes G et G' et $f : G \rightarrow G'$ un homomorphisme. Pour tout sous-groupe H de G , $f(H)$ est un sous-groupe de G' . Pour tout sous-groupe H' de G' , $f^{-1}(H')$ est un sous-groupe

A l'aide du théorème II.4.1, la vérification est immédiate. En particulier, si $f: G \rightarrow G'$ est un homomorphisme du groupe G dans le groupe G' , l'image de G notée $f(G)$ ou $\text{Im}(f)$ est un sous-groupe de G' , et $f^{-1}(e_{G'})$ et $f^{-1}(G') = G$ sont des sous-groupes de G .

DÉFINITION II.4.8

Soit deux groupes G et G' et $f: G \rightarrow G'$ un **homomorphisme**. Le sous-groupe de G , image réciproque du sous-groupe $\{e_{G'}\}$, est appelé le **noyau** de f et noté $\text{Ker } f$ ou $\text{Ker}(f)$.

THÉORÈME II.4.3

Soit deux groupes G et G' et $f: G \rightarrow G'$ un homomorphisme. Pour que l'homomorphisme f soit **injectif**, il faut et il suffit que son **noyau** soit réduit à $\{e_G\}$.

Démonstration :

La condition est évidemment nécessaire. Montrons qu'elle est suffisante. Si f n'était pas injectif, deux éléments distincts x_1 et x_2 de G auraient la même image, donc $f(x_1) = f(x_2)$, d'où

$$f(x_1^{-1}) f(x_1) = f(x_1^{-1}) f(x_2) \quad \text{d'où} \quad f(e_G) = f(x_1^{-1} x_2) \quad \text{et} \quad x_1^{-1} x_2 \neq e_G$$

appartiendrait à $\text{Ker } f$, contrairement à l'hypothèse. ■

Exercice 1 : Une partie S d'un monoïde M est appelée *sous-monoïde* de M lorsqu'elle possède l'élément neutre de M et qu'elle est stable pour la loi interne \top de M . Si c'est le cas, l'application $S \times S \rightarrow S$, $(x, y) \mapsto x \top y$ fait de S un monoïde, dont la loi est dite *induite* par \top sur S et se note encore \top .

a) Vérifier que toute intersection de sous-monoïdes de M en est un.

b) Montrer que si A est une partie non vide de M telle que $e_M \in A$, l'intersection $S(A)$ des sous-monoïdes de M contenant A est l'ensemble des composés itérés $a_1 \top a_2 \top \dots \top a_n$, où n parcourt \mathbb{N}^* et où les a_i sont quelconques dans A ; on dit que $S(A)$ est le sous-monoïde *engendré* par A .

Exercice 2 : On considère la suite strictement croissante $(p_k)_{k \in \mathbb{N}^*}$ des nombres premiers de \mathbb{N}^* et on fixe $n \in \mathbb{N}^*$. Démontrer que le sous-monoïde de (\mathbb{N}^*, \times) engendré par $\{1, p_1, p_2, \dots, p_n\}$ est isomorphe au monoïde $(\mathbb{N}^n, +)$ de l'exemple 3.

Exercice 3 : Démontrer que les monoïdes $(\mathbb{N}, +)$ et (\mathbb{N}^*, \times) ne sont pas isomorphes.

Indication : pour n donné dans \mathbb{N} , il y a exactement $n + 1$ couples $(u, v) \in \mathbb{N} \times \mathbb{N}$ tels que $u + v = n$.

Exercice 4 : Soit n et $p \in \mathbb{N}^*$, $n \neq p$. Démontrer que les monoïdes $(\mathbb{N}^n, +)$ et $(\mathbb{N}^p, +)$ ne sont pas isomorphes. Montrer également que $(\mathbb{N}^n, +)$ et (\mathbb{N}^*, \times) ne sont pas isomorphes.

Exercice 5 : Soit M un monoïde abélien, noté additivement, de neutre 0_M et soit I un ensemble non vide. On considère l'ensemble $\mathcal{F}(I, M)$ noté M^I muni de la loi $+$ suivante : pour $f, g \in M^I$, $f + g$ est l'application $I \rightarrow M$, $i \mapsto f(i) + g(i)$.

a) Vérifier que $(M^I, +)$ est un monoïde abélien.

b) Si $f \in M^I$, on appelle *support de f* l'ensemble des $i \in I$ tels que $f(i) \neq 0_M$. Montrer que le sous-ensemble (noté $M^{(I)}$) des $f \in M^I$ à *support fini* est un sous-monoïde de M^I (cf. exercice 1).

c) En utilisant la décomposition en facteurs premiers, montrer que le monoïde (\mathbb{N}^*, \times) est isomorphe à $\mathbb{N}^{(\mathbb{N}^*)}$.

Exercice 6 : Un monoïde régulier dans lequel tout élément est symétrisable à droite (resp. à gauche) est un groupe.

Exercice 7 : Un monoïde fini et régulier est un groupe.

Exercice 8 : Si $\alpha \in \mathbb{N}$, on note $\alpha\mathbb{N} = \{\alpha n\}_{n \in \mathbb{N}}$: c'est un sous-monoïde de $(\mathbb{N}, +)$.

a) Exhiber un sous-monoïde de $(\mathbb{N}, +)$ qui ne soit *pas* du type $\alpha\mathbb{N}$.

b) Si $\alpha \in \mathbb{N}^*$ et $\beta \in \mathbb{N}^*$, soit $d = \text{pgcd}(\alpha, \beta)$ (cf. le chapitre d'Arithmétique). Montrer qu'il existe $A \in \mathbb{N}^*$ tel que le sous-monoïde S de $(\mathbb{N}, +)$ engendré par $\{\alpha, \beta\}$ vérifie $S \cap \llbracket A, \rightarrow \llbracket = d\mathbb{N} \cap \llbracket A, \rightarrow \llbracket$.

c) Peut-on généraliser la propriété du b) ?

Exercice 9 : Si G est un groupe et si $k \in \mathbb{Z}$, l'application $f : G \rightarrow G, x \mapsto x^k$ est-elle un homomorphisme de groupes ?

Exercice 10 : a) On donne $n \in \mathbb{N}^*$; trouver tous les morphismes de monoïde de \mathbb{N}^n dans \mathbb{N} .

b) On donne un ensemble non vide I ; trouver les morphismes de monoïde de $\mathbb{N}^{(I)}$ (cf. exercice 5) dans \mathbb{N} .

c) Trouver les homomorphismes de groupe de \mathbb{Z}^n dans \mathbb{Z} et de $\mathbb{Z}^{(I)}$ dans \mathbb{Z} .

Exercice 11 : Soit G un groupe. On suppose qu'il existe $k \in \mathbb{N}^*$ tel que, pour tous $a \in G$ et $b \in G$: $(ab)^k = a^k b^k$ pour $i \in \{k-1, k, k+1\}$. Démontrer que G est abélien.

§ II.5 L'ANNEAU DES ENTIERS RELATIFS, LA STRUCTURE D'ANNEAU

Nous avons rencontré dans \mathbb{N} des opérations qui ne sont pas toujours possibles : la soustraction et la division. On généralise la notion de nombre pour pallier ce genre d'inconvénient : cela s'opère en plongeant \mathbb{N} dans des ensembles de nombres plus vastes où les opérations de base se comportent mieux à chaque étape de l'extension ; par exemple, dans \mathbb{Z} , la soustraction sera toujours possible ; dans \mathbb{Q} , la division par un élément non nul sera possible.

Entiers relatifs (ou entiers rationnels)

Sur l'ensemble $\mathbb{N} \times \mathbb{N}$, considérons la relation binaire \mathcal{R} définie par :

$$(a, b) \mathcal{R} (a', b') \text{ ssi } a + b' = a' + b.$$

On vérifie immédiatement que \mathcal{R} est une relation d'équivalence ; on note \mathbb{Z} l'ensemble quotient $\mathbb{N} \times \mathbb{N} / \mathcal{R}$, que l'on appelle ensemble des **entiers relatifs** (ou **entiers rationnels**). Nous allons rapidement étudier ces « nombres ». Pour cela, nous noterons $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{Z}$ la projection canonique. Rappelons que $\mathbb{N} \times \mathbb{N}$ est muni d'une loi additive naturelle, en posant : $(a, b) + (a', b') = (a + a', b + b')$ (cf. l'exemple 3 du § II.4), qui en fait un monoïde abélien. On vérifie :

$A \in \mathbb{Z}$ et $A' \in \mathbb{Z}$, pour $x \in A$ et $y \in A'$, l'élément $f(x + y)$ ne dépend que de A et A' et non du choix de x dans la classe d'équivalence A et de y dans A' (ce que l'on exprime en disant que la relation d'équivalence \mathcal{R} est *compatible avec l'addition* de $\mathbb{N} \times \mathbb{N}$).

On définit donc une loi interne sur \mathbb{Z} en posant, pour $A, A' \in \mathbb{Z} : A + A' =$ l'élément S de \mathbb{Z} tel que $(\forall x \in A, \forall y \in A'), f(x, y) = S$.

Nous laissons au lecteur le soin de vérifier dans les détails le théorème suivant, dont la démonstration est fastidieuse mais sans la moindre difficulté :

THÉORÈME II.5.1

|| Muni de la loi $+$ ci-dessus définie, l'ensemble \mathbb{Z} des entiers relatifs est un **groupe abélien**.

L'élément neutre de ce groupe (noté 0) n'est autre que $f(0, 0)$, ou plus précisément : $0_{\mathbb{Z}} = f(0_{\mathbb{N} \times \mathbb{N}})$. L'opposé de $f(a, b)$ est $f(b, a)$.

Une fois \mathbb{Z} construit, il est aisé de plonger \mathbb{N} dans \mathbb{Z} : pour $n \in \mathbb{N}$ posons $\varphi(n) = f((n, 0))$. On constate alors que φ est une injection de \mathbb{N} dans \mathbb{Z} ; si l'on note $-X$ l'opposé d'un élément $X \in \mathbb{Z}$, et si $I : \mathbb{Z} \rightarrow \mathbb{Z}$ est l'application $X \mapsto -X$, alors $\mathbb{Z} = \varphi(\mathbb{N}) \cup I(\varphi(\mathbb{N}))$, et φ est un morphisme de monoïdes de $(\mathbb{N}, +)$ dans le groupe $(\mathbb{Z}, +)$, c'est-à-dire :

$$(1) \quad (\forall (x, y) \in \mathbb{N} \times \mathbb{N}) \quad \varphi(x + y) = \varphi(x) + \varphi(y).$$

A l'aide de l'injection φ , on convient une fois pour toutes d'identifier \mathbb{N} à une partie de \mathbb{Z} : c'est ce que nous ferons désormais ; alors φ sera assimilée à l'injection canonique de \mathbb{N} dans \mathbb{Z} , et grâce à (1) l'addition de \mathbb{Z} prolonge celle de \mathbb{N} . Les éléments de \mathbb{N} s'appellent *éléments positifs* de \mathbb{Z} , et on pose aussi $\mathbb{N}^* = \mathbb{Z}_+^*$, $\mathbb{N} = \mathbb{Z}_+$; les éléments de $\mathbb{Z} \setminus \mathbb{Z}_+$ sont dits *strictement négatifs*, et on pose : $\mathbb{Z}_- = \mathbb{Z} \setminus \mathbb{Z}_+^*$, $\mathbb{Z}_-^* = \mathbb{Z} \setminus \mathbb{Z}_+$; l'application $n \mapsto -n$ définit une bijection de \mathbb{N}_+^* sur \mathbb{Z}_-^* .

Comme $(\mathbb{Z}, +)$ est un groupe abélien, la soustraction ne pose plus de problème dans \mathbb{Z} : si $x \in \mathbb{Z}$ et $y \in \mathbb{Z}$, il existe un, et un seul $d \in \mathbb{Z}$ tel que : $x + d = y$; c'est le nombre $d = y + (-x)$ que l'on note $y - x$ et qu'on appelle la différence de y et de x .

Si $x \in \mathbb{Z}$, un et un seul élément de $\{x, -x\}$ est dans \mathbb{N} : on l'appelle **valeur absolue de x** , et on le note $|x|$. Pour x, y dans \mathbb{Z} on a : $|x + y| \leq |x| + |y|$, avec égalité ssi x et y sont tous deux dans \mathbb{Z}_+ ou tous deux dans \mathbb{Z}_- .

Ensuite on munit \mathbb{Z} d'une relation d'ordre, appelée **ordre naturel** de \mathbb{Z} , notée \leq de la manière suivante : si x et $y \in \mathbb{Z}$, par définition $x \leq y$ ssi $y - x \in \mathbb{N}$. On constate que le sous-ensemble ordonné \mathbb{N} de (\mathbb{Z}, \leq) n'est autre que \mathbb{N} muni de son ordre naturel, ce qu'on exprime en disant que l'ordre naturel de \mathbb{Z} **prolonge** celui de \mathbb{N} . L'ordre naturel de \mathbb{Z} est total, compatible avec l'addition de \mathbb{Z} (c'est-à-dire $x \leq y \Leftrightarrow x + z \leq y + z$) et avec la multiplication par un entier > 0 (si $r \in \mathbb{N}^*$, $xr \leq yr \Leftrightarrow x \leq y$). Cet ordre est *archimédien*, c'est-à-dire : si $a \in \mathbb{Z}$ et si $b \in \mathbb{N}^*$, alors il existe $c \in \mathbb{Z}$ tel que $b \mid c \mid \geq |a|$.

Mais il est important de remarquer que (\mathbb{Z}, \leq) n'est pas bien ordonné (il n'y a pas de plus petit élément).

On définit les *intervalles* de \mathbb{Z} de la même manière que dans \mathbb{N} et on les note avec les symboles $[[$ et $]]$; on constate que, pour tout $a \in \mathbb{Z}$; les ensembles ordonnés $[[\leftarrow, a]]$ et $[[a, \rightarrow]]$ sont tous deux isomorphes à (\mathbb{N}, \leq) ; cette remarque

un peu d'attention, d'étendre aux entiers relatifs toutes les techniques de *réurrence* vues au § II.2.

Homomorphismes du groupe \mathbb{Z} dans un groupe

On considère un groupe G (noté multiplicativement sans symbole) ; si $x \in G$ et $m \in \mathbb{Z}$ on pose : $x^m = xxx \dots x$ (m fois) si $m \geq 1$, $x^m = e_G$ si $m = 0$ et $x^m = (x')^{-m}$ où x' est le symétrique de x si $m < 0$. L'associativité de la loi de groupe de G entraîne :

$$(1) \quad (\forall x \in G, \forall (m, n) \in \mathbb{Z}^2) \quad x^{m+n} = x^m x^n,$$

ce qui signifie :

Pour $x \in G$, l'application $\varphi_x : \mathbb{Z} \rightarrow G, m \mapsto x^m$ est un homomorphisme de groupes. En raison de (1) on voit que le symétrique x' de x est x^{-1} , ce qui justifie *a posteriori* cette notation du symétrique ; le symétrique de x^n est x^{-n} .

Réciproquement, soit $\varphi : \mathbb{Z} \rightarrow G$ un homomorphisme de groupes. Posons $x = \varphi(1)$. Par récurrence on a d'abord $\varphi(n) = x^n$ pour $n \in \mathbb{N}$; puis

$$\begin{aligned} \varphi(-n + n) &= \varphi(0) = e_G = \varphi(-n) \varphi(n) = \varphi(-n) x^n \\ &\text{d'où} \quad \varphi(-n) = x^{-n}. \end{aligned}$$

Autrement dit $\varphi = \varphi_x$. Donc :

Les seuls homomorphismes de \mathbb{Z} dans G sont ceux du type φ_x .

Dans le cas où G est noté additivement, on écrit mx au lieu de x^m , et (1) s'écrit

$$(2) \quad (\forall x \in G, \forall (m, n) \in \mathbb{Z}^2) \quad (m+n)x = mx + nx.$$

L'opposé de $x \in G$ est $(-1)x$; celui de mx est $(-m)x$.

Structure d'anneau de \mathbb{Z}

On munit le groupe abélien \mathbb{Z} d'une deuxième loi interne, appelée *multiplication*, et notée \times ou sans symbole, définie à partir de la multiplication dans \mathbb{N} par la règle des signes :

Si x et $y \in \mathbb{Z}$, par définition $xy = |x| |y|$ si x et y sont tous deux dans \mathbb{Z}_+ ou dans \mathbb{Z}_- , et $xy = -(|x| |y|)$ dans le cas contraire. On vérifie immédiatement le :

THÉORÈME II.5.2

|| La multiplication de \mathbb{Z} vérifie les propriétés suivantes :
 || (I) Elle **prolonge** celle de \mathbb{N} (pour x et $y \in \mathbb{N}$ leur produit dans \mathbb{Z} est celui dans \mathbb{N}).

(II) Elle est **associative, commutative** et admet **1** pour l'élément neutre.

(III) Elle est **distributive par rapport à l'addition**, c'est-à-dire :

$$(\forall (x, y, z) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}) \quad z(x + y) = zx + zy$$

et

$$(x + y)z = xz + yz.$$

On en déduit les règles de calcul usuelles dans \mathbb{Z} telles que :

pour $x \in \mathbb{Z}$: $0x = x0 = 0$; $-x = (-1)x$;

pour x et $y \in \mathbb{Z}$:

$$\begin{aligned} (-x)y &= x(-y) = -(xy) \\ (-x)(-y) &= xy = -[(-x)y] = -[-xy]. \end{aligned}$$

Si on définit x^m par récurrence pour $m \in \mathbb{N}^*$, en complétant par la convention $x^0 = 1$ pour tout $x \in \mathbb{Z}$, alors $x^{m+n} = x^m x^n$ pour m et $n \in \mathbb{N}$, et $(xy)^m = x^m y^m$, pour x et $y \in \mathbb{Z}$ et $m \in \mathbb{N}$.

On traduit les énoncés des théorèmes II.5.1 et II.5.2 en disant que : **l'addition et la multiplication de \mathbb{Z} définissent sur \mathbb{Z} une structure d'anneau commutatif, dont 0 est l'élément nul et dont 1 est l'élément unité.**

Il est immédiat que dans \mathbb{Z} la relation $xy = 0$ équivaut à ($x = 0$ ou $y = 0$). On traduit ce fait en disant que **\mathbb{Z} est un anneau intègre.**

DÉFINITION II.5.1

On appelle **anneau** un ensemble A muni de deux lois de composition interne : une addition (notée en général $+$) et une multiplication (notée en général \times ou sans symbole) satisfaisant aux axiomes suivants :

(\mathcal{A}_1) $(A, +)$ est un groupe abélien (son neutre, noté 0 ou 0_A est l'élément nul de A).

(\mathcal{A}_2) La multiplication est associative et distributive par rapport à l'addition, c'est-à-dire

(I) $(\forall (a, b, c) \in A^3) \quad (ab)c = a(bc)$

(II) $(\forall (a, b, c) \in A^3) \quad a(b+c) = ab+ac \text{ et } (a+b)c = ac+bc.$

(\mathcal{A}_3) La multiplication admet un élément neutre (noté en général 1_A , ou 1), appelé l'élément unité de A .

Un anneau est dit **commutatif** ssi sa multiplication est commutative.

Exemple 1 : Si A est réduit à un seul élément, noté 0, les lois $0+0=0$ et $0 \times 0=0$ définissent sur A une structure d'anneau, dans lequel $0_A = 1_A = 0$; cet anneau est appelé **anneau nul**.

Un anneau dans lequel $0_A = 1_A$ est certainement un anneau nul. Les anneaux pour lesquels $0_A \neq 1_A$ sont dits **non nuls**.

Exemple 2 : $(\mathbb{Z}, +, \times)$ est un anneau commutatif non nul.

Dans tout anneau A , on a les règles de calcul suivantes :

- $(\forall x \in A) \quad 0_A x = x 0_A = 0_A$

(car
$$x 0_A = x(0_A + 0_A) = x 0_A + x 0_A,$$

d'où en simplifiant par 0_A dans le groupe abélien $(A, +)$, $x 0_A = 0_A$, et de même pour $0_A x$).

- $(\forall x \in A) \quad -x = (-1_A)x = x(-1_A)$

(car
$$x + (-1_A)x = 1_A x + (-1_A)x = [1_A + (-1_A)]x = 0_A x = 0_A$$

et de même pour $x + x(-1_A)$).

- $(\forall (x, y) \in A^2) \quad (-x)y = x(-y) = -(xy);$

$$(-x)(-y) = -((-x)y) = -(x(-y)) = -(-(xy)) = xy.$$

On reconnaît la « règle des signes ».

- En définissant x^m par récurrence pour $m \in \mathbb{N}$ à partir de $x^0 = 1_A$ et $x^{m+1} = x^m x$ on obtient $(\forall (x, y) \in A^2) \quad (xy)^m = x^m y^m$ seulement si x et y sont permutables ($xy = yx$) tandis que $x^m x^n = x^{m+n}$ pour m et $n \in \mathbb{N}$ est toujours valable.

- Si $x \in A$ et si $m \in \mathbb{Z}$ on voit, par récurrence, que :

$$mx = (m 1_A)x = x(m 1_A).$$

Eléments réguliers, éléments inversibles d'un anneau

On considère ici un anneau non nul A .

Un élément $a \in A$ est dit **régulier à gauche** (resp. **à droite**) ssi l'application $x \mapsto ax$ (resp. $x \mapsto xa$) de A dans A est *injective*. Il revient au même de dire que $ax = ay \Rightarrow x = y$, ou que $ax = 0 \Rightarrow x = 0$ (pour x et $y \in A$).

Un élément régulier à gauche (resp. à droite) est nécessairement non nul. Mais un élément peut être régulier à gauche sans l'être à droite, et *vice versa*.

L'élément $a \in A$ est dit **régulier** ssi il est régulier à gauche **et** à droite.

Un élément $a \in A$ est dit **diviseur de zéro à gauche** (resp. à droite) ssi : $a \neq 0$ et a n'est pas régulier à gauche (resp. à droite), ce qui signifie $\exists x \in A \setminus \{0\}$ tel que $ax = 0$. Lorsque A est commutatif un élément quelconque est : soit nul, soit régulier, soit diviseur de zéro, sans considération de droite ou de gauche.

DÉFINITION II.5.2

Un anneau A est dit **intègre** ssi il est **non nul**, **commutatif**, et **sans diviseur de zéro**.

C'est un anneau commutatif, non nul, vérifiant :

$$(3) \quad (\forall (x, y) \in A^2) \quad xy = 0 \Leftrightarrow (x = 0 \text{ ou } y = 0).$$

Les anneaux non nuls vérifiant (3) mais non présumés commutatifs seront simplement appelés **anneaux sans diviseur de zéro**.

Exemple 3 : L'anneau $(\mathbb{Z}, +, \times)$ est intègre.

Soit A un anneau non nul. Un élément $a \in A$ est dit **inversible à droite** (resp. **à gauche**) ssi il est symétrisable à droite (resp. à gauche) pour la multiplication, c'est-à-dire :

$$\exists b \in A \mid ab = 1_A \quad (\text{resp. } \exists c \in A \mid ca = 1_A);$$

il est dit **inversible** ssi il l'est à droite et à gauche. Cela est conforme aux définitions du § II.4. D'après la proposition II.4.1, si $a \in A$ est inversible, il a un seul inverse à droite, un seul inverse à gauche, et ils sont égaux : l'élément ainsi défini s'appelle **l'inverse de a** et se note a^{-1} (parfois $\frac{1_A}{a}$).

Exemple 4 : Dans \mathbb{Z} , les seuls éléments inversibles sont -1 et 1 .

On démontre facilement le résultat suivant :

THÉORÈME II.5.3

Soit A un anneau non nul ; l'ensemble \mathcal{U}_A des éléments inversibles de A est **stable** pour la multiplication ; sur cet ensemble, la multiplication induit une loi de groupe.

Le groupe (\mathcal{U}_A, \times) s'appelle **groupe des éléments inversibles** de A ou parfois **groupe des unités** de l'anneau A . Bien sûr, on a : $\mathcal{U}_A \subset A \setminus \{0_A\}$.

Homomorphismes d'anneaux**DÉFINITION II.5.3**

Soit A et B deux anneaux ; on appelle **homomorphisme de A dans B** toute application $\rho : A \rightarrow B$ telle que :

$$(I) \quad \rho(1_A) = 1_B.$$

~ (II) ρ est un morphisme pour l'addition et un morphisme pour la multiplication, autrement dit :

$$(\forall (x, y) \in A^2) \quad \rho(x + y) = \rho(x) + \rho(y), \quad \rho(xy) = \rho(x) \rho(y).$$

~ On appelle **isomorphisme** d'anneaux de A sur B toute bijection $\rho : A \rightarrow B$ telle que ρ et ρ^{-1} soient des homomorphismes d'anneaux. On appelle **automorphisme** de l'anneau A tout isomorphisme de A sur A , et **endomorphisme** de A tout homomorphisme de A dans A .

Par exemple $\text{Id}_A : A \rightarrow A$ est un automorphisme. $\text{Id}_{\mathbb{Z}}$ est le seul automorphisme de \mathbb{Z} .

Le composé de deux homomorphismes (resp. isomorphismes) en est un.

Si $\rho : A \rightarrow B$ est un isomorphisme d'anneaux, $\rho^{-1} : B \rightarrow A$ en est un.

Si $\rho : A \rightarrow B$ est un homomorphisme *bijectif* d'anneaux, c'est un isomorphisme.

Exemple 5 : Soit A un anneau et I un ensemble ; sur $\mathcal{F}(I, A) = A^I$, on définit une structure d'anneau (dite *naturelle*) en posant :

$$\begin{aligned} \text{si } f \text{ et } g \in A^I, \quad f + g & \text{ est l'application } I \rightarrow A, \quad i \mapsto f(i) + g(i); \\ fg & \text{ est l'application } I \rightarrow A, \quad i \mapsto f(i) g(i). \end{aligned}$$

L'élément unité de A^I est l'application *constante* $u : I \rightarrow A, x \mapsto 1_A$.

Si A est commutatif, A^I l'est aussi.

Si $I = \llbracket 1, n \rrbracket$ avec $n \in \mathbb{N}^*$ on écrit A^n au lieu de A^I .

Cela dit, pour chaque $i \in I$, l'application $p_i : A^I \rightarrow A, f \mapsto f(i)$ est un homomorphisme d'anneaux.

DÉFINITION II.5.4

~ Soit A un anneau. On appelle **sous-anneau** de A toute partie $B \subset A$ qui est stable pour la multiplication, qui est un sous-groupe additif de A et telle que $1_A \in B$.

Si B est un sous-anneau de A , l'addition et la multiplication de A induisent sur B une structure d'anneau ; on se réfère toujours, implicitement à cette structure induite ; l'injection canonique $j : B \rightarrow A$ est alors un homomorphisme d'anneaux.

Il est clair que toute intersection de sous-anneaux d'un anneau A en est encore un et que si $\rho : R \rightarrow S$ est un homomorphisme de l'anneau R dans l'anneau S , alors $\text{Im}(\rho)$ est un sous-anneau de S .

Notons que si A est commutatif (resp. intègre), tout sous-anneau de A l'est encore, mais il se peut qu'un anneau A non commutatif (resp. non intègre) admette des sous-anneaux commutatifs (resp. intègres).

Exercice 1 : Soit $r \in \mathbb{Z}^*$; on munit le groupe abélien \mathbb{Z}^2 de la multiplication $*$ définie, pour $(x_1, x_2) \in \mathbb{Z}^2$ et $(y_1, y_2) \in \mathbb{Z}^2$, par :

$$(x_1, x_2) * (y_1, y_2) = (x_1 y_1 + r x_2 y_2, x_1 y_2 + x_2 y_1)$$

- Vérifier que $(\mathbb{Z}^2, +, *)$ est un anneau commutatif.
- Pour quels entiers r cet anneau est-il intègre ?

Exercice 2 : Si I est un ensemble non vide, quels sont, dans l'anneau \mathbb{Z}^I , les diviseurs de zéro ? les éléments inversibles ?

Exercice 3 : Soit A un anneau ; pour x et $y \in A$, on pose $[x, y] = xy - yx$;

- le groupe abélien A , muni du « produit » $(x, y) \mapsto [x, y]$, est-il un anneau ?
- vérifier : $(\forall (x, y, z) \in A^3) [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

Exercice 4 : Soit G un groupe abélien, noté additivement ; on désigne par φ l'injection canonique de \mathbb{N} dans \mathbb{Z} ; on donne une application $f : \mathbb{N} \rightarrow G$ qui soit un morphisme pour l'addition : $(\forall (x, y) \in \mathbb{N}^2) f(x + y) = f(x) + f(y)$. Montrer qu'il existe un, et un seul, homomorphisme de groupes $\bar{f} : \mathbb{Z} \rightarrow G$, tel que $\bar{f} \circ \varphi = f$; que si f est injectif, alors \bar{f} est injectif ; et que si f est injectif et qu'en plus, pour tout $x \in G$, on a : $x \in \text{Im}(f)$ ou $-x \in \text{Im}(f)$, alors \bar{f} est un isomorphisme.

Exercice 5 : Pour $n \in \mathbb{N}^*$, trouver les homomorphismes d'anneaux de \mathbb{Z}^n dans \mathbb{Z} .

Exercice 6 : Trouver tous les sous-anneaux de \mathbb{Z}^2 .

Exercice 7 : Dans un anneau A , on suppose : $(\forall x \in A) x^2 = x$. Montrer que A est commutatif. Si A contient au moins 3 éléments distincts montrer qu'il a nécessairement des diviseurs de zéro. Montrer qu'en fait A ne peut avoir exactement 3 éléments. *Nota :* A est appelé un *anneau de Boole*.

Exercice 8 (plus difficile). Dans un anneau A , on suppose : $(\forall x \in A) x^3 = x$. Montrer que A est commutatif.

Exercice 9 : a) Si B est un sous-anneau de l'anneau A , on note \mathcal{U}_A et \mathcal{U}_B les groupes d'éléments inversibles de A et de B ; comparer les ensembles \mathcal{U}_B et $\mathcal{U}_A \cap B$.

b) Plus généralement soit $\rho : B \rightarrow A$ un homomorphisme d'un anneau B dans un anneau A ; comparer $\rho(\mathcal{U}_B)$ et $\mathcal{U}_A \cap \text{Im}(\rho)$.

Exercice 10 : Soit \mathcal{S}_l le sous-ensemble de l'anneau $\mathbb{Z}^{\mathbb{N}}$ formé des *suites stationnaires* d'entiers relatifs.

- Vérifier que \mathcal{S}_l est un sous-anneau de $\mathbb{Z}^{\mathbb{N}}$.
- Trouver tous les homomorphismes d'anneaux : $\mathcal{S}_l \rightarrow \mathbb{Z}$.

Exercice 11 : On munit le groupe abélien \mathbb{Z}^2 de la multiplication \cdot définie, pour $(x_1, x_2) \in \mathbb{Z}^2$ et $(y_1, y_2) \in \mathbb{Z}^2$ par $(x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, 0)$. Obtient-on un anneau ? Quels sont les diviseurs de zéro ?

Exercice 12 : a) Donner des exemples d'une partie B d'un anneau A , non vide et stable pour l'addition et la multiplication de A , telle que ces lois induisent sur B une structure d'anneau, *mais que B ne soit pas un sous-anneau de A* (cf. aussi Chap. XI, exercice 9 du § IX.6). Trouver toutes les parties B de ce type si A est l'anneau $\mathbb{Z}^n = \mathcal{F}([1, n], \mathbb{Z})$ avec $n \geq 2$.

b) Soit B une partie d'un anneau A vérifiant les conditions de a), et qui ne soit pas un sous-anneau de A . On note 1_B l'élément unité de B . Que dire de 1_B ? Qu'en conclure si A est un anneau intègre ?

Exercice 13 : Soit A un anneau. On suppose, pour tout couple $(a, b) \in A \times A$:

$$(ab)^2 = a^2 b^2.$$

Démontrer que A est commutatif.

(Indication : penser à l'élément unité, qui est permutable avec tout élément.)

§ II.6 LES NOMBRES RATIONNELS, LA STRUCTURE DE CORPS

Dans l'anneau \mathbb{Z} la division exacte n'est pas toujours possible ; on est ainsi amené à étendre encore la notion de nombre, pour obtenir un ensemble plus vaste où cette opération ne posera plus un tel problème, et qui contiendra \mathbb{Z} avec son addition et sa multiplication prolongées au nouvel ensemble.

Nombres rationnels

Sur l'ensemble $\mathbb{Z} \times \mathbb{Z}^*$, on vérifie que la relation binaire \mathcal{R} définie par : $(\forall (a, b) \in \mathbb{Z} \times \mathbb{Z}^*, \forall (a', b') \in \mathbb{Z} \times \mathbb{Z}^*) (a, b) \mathcal{R} (a', b') \text{ ssi } ab' = ba' \text{ dans } \mathbb{Z}$, est une relation d'équivalence.

On note \mathbb{Q} l'ensemble $\mathbb{Z} \times \mathbb{Z}^* / \mathcal{R}$ et c'est cet ensemble qu'on appelle **ensemble des nombres rationnels**.

Pour l'étudier brièvement, nous désignerons par $(a, b) \mapsto \frac{a}{b}$ (ou a/b) la projection canonique $\mathbb{Z} \times \mathbb{Z}^* \rightarrow \mathbb{Q}$. On voit tout d'abord que \mathbb{Z} se plonge de façon naturelle dans \mathbb{Q} : si $x \in \mathbb{Z}$ il suffit de poser $\psi(x) = \frac{x}{1}$. Alors ψ est injective ; à l'aide de cette injection on identifie \mathbb{Z} à un sous-ensemble de \mathbb{Q} , et alors ψ s'identifie à l'injection canonique de \mathbb{Z} dans \mathbb{Q} .

On munit facilement \mathbb{Q} d'une structure d'anneau telle que \mathbb{Z} sera un sous-anneau de cet anneau, et que la structure d'anneau induite par \mathbb{Q} sur \mathbb{Z} sera sa structure d'anneau antérieurement définie. Le lecteur, ayant déjà calculé sur des fractions, ne sera pas étonné qu'on opère ainsi : si $A \in \mathbb{Q}$ et $B \in \mathbb{Q}$, on commence par vérifier que pour $(a, a') \in A$, $(b, b') \in B$, l'élément $\frac{ab' + a'b}{a'b'}$ ne dépend que de A et de B . Il en est de même de l'élément $\frac{ab}{a'b'}$, en observant que le « dénominateur » $a'b'$ n'est pas nul car l'anneau \mathbb{Z} est intègre et que $a' \neq 0$ et $b' \neq 0$. On définit donc deux lois internes sur \mathbb{Q} , notées $+$ et \times (la deuxième est souvent notée sans symbole), par : $(\forall A \in \mathbb{Q}, \forall B \in \mathbb{Q}) A + B$ est l'élément S de \mathbb{Q} tel que $\forall (a, a') \in A, \forall (b, b') \in B, S$ est la classe d'équivalence dont $(ab' + ba', a'b')$ est un élément, ce que l'on écrit $S = \frac{ab' + ba'}{a'b'}$; et de même AB est l'élément P de \mathbb{Q} tel que

$$(\forall (a, a') \in A, \forall (b, b') \in B), P = \frac{ab}{a'b'}.$$

Cela fait, il est facile de vérifier le théorème fondamental :

THÉORÈME II.6.1

*L'addition et la multiplication définies ci-dessus munissent \mathbb{Q} d'une structure d'anneau commutatif intègre ; \mathbb{Z} étant muni de sa structure d'anneau, l'injection canonique $\psi : \mathbb{Z} \rightarrow \mathbb{Q}$ est un **homomorphisme** d'anneaux. En d'autres termes, \mathbb{Z} est un sous-anneau de \mathbb{Q} , et la structure d'anneau induite par \mathbb{Q} sur \mathbb{Z} est celle définie sur \mathbb{Z} au § II.5.*

Si $A \in \mathbb{Q}$, les **représentants** de A , i.e. les couples $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$ tels que $a/b = A$, sont appelés **fractions** qui représentent A (a est le *numérateur* et b le *dénominateur* de cette fraction). L'identification de \mathbb{Z} avec $\psi(\mathbb{Z})$ consiste à confondre tout entier $x \in \mathbb{Z}$ avec la fraction $x/1$.

Il reste à vérifier que cette construction a permis de résoudre le problème posé au départ. C'est l'objet de la proposition suivante :

PROPOSITION II.6.1

|| Dans l'anneau \mathbb{Q} , tout élément **non nul** est inversible ; de plus l'application $\mathbb{Z} \times \mathbb{Z}^* \rightarrow \mathbb{Q}, (x, y) \mapsto xy^{-1}$ est surjective.

Démonstration :

Les éléments non nuls de \mathbb{Q} sont ceux représentés par les fractions $\frac{a}{b}$ avec $(a, b) \in \mathbb{Z}^* \times \mathbb{Z}^*$; il est alors clair que l'inverse de $\frac{a}{b}$ existe et que c'est $\frac{b}{a}$. En particulier, l'inverse de $a = \frac{a}{1} \in \mathbb{Z}^*$ est $\frac{1}{a}$; si $r \in \mathbb{Q}, r = \frac{a}{b}$ ($a \in \mathbb{Z}, b \in \mathbb{Z}^*$), on a donc $r = a \times \frac{1}{b} = a \times b^{-1}$, ce qui prouve la dernière assertion. ■

On sait dorénavant quelle signification accorder au quotient de deux entiers a et b ($b \neq 0$) : c'est le nombre rationnel $r = \frac{a}{b}$. Mais il y a plus : dans l'anneau commutatif intègre \mathbb{Q} , tout élément non nul est inversible, ce qui entraîne qu'on peut toujours diviser le rationnel A par le rationnel $B \neq 0$. On traduit toutes ces propriétés en disant que \mathbb{Q} (muni des lois ci-dessus) est un **corps commutatif**. L'ensemble $\mathbb{Q} \setminus \{0\}$, noté \mathbb{Q}^* , coïncide alors avec le groupe des éléments inversibles de l'anneau \mathbb{Q} .

Ordre naturel dans \mathbb{Q}

Si $r \in \mathbb{Q}^*$ il n'y a que deux cas possibles : ou bien toutes les fractions $\frac{a}{b}$ telles que $r = \frac{a}{b}$ sont telles que $ab \in \mathbb{N}^*$, ou bien toutes vérifient $-ab \in \mathbb{N}^*$.

On appelle rationnels **strictement positifs** ceux qui relèvent du premier cas, ensemble noté \mathbb{Q}_+^* . Si on leur adjoint 0, on obtient les rationnels **positifs** dont l'ensemble est noté \mathbb{Q}_+ . Cela dit, la relation binaire, notée \leq , sur \mathbb{Q} définie par : $x \leq y$ ssi $y - x \in \mathbb{Q}_+$ est une *relation d'ordre total* sur \mathbb{Q} , qui prolonge celle de \mathbb{Z} , et appelée *ordre naturel*, ou *ordre usuel*, de \mathbb{Q} . Cet ordre vérifie :

$$(\forall (x, y, z) \in \mathbb{Q}^3) \quad x \leq y \Leftrightarrow x + z \leq y + z$$

(compatibilité avec l'addition)

$$(\forall (x, y) \in \mathbb{Q}^2, \forall z \in \mathbb{Q}_+^*) \quad x \leq y \Leftrightarrow xz \leq yz$$

(compatibilité avec la multiplication par un rationnel *strictement positif*).

Cependant l'ordre usuel de \mathbb{Q} a une structure beaucoup plus complexe que celui de \mathbb{Z} : si on y définit, de la manière habituelle, les intervalles, pour des bornes A et $B \in \mathbb{Q}$ (avec $A < B$) on remarque que l'intervalle noté $[A, B]$ (et par voie de conséquence l'intervalle $]A, B[$) possède toujours une infinité d'éléments (contrairement à $[a, b]$ où a et $b \in \mathbb{Z}$) ; cela résulte du fait que $A < \frac{A+B}{2} < B$.

aussi que $]A, B[$ n'a pas de borne inférieure dans \mathbb{Q}^2 (contrairement à $]a, b[$). Cependant la propriété d'Archimède subsiste : en désignant par $|r|$ celui des deux nombres r ou $-r$ qui appartient à \mathbb{Q}_+ (appelé **valeur absolue** de r), on a : $(\forall a \in \mathbb{Q}_+, \forall r \in \mathbb{Q}) \exists n \in \mathbb{N}$ tel que $na \geq |r|$.

La structure de corps

DÉFINITION II.6.1

*On appelle **corps** tout anneau K non nul dans lequel tout élément non nul est inversible dans K . Un corps est dit **commutatif** ssi c'est un corps et un anneau commutatif.*

Comme nous venons de le voir, \mathbb{Q} est un corps ; mais \mathbb{Z} n'en est pas un.

Un corps non commutatif est parfois appelé une **algèbre à division**. (Mais il ne faudra pas oublier d'y distinguer quotient à droite et quotient à gauche).

Un anneau K est un corps ssi le groupe \mathcal{U}_K de ses éléments inversibles est $K \setminus \{0\}$; \mathcal{U}_K se note alors K^* . Si K est commutatif, le groupe (K^*, \times) est abélien.

Lorsque K est un corps, le groupe (K^*, \times) s'appelle **groupe multiplicatif de ce corps** (et se note K^*).

On trouvera dans l'exercice 30 du § XIII.5 l'exemple le plus simple d'algèbre à division. Il n'existe pas d'algèbre à division finie. Autrement dit, *tout corps fini est commutatif* (Th. de Wedderburn⁽¹⁾). On trouvera une démonstration de cette propriété dans l'exercice 15 du § IX.4.

Si un anneau est un corps, alors il n'admet pas de diviseur de zéro ; en particulier, un corps commutatif est nécessairement un anneau intègre.

DÉFINITION II.6.2

*On appelle **sous-corps** d'un corps K tout sous-anneau L de K tel que la structure d'anneau induite par K sur L soit une structure de corps.*

Si L est une partie de K , il y a équivalence entre les propriétés suivantes :

(\mathcal{C}_1) : L est un sous-corps de K ;

(\mathcal{C}_2) : L est un sous-anneau de K et pour tout $x \in L \setminus \{0\}$, on a : $x^{-1} \in L$ (x^{-1} désignant l'inverse de x dans K) ;

(\mathcal{C}_3) : L est un sous-anneau de K , et de groupe \mathcal{U}_L des éléments inversibles de L est $\mathcal{U}_L = L \cap K^*$.

Il est clair que tout sous-corps d'un corps commutatif est commutatif. Etudions maintenant les homomorphismes d'anneaux entre corps : un

⁽¹⁾ Wedderburn Joseph Henry *Maclagan*, Mathématicien écossais (1882-1948) auteur de travaux en théorie des groupes, fut le premier à démontrer que tout corps fini

isomorphisme d'anneaux entre deux corps sera appelé **isomorphisme de corps**. Un isomorphisme d'un corps K sur lui-même est appelé un **automorphisme du corps K** .

PROPOSITION II.6.2

Soit K un corps et L un anneau non nul. Soit $\rho : K \rightarrow L$ un homomorphisme d'anneaux ; alors ρ est injectif, le sous-anneau $K' = \rho(K)$ de L est un corps et $\rho|^{K'} : K \rightarrow K'$ est un isomorphisme de corps.

Démonstration :

L'important est de voir que ρ est injectif, tout le reste en découle. Or, soit x_1 et $x_2 \in K$ tels que $\rho(x_1) = \rho(x_2)$; alors

$$\rho(x_1 - x_2) = \rho(x_1) - \rho(x_2) = 0_L ;$$

posons $x = x_1 - x_2$; si x était $\neq 0_K$, on aurait

$$\rho(x^{-1}) \rho(x) = \rho(x^{-1}) 0_L = 0_L = \rho(x^{-1}x) = \rho(1_K) = 1_L ,$$

ce qui est contradictoire. Donc $x = 0_K$ et $x_1 = x_2$. ■

Soit K et L deux corps commutatifs. La proposition II.6.2 justifie le langage suivant :

un **isomorphisme de K dans L** est un homomorphisme d'anneaux de K dans L ; un **isomorphisme de K sur L** est un isomorphisme du corps K sur le corps L .

L'expression « homomorphisme de corps » est à proscrire.

DÉFINITION II.6.3

Soit K un corps commutatif ; on appelle **extension de K** tout couple (j, L) , où L est un corps commutatif et où j est un isomorphisme de K dans L .

Si (j, L) est une extension de K et si $K' = j(K)$, on peut d'après la proposition II.6.2 identifier, à l'aide de j , K au sous-corps K' de L . Inversement, un corps commutatif L peut être considéré comme extension de chacun de ses sous-corps M , à l'aide de l'injection canonique $M \rightarrow L$.

Il existe de très nombreuses extensions du corps \mathbb{Q} des rationnels, l'une des plus importantes étant le **corps des nombres réels** qui sera construit dans le tome d'Analyse et que l'on note \mathbb{R} . Dans ce tome d'Algèbre, nous supposons acquises les propriétés essentielles de \mathbb{R} ainsi qu'une certaine pratique du calcul sur les nombres réels. On a défini au chapitre I les notations \mathbb{R}_+ , \mathbb{R}_+^* , \mathbb{R}_- , \mathbb{R}_-^* et \mathbb{R}^* et, si besoin est, nous utiliserons les notations usuelles des intervalles telles que $[a, b]$, $[a, \rightarrow$, $[,]$

En revanche nous reviendrons plus en détail au chapitre VI sur une extension très remarquable de \mathbb{R} (et par conséquent de \mathbb{Q}), à savoir le **corps des nombres complexes** noté \mathbb{C} , en supposant également acquise une certaine pratique du maniement de ces nombres. Le corps \mathbb{C} est par excellence le corps de base dans une branche très vivante des mathématiques contemporaines : la **Géométrie algébrique**.

Revenons sur le procédé qui a permis de passer de \mathbb{Z} à \mathbb{Q} car il se généralise considérablement :

THÉOREME II.6.2

Soit A un anneau intègre ; il existe un couple (j, K) , où K est un corps commutatif et où $j : A \rightarrow K$ est un homomorphisme d'anneaux injectif, possédant les propriétés suivantes : l'application $A \times (A \setminus \{0\}) \rightarrow K$, $(x, y) \mapsto j(x) (j(y))^{-1}$ est surjective. De plus, si (j_1, K_1) et (j_2, K_2) sont deux tels couples, il existe un et un seul isomorphisme $\varphi : K_1 \rightarrow K_2$ tel que $j_2 = \varphi \circ j_1$.

Démonstration :

a) *Existence de (j, K) .* On procède exactement comme pour le théorème II.6.1. Sur l'ensemble $\mathcal{E} = A \times (A \setminus \{0\})$, on vérifie que la relation binaire \mathcal{R} , définie par : $(a, b) \mathcal{R} (a', b')$ ssi $ab' = a'b$ dans A , est une relation d'équivalence. On note K l'ensemble quotient $\mathcal{E} / \mathcal{R}$, et $(x, y) \mapsto \frac{x}{y}$ la projection canonique de \mathcal{E} sur K .

Les opérations $(a, b) \oplus (a', b') = \frac{ab' + a'b}{bb'}$ et $(a, b) \otimes (a', b') = \frac{aa'}{bb'}$ sont définies sur \mathcal{E} , compatibles avec \mathcal{R} . Cela permet de définir sur K deux lois de composition notées $+$ et \times (la deuxième aussi notée sans symbole) ; une addition : si X et $Y \in K$, $X + Y$ est l'élément S de K tel que $\forall x \in X, \forall y \in Y, Z = x \oplus y$; une multiplication : si X et $Y \in K$, XY est l'élément P de K tel que $\forall x \in X, \forall y \in Y, P = x \otimes y$. On voit facilement que ces deux lois munissent K d'une structure de corps commutatif ; l'élément nul est $0_K = \frac{0_A}{1_A}$, l'unité est $1_K = \frac{1_A}{1_A}$. Si

$X = \frac{a}{b} \in K$, avec $(a, b) \in \mathcal{E}$, alors $X \neq 0_K \Leftrightarrow a \neq 0_A$. Dans ce cas l'inverse de X est $\frac{b}{a}$. L'injection j est $A \rightarrow K, x \mapsto \frac{x}{1_A}$. Si $X = \frac{a}{b} \in K$ avec $(a, b) \in \mathcal{E}$, alors $X = j(a) (j(b))^{-1}$. Le couple (j, K) ainsi obtenu répond à la question.

b) Soit (j_1, K_1) et (j_2, K_2) répondant à la question. Si φ existe, on a, pour $(a, b) \in \mathcal{E}$, $\varphi \left(\frac{a}{b} \right) = j_2(a) (j_2(b))^{-1}$; réciproquement, si $X \in K_1$, par hypothèse il existe $(a, b) \in \mathcal{E}$ tel que $X = \frac{a}{b}$, l'élément $j_2(a) (j_2(b))^{-1}$ ne dépend que de X et non du choix du représentant (a, b) . Notons le $\varphi(X)$: on vérifie sans peine que l'application $\varphi : K_1 \rightarrow K_2$ est un isomorphisme du corps K_1 sur le corps K_2 et que $j_2 = \varphi \circ j_1$. ■

DÉFINITION II.6.4

Si A est un anneau intègre, le corps K construit dans la démonstration du théorème II.6.2 s'appelle **corps des fractions** de A , et $j : A \rightarrow K$ est l'**injection canonique** de A dans K .

En fait nous avons vu que de tous les corps contenant A , K est en quelque sorte « le plus petit possible », ce qui garantit son unicité, à un isomorphisme près. Bien sûr, à l'aide de j , on identifie A à un sous-anneau de K . On peut dire par exemple que \mathbb{Q} est le *corps des fractions* de l'anneau intègre \mathbb{Z} mais on remarquera que les nombres rationnels, éléments de \mathbb{Q} , ne sont pas à proprement parler des « fractions », ce terme désignant non pas un objet mathématique, mais un mode d'écriture, le même nombre rationnel r pouvant s'écrire d'une infinité de façons sous forme fractionnaire.

THÉORÈME II.6.3

Soit A un anneau intègre, K son corps des fractions, $j : A \rightarrow K$ l'injection canonique, et $f : A \rightarrow L$ un homomorphisme d'anneaux **injectif** de A dans un corps commutatif L .
Alors il existe un, et un seul, homomorphisme d'anneaux $\bar{f} : K \rightarrow L$ qui **prolonge** f , c'est-à-dire tel que $\bar{f} \circ j = f$, et \bar{f} est injectif, c'est un **isomorphisme de K dans L** .

Démonstration :

L'injectivité de \bar{f} vient de la proposition II.6.2.

Analyse : Si \bar{f} existe, on a, pour tout représentant a/b de $x \in K$: (où $a \in A, b \in A \setminus \{0\}$)

$$\bar{f}(b) \bar{f}(a/b) = \bar{f}(a) = f(a) = f(b) \bar{f}(a/b).$$

Puisque f est injectif, $b \neq 0_A$ entraîne $f(b) \neq 0_L$, d'où puisque L est un corps :

$$\bar{f}(a/b) = f(a) (f(b))^{-1}.$$

D'où l'unicité de \bar{f} .

Synthèse : Soit $x \in K$; si $x = a/b = a'/b'$ avec $(a, b) \in A \times (A \setminus \{0\})$,

$$(a', b') \in A \times (A \setminus \{0\}),$$

on en déduit : $ab' = ba'$, d'où : $f(a) f(b') = f(b) f(a')$, et, puisque $f(b) \neq 0_L$,

$$f(b') \neq 0_L : f(a) (f(b))^{-1} = f(a') (f(b'))^{-1}.$$

On peut donc définir une application $\bar{f} : K \rightarrow L, x \mapsto$ valeur commune des $f(a) (f(b))^{-1}$ pour les $(a, b) \in A \times (A \setminus \{0\})$ tels que $x = a/b$.

alors facilement que \bar{f} ainsi définie est un homomorphisme d'anneaux, et que $\bar{f} \circ j = f$, d'où le théorème. ■

Exercice 1 : Soit A un sous-anneau de \mathbb{Q} ; montrer que $\mathbb{Z} \subset A$ et que le corps des fractions de A est \mathbb{Q} . Plus généralement, si A est un anneau intègre, K son corps des fractions, et B un sous-anneau de K tel que $A \subset B \subset K$, alors K est le corps des fractions de B .

Exercice 2 : Le seul sous-corps de \mathbb{Q} est \mathbb{Q} .

Exercice 3 : Désignons par $E(x)$ la partie entière d'un réel x ; montrer que :

$$(\forall n \in \mathbb{N}^*, \forall t \in \mathbb{R}) \quad E(nt) = \sum_{k=0}^{n-1} E\left(t + \frac{k}{n}\right);$$

montrer aussi $E\left(\frac{E(nx)}{n}\right) = E(x)$.

Exercice 4 : Trouver tous les triplets $(x, y, z) \in (\mathbb{N}^*)^3$ tels que $\frac{1}{x} + \frac{1}{y} + \frac{1}{z} = 1$.

Exercice 5 : Trouver tous les triplets $(x, y, z) \in (\mathbb{N}^*)^3$ tels que $xyz = 1 + x + y + z$.

Exercice 6 : Montrer que $(\forall n \in \mathbb{N}) \quad \left(1 + \frac{1}{1}\right) \left(1 + \frac{1}{3}\right) \cdots \left(1 + \frac{1}{2n+1}\right) > \sqrt{2n+3}$.

Exercice 7 : Soit a et $b \in \mathbb{N}^*$, $a > b$. Démontrer que $\frac{a^2 + b^2}{a^2 - b^2} \notin \mathbb{N}$.

Exercice 8 : a) Utiliser l'identité

$$x^3 + y^3 + z^3 - 3xyz = (x + y + z)(x^2 + y^2 + z^2 - yz - zx - xy)$$

pour prouver que l'application $j: \mathbb{Z}^3 \rightarrow \mathbb{R}$, $(a, b, c) \mapsto a + b^3\sqrt{2} + c^3\sqrt{4}$ est injective ; en déduire que $\bar{j}: \mathbb{Q}^3 \rightarrow \mathbb{R}$, $(a, b, c) \mapsto a + b^3\sqrt{2} + c^3\sqrt{4}$ est injective.

b) Démontrer que $K = \text{Im}(\bar{j})$ est un sous-corps de \mathbb{R} , et prouver que le seul automorphisme de ce corps est Id_K .

Exercice 9 : Démontrer, pour $n \in \mathbb{N}$, $n \geq 2$, que $S_n = \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \notin \mathbb{N}$; généraliser pour $T_n = \frac{1}{a+b} + \frac{1}{a+2b} + \cdots + \frac{1}{a+nb}$, où a et $b \in \mathbb{N}^*$.

Exercice 10 : Résoudre dans \mathbb{Q}_+^* l'équation $x^y = y^x$, où $x < y$.

Réponse : $x = \left(1 + \frac{1}{n}\right)^n$, $y = \left(1 + \frac{1}{n}\right)^{n+1}$ avec $n \in \mathbb{N}^*$.

Exercice 11 : Soit $\alpha \in \mathbb{Q}$, $\alpha \in]0, 1[$; trouver $q \in \mathbb{N}^*$ tel que $\frac{1}{q+1} \leq \alpha < \frac{1}{q}$. En déduire que α s'écrit comme somme finie de termes du type $\frac{1}{q}$.

Exercice 12 : Résoudre dans $(\mathbb{N}^*)^3$ l'équation $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$.

Exercice 13 : Résoudre dans \mathbb{N}^2 l'équation $\frac{x-y}{x^2+xy+y^2} = \frac{2}{67}$, avec $(x, y) \neq (0, 0)$.

Exercice 14 : Si $n \in \mathbb{N}^*$, montrer que

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}.$$

Exercice 15 : Soit A un anneau intègre, et L un corps commutatif ; on su

homomorphisme d'anneaux injectif $\rho : A \rightarrow L$; montrer qu'il existe un, et un seul, isomorphisme du corps des fractions K de A dans L qui prolonge ρ .

Exercice 16 : L'intersection d'une famille non vide de sous-corps d'un corps K est un sous-corps de K .

Exercice 17 : a) Le seul automorphisme du corps \mathbb{Q} est $\text{Id}_{\mathbb{Q}}$, et c'est le seul isomorphisme de \mathbb{Q} dans \mathbb{Q} .

b) Montrer que le seul isomorphisme du corps \mathbb{R} dans lui-même est l'automorphisme $\text{Id}_{\mathbb{R}}$. (*Indication :* si $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ est un tel isomorphisme, on a : $\sigma(x) = x$ pour $x \in \mathbb{Q}$; puis $\sigma(x^2) = (\sigma(x))^2$ pour tout $x \in \mathbb{R}$, d'où l'on déduira que σ est une fonction *croissante*, et on conclura.)

Chapitre III

BASES DU CALCUL ALGÈBRIQUE ET COMBINATOIRE

§ III.1 ITÉRATION D'UNE LOI DE COMPOSITION

Composés itérés

Nous considérons ci-après un ensemble non vide M , muni d'une loi interne $\top : (x, y) \mapsto x \top y$.

DÉFINITION III.1.1

On appelle **composé d'éléments** a_0, a_1, \dots, a_n de M pris dans cet ordre (où $n \in \mathbb{N}$), le terme A_n de la suite $(A_k)_{0 \leq k \leq n}$ définie par récurrence ainsi :

$$A_0 = a_0 \quad \text{et} \quad \forall k \in \llbracket 0, n-1 \rrbracket \quad A_{k+1} = A_k \top a_{k+1}.$$

Le composé A_n est noté :

$$(I) \quad A_n = \bigtop_{i=0}^n a_i.$$

Dans la notation (I), l'indice i est « muet », ce qui signifie : le résultat A_n ne dépend pas de la lettre i , qu'on peut remplacer par n'importe quel autre symbole. Il convient de noter que, pour tout $k \in \llbracket 0, n \rrbracket$, le terme A_k intermédiaire figurant dans la définition III.1.1 est $\bigtop_{i=0}^k a_i$; en particulier,

si $n \geq 1$, $A_n = \left(\bigtop_{i=0}^{n-1} a_i \right) \top a_n$. Si la loi interne de M est notée additivement

(resp. multiplicativement), c'est-à-dire avec les symboles $+$ (resp. \times) au lieu de \top , on écrit

$$A_n = \sum_{i=0}^n a_i \quad \left(\text{resp. } A_n = \prod_{i=0}^n a_i \right) \quad \text{au lieu de (I).}$$

Justifions rigoureusement la définition III.1.1, indispensable si l'on veut rendre le traitement du calcul demandé automatique, par exemple pour une programmation sur ordinateur. Pour cela, si $n \geq 1$, on utilise le théorème II.1.5 de la manière suivante : prolongeons l'application $\llbracket 0, n \rrbracket \rightarrow M, k \mapsto a_k$ en une application $f: \mathbb{N} \rightarrow M$; considérons alors l'application

$$G: \mathbb{N} \times M \rightarrow \mathbb{N} \times M, (k, x) \mapsto (k+1, x \top f(k+1)) ;$$

d'après le théorème II.1.5, on a une, et une seule application $H: \mathbb{N} \rightarrow \mathbb{N} \times M$ telle que $H(0) = (1, f(0) \top f(1))$ et $(\forall p \in \mathbb{N}) H(p+1) = G(H(p))$. Cela dit, on vérifie par récurrence que, pour tout p , le terme $H(p)$ ne dépend que des éléments $(f(i))_{0 \leq i \leq p}$. En particulier, $H(n-1)$ ne dépend que de a_0, a_1, \dots, a_n ; on obtient le terme cherché A_n par la formule : $A_n = \text{pr}_2(H(n-1))$.

Le lecteur qui serait dérouté par ce qui précède pourra se convaincre que s'il veut programmer la construction de A_n , il sera bel et bien amené à procéder ainsi puisqu'il lui faudra deux « variables », une pour les *valeurs* a_i , une autre pour les *indices* i .

THÉORÈME III.1.1 (associativité)

Soit $n \in \mathbb{N}$ et $a_0, a_1, a_2, \dots, a_N$ des éléments de M . Soit $p \in \mathbb{N}$, et (n_0, n_1, \dots, n_p) une suite finie strictement croissante d'entiers naturels telle que $n_p = N$. Pour $i = 0$, posons $B_0 = \bigtop_{j=0}^{n_0} a_j$, et, pour $i \in \llbracket 1, N \rrbracket$ posons $B_i = \bigtop_{j=n_{i-1}+1}^{n_i} a_j$. Si la loi \top est associative, alors $\bigtop_{k=0}^N a_k = \bigtop_{i=0}^p B_i$.

Démonstration :

On raisonne par récurrence sur N ; la propriété étant évidente pour $N = 0$, il reste à prouver qu'elle est héréditaire. Il n'y a qu'à relire les notations pour régler le cas où $p = 0$. Supposons $p \geq 1$ et posons : $A = \bigtop_{j=0}^{N+1} a_j$, $A' = \bigtop_{j=0}^N a_j$, d'où $A = A' \top a_{n+1}$. Par l'hypothèse de récurrence $A' = \bigtop_{i=0}^{p-1} B_i$ si $n_{p-1} = N$, et

$$A' = \left(\bigtop_{i=0}^{p-1} B_i \right) \top V \text{ avec } V = \bigtop_{j=n_{p-1}+1}^N a_j \text{ si } n_{p-1} < N.$$

Si $n_{p-1} = N$, alors $B_p = a_{N+1}$, et

$$A = A' \top B_p = \left(\bigtop_{i=0}^{p-1} B_i \right) \top B_p = \bigtop_{i=0}^p B_i.$$

Si $n_{p-1} < N$, alors $B_p = V \top a_{n+1}$, et

$$A = A' \top a_{n+1} = \left(\left(\bigtop_{i=0}^{p-1} B_i \right) \top V \right) \top a_{n+1} = \left(\bigtop_{i=0}^{p-1} B_i \right) \top ($$

en utilisant l'associativité de \top ,

$$= \left(\bigtop_{i=0}^{p-1} B_i \right) \top B_p = \bigtop_{i=0}^p B_i,$$

ce qui achève la démonstration. ■

THÉORÈME III.1.2

Soit $N \in \mathbb{N}$ et $a_0, a_1, a_2, \dots, a_N$ des éléments de M . Si la loi \top est associative et commutative, alors, pour toute bijection

$$\sigma : \llbracket 0, N \rrbracket \rightarrow \llbracket 0, N \rrbracket, \text{ on a } \bigtop_{i=0}^N a_i = \bigtop_{i=0}^N a_{\sigma(i)}.$$

Démonstration :

Pour $N = 0$, σ est la bijection identique, et pour $N = 1$ le théorème résulte de la commutativité de \top . Supposant le théorème vrai à l'ordre $N \geq 1$ on va le montrer vrai à l'ordre $N + 1$. Posons $A = \bigtop_{i=0}^{N+1} a_i$, $B = \bigtop_{i=0}^{N+1} a_{\sigma(i)}$,

$A' = \bigtop_{i=0}^N a_i$. Soit r l'entier dans $\llbracket 0, N + 1 \rrbracket$ tel que $\sigma(r) = N + 1$.

• Si $r = N + 1$, alors $A' = \bigtop_{i=0}^N a_{\sigma(i)}$ car $\sigma \parallel \llbracket 0, N \rrbracket$ est une bijection de

$\llbracket 0, N \rrbracket$ sur lui-même, d'où $A = A' \top a_{N+1} = \left(\bigtop_{i=0}^N a_{\sigma(i)} \right) \top a_{\sigma(N+1)} = B$.

• Si $r = 0$, alors en utilisant le théorème III.1.1, $B = a_{N+1} \top \left(\bigtop_{j=1}^{N+1} a_{\sigma(j)} \right) = a_{N+1} \top U$ (où $U = \bigtop_{j=1}^{N+1} a_{\sigma(j)}$). En utilisant la commutativité de \top , $B = U \top a_{N+1}$. Mais U peut s'écrire $\bigtop_{k=0}^N b_k$ (avec $b_k = a_{\sigma(k+1)}$). Soit φ la bijection de $\llbracket 0, N \rrbracket$ dans $\llbracket 0, N \rrbracket$,

$k \mapsto \sigma^{-1}(k) - 1$. Alors l'hypothèse de récurrence entraîne que $V = \bigtop_{k=0}^N b_{\varphi(k)} = A'$, d'où $A = U \top a_{N+1} = A' \top a_{N+1} = B$.

• Si $1 \leq r \leq N$, posons $V = \bigtop_{i=r+1}^{N+1} a_{\sigma(i)}$; par commutativité de \top , on a :

$a_{n+1} \top V = V \top a_{n+1}$ d'où, d'après le théorème III.1.1

$$B = \left(\biguplus_{i=0}^{r-1} a_{\sigma(i)} \right) \top (a_{N+1} \top V) =$$

$$= \left(\left(\biguplus_{i=0}^{r-1} a_{\sigma(i)} \right) \top V \right) \top a_{N+1} = W \top a_{N+1},$$

avec $W = \left(\biguplus_{i=0}^{r-1} a_{\sigma(i)} \right) \top V$. En réutilisant le théorème III.1.1, $W = \biguplus_{k=0}^N b_k$

avec $b_k = a_{\sigma(k)}$ si $k < r$, $b_k = a_{\sigma(k+1)}$ si $k \geq r$. Soit φ la bijection $\llbracket 0, N \rrbracket \rightarrow \llbracket 0, N \rrbracket$, $k \mapsto \sigma^{-1}(k)$ si $k < r$, $k \mapsto \sigma^{-1}(k) - 1$ si $k \geq r$; par l'hypothèse de récurrence, $W = \biguplus_{k=0}^N b_{\varphi(k)} = A'$, et enfin

$$B = A' \top a_{N+1} = A. \quad \blacksquare$$

Lois associatives et commutatives : principe de Fubini ⁽¹⁾

Ci-dessous, nous considérons un **monoïde abélien** (M, \top) , dont l'élément neutre sera noté e .

Soit I un ensemble fini non vide, de cardinal n , et soit $(a_i)_{i \in I}$ une famille d'éléments de M . D'après le théorème III.1.2, si $\varphi : \llbracket 1, n \rrbracket \rightarrow I$ est une bijection, l'élément $\biguplus_{k=1}^n a_{\varphi(k)}$ de M ne dépend pas de φ , d'où :

DÉFINITION III.1.2

On appelle **composé des a_i par \top** , et on note $\biguplus_{i \in I} a_i$, l'élément $A \in M$ tel que $A = \biguplus_{k=1}^n a_{\varphi(k)}$ pour toute bijection $\varphi : \llbracket 1, n \rrbracket \rightarrow I$.

On complète cette définition en convenant que si $I = \emptyset$, alors $\biguplus_{i \in I} a_i = e$. Grâce aux théorèmes III.1.1 et III.1.2, on a immédiatement :

THÉORÈME III.1.3

Soit I et L deux ensembles finis non vides et soit $(J_\lambda)_{\lambda \in L}$ un partage de I . Soit $(a_i)_{i \in I}$ une famille dans M . Pour chaque $\lambda \in L$, posons $A_\lambda = \biguplus_{i \in J_\lambda} a_i$. Alors on a l'égalité :

$$\biguplus_{i \in I} a_i = \biguplus_{\lambda \in L} A_\lambda = \biguplus_{\lambda \in L} \left(\biguplus_{i \in J_\lambda} a_i \right).$$

⁽¹⁾ Fubini (Ghirin), mathématicien italien (1879-1943).

De façon imagée, nous dirons que dans un monoïde abélien on peut opérer « par paquets ». Voici une application très importante de ce théorème. Considérons deux ensembles finis non vides I et J , et une partie non vide L de $I \times J$. Donnons-nous une famille $(a_{i,j})_{(i,j) \in L}$ d'éléments de M . Pour chaque $i \in I$ (resp. $j \in J$), notons $L_{(i)} = \{j \in J \mid (i,j) \in L\}$ (resp. $L^{(j)} = \{i \in I \mid (i,j) \in L\}$).

Alors $(L_{(i)})_{i \in I}$ et $(L^{(j)})_{j \in J}$ sont des partages de L . En leur appliquant le théorème III.1.3, on obtient le **principe de Fubini** qui constitue un puissant outil de calcul, et se traduit par la formule

$$(1) \quad \prod_{(i,j) \in L} a_{i,j} = \prod_{i \in I} \left(\prod_{j \in L_{(i)}} a_{i,j} \right) = \prod_{j \in J} \left(\prod_{i \in L^{(j)}} a_{i,j} \right).$$

Il va de soi qu'au deuxième et au troisième membres de (1), on peut ne retenir que celles, parmi les expressions $\prod_{j \in L_{(i)}} a_{i,j}$ (resp. $\prod_{i \in L^{(j)}} a_{i,j}$), qui sont telles que $L_{(i)} \neq \emptyset$ (resp. $L^{(j)} \neq \emptyset$), ce qui abrège les calculs. On voit maintenant d'où sort la convention qui suit la définition III.1.2.

Exemple 1. Evaluer la somme d'entiers $S_n = \sum_{\substack{(p,q) \in \mathbb{N} \times \mathbb{N} \\ p+q \leq n}} (p+q)$. Utilisons la formule (1) :

$$\begin{aligned} S_n &= \sum_{p=0}^n \left(\sum_{q=0}^{n-p} (p+q) \right) = \sum_{p=0}^n \left[p(n-p+1) + \sum_{q=0}^{n-p} q \right] = \\ &= \frac{1}{2} \sum_{p=0}^n (n+p)(n-p+1) \end{aligned}$$

d'où

$$\begin{aligned} S_n &= \frac{n(n+1)}{2} \times (n+1) - \sum_{p=1}^n \frac{p(p-1)}{2} = n(n+1) \left[\frac{n+1}{2} - \frac{n-1}{6} \right] = \\ &= \frac{n(n+1)(n+2)}{3}. \end{aligned}$$

En sommant par paquets, on obtiendrait

$$S_n = 1 \cdot 0 + 2 \cdot 1 + \dots + (k+1)k + \dots + (n+1)n = \frac{1}{3} n(n+1)(n+2)$$

en tenant compte du fait que si $f(k) = \frac{1}{3} k(k+1)(k+2)$, on a

$$f(k) - f(k-1) = \frac{1}{3} k(k+1)(k+2) - \frac{1}{3} (k-1)k(k+1) = k(k+1),$$

$$\text{d'où } \sum_{k=0}^n f(k) = f(n).$$

Exercice 1 : Calculer $S_n = \sum p^2 q^2$, la somme étant étendue aux entiers naturels p et q tels que $p + q = n$.

§ III.2 CALCUL DANS UN ANNEAU

Nous considérons ci-après un anneau A , et nous allons, grâce à ce qui précède, développer dans A les produits de sommes finies. Par une récurrence immédiate, on voit d'abord que si a, b_1, b_2, \dots, b_n sont dans A alors $a \left(\sum_{i=1}^n b_i \right) = \sum_{i=1}^n (ab_i)$, et $\left(\sum_{i=1}^n b_i \right) a = \sum_{i=1}^n b_i a$; d'où, pour toute famille finie $(b_i)_{i \in I}$ d'éléments de A et tout $a \in A$,

$$(1) \quad a \left(\sum_{i \in I} b_i \right) = \sum_{i \in I} (ab_i); \quad \left(\sum_{i \in I} b_i \right) a = \sum_{i \in I} (b_i a).$$

En appliquant le *principe de Fubini*, on en déduit que si $(a_i)_{i \in I}$ et $(b_j)_{j \in J}$ sont deux familles finies dans A , on a :

$$(2) \quad \boxed{\left(\sum_{i \in I} a_i \right) \left(\sum_{j \in J} b_j \right) = \sum_{(i,j) \in I \times J} (a_i b_j)}.$$

A partir de (2) on obtient facilement le développement d'un produit fini de sommes finies dans A (toute succession d'additions et de soustractions en nombre fini peut, rappelons-le, être mise sous forme de somme) : soit I_1, I_2, \dots, I_N des ensembles finis ; pour chaque $i \in \llbracket 1, N \rrbracket$, soit $(a_{i,j})_{j \in I_i}$ une famille d'éléments de A , et posons $A_i = \sum_{j \in I_i} a_{i,j}$. Alors :

$$(3) \quad \prod_{i=1}^N A_i = \sum_{(i_1, i_2, \dots, i_N) \in I_1 \times I_2 \times \dots \times I_N} a_{1, i_1} a_{2, i_2} \dots a_{N, i_N},$$

ce qui s'écrit aussi

$$(4) \quad \prod_{i=1}^N \left(\sum_{j \in I_i} a_{i,j} \right) = \sum_{(i_1, \dots, i_N) \in I_1 \times \dots \times I_N} \left(\prod_{k=1}^N a_{k, i_k} \right).$$

Attention ! comme nous n'avons pas supposé l'anneau commutatif, il importe que, dans les relations (3) et (4), pour chaque terme $\prod_{k=1}^N a_{k,i_k}$ de la

grande somme, l'ordre des facteurs a_{k,i_k} soit strictement respecté. C'est aussi la raison pour laquelle il serait incorrect de noter ce terme

$\prod_{k \in \llbracket 1, N \rrbracket} a_{k,i_k}$ puisque nous avons réservé cette écriture $\left(\prod_{\lambda \in L} b_\lambda \right)$ au cas où

le produit est associatif **et commutatif**.

Cependant, même dans un anneau non commutatif, il peut arriver que deux éléments a et b de A soient **permutables** (c'est-à-dire tels que $ab = ba$).

Alors on obtient les égalités

$$(5) \quad \boxed{\begin{aligned} a^n - b^n &= (a - b) \sum_{k=0}^{n-1} a^{n-1-k} b^k \\ a^n - b^n &= (a - b)(a^{n-1} + a^{n-2}b + \dots + b^{n-1}) \end{aligned}}$$

qui se vérifient facilement par récurrence, pour $n \in \mathbb{N}^*$. En effet, pour $k \in \mathbb{N}$, $a^{k+1} - b^{k+1} = a(a^k - b^k) + (a - b)b^k$, et on utilise le fait que $ab = ba \Rightarrow (ab)^k = a^k b^k$ ($(ab)^2 = abab = aabb = a^2 b^2$, etc. Voir théorème III.1.2).

On retiendra tout particulièrement la formule suivante, valable pour tout $a \in A$:

$$(6) \quad \boxed{a^n - 1_A = (a - 1_A)(a^{n-1} + a^{n-2} + \dots + a + 1_A)}$$

qui résulte de (5) car a et 1_A sont évidemment permutables.

On remarque que si $1_A - a$ est inversible dans A , la relation (6) fournit une **formule sommatoire** pour la **somme géométrique finie** $1_A + a + a^2 + \dots + a^{n-1}$, à savoir

$$(7) \quad \sum_{k=0}^{n-1} a^k = (1_A - a)^{-1} \times (1_A - a^n).$$

Lorsque A est un corps, (7) s'écrit plutôt

$$(8) \quad (\forall a \in A), \quad a \neq 1_A : \quad \sum_{k=0}^{n-1} a^k = \frac{1 - a^n}{1 - a}.$$

Nous utiliserons fréquemment (8), en particulier si $a \in \mathbb{Q}$,

§ III.3 COMPOSÉ DE FAMILLES A SUPPORT FINI. NUMÉRATION

Reprenons un **monoïde abélien** (M, \top) d'élément neutre e . Soit I un ensemble quelconque, et $a = (a_i)_{i \in I}$ une famille d'éléments de M .

Nous appellerons **support de la famille a** , et nous noterons $\text{supp}(a)$, l'ensemble $\{i \in I \mid a_i \neq e\}$. Cette notion est étroitement liée à la loi de composition \top et en particulier à son élément neutre e .

Cela dit, on peut donner un sens au composé des a_i lorsque le support de a est fini. De manière précise :

(1) **si le support de a est fini**, on pose : $\prod_{i \in I} a_i = \prod_{i \in \text{supp}(a)} a_i$, cette dernière

écriture ayant un sens d'après la définition III.1.2. Si $\text{supp}(a) = \emptyset$, rappelons que l'élément ainsi défini est e . Avec un peu d'attention, il est facile d'étendre le théorème III.1.3 et le principe de Fubini aux familles à support fini, en prenant la précaution de n'écrire que des symboles qui représentent, en fait, des *composés finis*. On obtient ainsi :

PROPOSITION III.3.1

|| Soit $(a_i)_{i \in I}$ une famille d'éléments de M et $(J_\lambda)_{\lambda \in L}$ un partage de I .
|| Si la famille $(a_i)_{i \in I}$ est à **support fini**, alors

$$\prod_{i \in I} a_i = \prod_{\lambda \in L} \left(\prod_{i \in J_\lambda} a_i \right).$$

PROPOSITION III.3.2

|| Soit I et J deux ensembles et L une partie de $I \times J$. Si la famille
|| $(a_{i,j})_{(i,j) \in L}$ est à **support fini**, on a

$$(I) \quad \prod_{(i,j) \in L} a_{i,j} = \prod_{i \in I} \left(\prod_{j \in L(i)} a_{i,j} \right) = \prod_{j \in J} \left(\prod_{i \in L(j)} a_{i,j} \right).$$

Dans cet énoncé, les symboles $\prod_{j \in L(i)} a_{i,j}$ et $\prod_{i \in L(j)} a_{i,j}$ ont un sens parce que

toute sous-famille d'une famille à support fini l'est encore.

Soit $(a_i)_{i \in I}$ et $(b_i)_{i \in I}$ deux familles à support fini dans M ; alors la famille $(a_i \top b_i)_{i \in I}$ est à support fini (contenu dans $\text{supp}(a) \cup \text{supp}(b)$) et on a :

$$(2) \quad \prod_{i \in I} (a_i \top b_i) = \left(\prod_{i \in I} a_i \right) \top \left(\prod_{i \in I} b_i \right)$$

comme cas particulier de la proposition III.3.2.

L'intérêt majeur de l'écriture (2) est que figure partout le *même ensemble d'indices I* , ce qui évite une écriture très lourde et d'un maniement peu commode.

Exemple 1 : Numération des entiers. On se propose de représenter n'importe quel entier à l'aide d'un nombre fini $q \geq 2$ de symboles notés $0, 1, \dots$, et éventuellement du signe $-$ (pour les entiers négatifs). Ce problème, dont une solution pratique est connue de tous (par l'adoption universelle du système décimal), se trouve déjà exposé dans l'*Arénaire* d'Archimède, qui montre bien où réside la difficulté : la représentation doit rester valable même pour des entiers « très grands ». Montrons ici le fondement théorique sur lequel repose cette solution.

THÉOREME III.3.1

|| A chaque suite à support fini $s = (s_n)_{n \in \mathbb{N}}$ d'entiers appartenant à $\llbracket 0, q-1 \rrbracket$, associons l'entier naturel $\Phi_q(s) = \sum_{n \in \mathbb{N}} s_n q^n$. L'application Φ_q ainsi définie est une bijection de l'ensemble \mathcal{S}_q des suites à support fini à valeurs dans $\llbracket 0, q-1 \rrbracket$ sur l'ensemble \mathbb{N} .

Démonstration :

a) *Injectivité.* Si $\Phi_q(s) = \Phi_q(s')$, supposons par exemple $s_0 \geq s'_0$; alors : $s_0 - s'_0 = \sum_{n \geq 1} (s'_n - s_n) q^n$. Le premier membre appartient à $\llbracket 0, q-1 \rrbracket$, et le second est multiple entier de q , donc $s_0 = s'_0$; alors en divisant par q , il reste : $\sum_{n \geq 1} s_n q^{n-1} = \sum_{n \geq 1} s'_n q^{n-1}$, ce qui permet de recommencer le raisonnement, d'où, par récurrence, $s_n = s'_n$ pour tout n .

b) *Surjectivité.* Soit $N \in \mathbb{N}^*$, et supposons que tous les entiers de $\llbracket 1, N \rrbracket$, appartiennent à l'image de Φ_q . Qu'en est-il de $N+1$? La division euclidienne de $N+1$ par q donne : $N+1 = aq + r$ avec $r \in \llbracket 0, q-1 \rrbracket$. L'hypothèse de récurrence (du fait que $q \geq 2$) montre alors que $a \in \text{Im}(\Phi_q)$, d'où évidemment $N+1 = aq + r \in \text{Im}(\Phi_q)$. Comme $0 \in \text{Im}(\Phi_q)$ et $1 \in \text{Im}(\Phi_q)$ de toute évidence, la surjectivité est établie. ■

On remarque le rôle joué par le fait que \mathbb{N} est *archimédien*, car c'est cette propriété qui a permis de définir la division euclidienne dans \mathbb{N} . On sait que dans le système de numération à base DIX, les dix « chiffres » arabes utilisés pour représenter les dix premiers entiers naturels sont $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$. Le système de numération *binnaire*, utilisé dans le langage machine des ordinateurs, est plus économique en symboles : il lui suffit des « chiffres » 0 et 1 .

Exercice 1 (numération anglaise) : Soit $(r_k)_{k \geq 1}$ une suite d'entiers > 1 ; on considère l'ensemble \mathcal{A}_r des suites à support fini $s = (s_i)_{i \in \mathbb{N}}$ d'entiers tels que $s_i < r_{i+1}$ pour tout $i \geq 0$. Montrer que l'application

$$\mathcal{A}_r \mapsto \mathbb{N}, \quad s \mapsto \sum_{i \in \mathbb{N}} s_i \left(\prod_{k=0}^i r_k \right) = s_0 + s_1 r_1 + s_2 r_1 r_2 + \dots$$

est bijective.

Exercice 2 : Dans quel système de numération le nombre qui s'écrit 1986 en base dix s'écrit-il 1201002 ? (*Indication :* une base $q < 0$ n'est pas interdite !)

§ III.4 DÉNOMBREMENT

Pour tout entier $n \in \mathbb{N}^*$, l'entier $\prod_{k \in \llbracket 1, n \rrbracket} k = 1 \times 2 \times \cdots \times n$ se note $n!$ et s'appelle *factorielle de n* . On convient que $0! = 1$. Cette fonction de \mathbb{N} dans \mathbb{N} peut aussi être définie de façon *réursive* par $0! = 1$ et $n! = n \times (n-1)!$ pour $n \geq 1$.

Le principe (ou lemme) des bergers

Soit E un ensemble fini de cardinal $n \geq 1$. Considérons la fonction constante $\chi_E : E \rightarrow \mathbb{N}$, $x \mapsto 1$. On vérifie par récurrence sur n , la relation fondamentale

$$(1) \quad \text{card}(E) = \sum_{x \in E} \chi_E(x),$$

que l'on écrit parfois sous la forme moins explicite : $\text{card}(E) = \sum_{x \in E} 1$. En appliquant le théorème III.1.3, on déduit de (1) :

THÉORÈME III.4.1 (principe des bergers)

$$\left\| \begin{array}{l} \text{Soit } E \text{ un ensemble fini et } (F_i)_{i \in I} \text{ un } \textbf{partage} \text{ de } E \text{ (où } I \text{ est un} \\ \text{ensemble fini). Alors} \end{array} \right. \quad \text{card}(E) = \sum_{i \in I} \text{card}(F_i).$$

En particulier, si \mathcal{F} est une partie de $\mathcal{P}(E)$ formée d'ensembles *disjoints* tels que $\bigcup_{A \in \mathcal{F}} A = E$, on a :

$$(2) \quad \boxed{\text{card}(E) = \sum_{A \in \mathcal{F}} \text{card}(A)}.$$

Les relations (1) et (2), le principe des bergers ainsi que les théorèmes II.3.6 et II.3.7 donnant le cardinal d'une réunion et d'un produit cartésien, aussi évidents qu'ils paraissent, sont les outils de base de la branche des mathématiques actuellement très vivante appelée *Analyse combinatoire*. En voici les premiers rudiments :

THÉORÈME III.4.2

$$\left\| \begin{array}{l} \text{Soit } E \text{ et } F \text{ deux ensembles finis, de cardinaux respectifs } n \geq 1 \text{ et} \\ p \geq 1. \text{ Alors } \mathcal{F}(E, F) \text{ est un ensemble fini, et} \\ \text{card}(\mathcal{F}(E, F)) = p^n. \end{array} \right.$$

Démonstration :

Par récurrence sur n .

Pour $n = 1$, $\mathcal{F}(E, F)$ et F sont équipotents, d'où le résultat. Supposons le théorème vrai à l'ordre $n \in \mathbb{N}^*$, et que $\text{card}(E) = n + 1$; fixons $a \in E$; pour chaque $b \in F$, notons \mathcal{E}_b l'ensemble des applications $f: E \rightarrow F$ telles que $f(a) = b$. On contrôle que $(\mathcal{E}_b)_{b \in F}$ est une partition de $\mathcal{F}(E, F)$, d'où, d'après (2) $\text{card}(\mathcal{F}(E, F)) = \sum_{b \in F} \text{card}(\mathcal{E}_b)$.

Mais chacun des \mathcal{E}_b est équipotent à $\mathcal{F}(E \setminus \{a\}, F)$, de cardinal p^n par l'hypothèse de récurrence. D'où

$$\text{card}(\mathcal{F}(E, F)) = \sum_{b \in F} p^n = p^n \text{card}(F) = p^{n+1}. \quad \blacksquare$$

COROLLAIRE

|| Si E est un ensemble fini de cardinal $n \geq 1$, alors $\mathcal{P}(E)$ est fini, de cardinal 2^n .

En effet, à chaque partie A de E , associons la fonction $\chi_A: E \rightarrow \{0, 1\}$ $x \mapsto 0$ si $x \notin A$, $x \mapsto 1$ si $x \in A$ (fonction indicatrice de la partie A). On vérifie que $A \mapsto \chi_A$ est une bijection de $\mathcal{P}(E)$ sur $\mathcal{F}(E, \{0, 1\})$, d'où le résultat (en appliquant le théorème III.4.2, qui s'écrit aussi de façon plus suggestive : $\text{card}(F^E) = (\text{card}(F))^{\text{card}(E)}$).

THÉORÈME III.4.3

|| Soit E et F deux ensembles finis de cardinaux respectifs $n \geq 1$, $p \geq 1$, avec $p \leq n$. Le nombre des applications **injectives** de F dans E est $\frac{n!}{(n-p)!} = n(n-1) \dots (n-p+1) = \prod_{i=1}^p (n-p+i)$.

Démonstration :

Notons $\mathcal{J}(X, Y)$ l'ensemble des injections d'un ensemble X dans un ensemble Y et raisonnons par récurrence sur p .

Pour $p = 1$, le théorème est évident.

Supposons-le vrai à l'ordre $p-1 \geq 1$. Fixons un élément $b \in F$ et notons A_n^p le cardinal de $\mathcal{J}(F, E)$. À chaque $\alpha \in \mathcal{J}(F, E)$, associons $f(\alpha) = \alpha(b)$, élément de E . Alors $(f^{-1}(y))_{y \in E}$ est un partage de $\mathcal{J}(F, E)$, d'où $A_n^p = \sum_{y \in E} \text{card}(f^{-1}(y))$.

Fixons $y \in E$; l'application

$$f^{-1}(y) \rightarrow \mathcal{J}(F \setminus \{b\}, E \setminus \{y\}),$$

$\alpha \mapsto \alpha \Big|_{F \setminus \{b\}}^{E \setminus \{y\}}$ est bijective, d'où d'après l'hypothèse de récurrence, $\text{card}(f^{-1}(y)) = \frac{(n-1)!}{(n-p)!}$. Il s'ensuit $A_n^p = \sum_{y \in E} \frac{(n-1)!}{(n-p)!} = \frac{n!}{(n-p)!}$.

L'entier $\prod_{i=1}^p (n - p + i) = \frac{n!}{(n-p)!}$ que nous avons noté A_n^p s'appelle aussi le *nombre d'arrangements* (sans répétition) de n objets pris p à p , n^p étant le nombre d'arrangements de n objets p à p avec répétition possible (penser au tirage dans une « population » E de taille n d'un « échantillon » ordonné de taille p , avec ou sans « remise »).

En tenant compte des résultats qui suivent le théorème III.3.4, on en déduit :

COROLLAIRE

|| Si E et F sont deux ensembles finis **de même cardinal** $n \geq 1$, le nombre des **bijections** de F sur E (ou de E sur F) est : $n!$

En particulier, lorsque $\text{card}(E) = n$, le cardinal de l'ensemble (noté \mathfrak{S}_E) des **permutations** de E (c'est-à-dire des bijections de E sur E) est $n!$

Combinaisons

THÉORÈME III.4.4

|| Soit E un ensemble fini à n éléments et p un naturel, $1 \leq p \leq n$. Le nombre des parties à p éléments de E est

$$\frac{n!}{p! (n-p)!} = \frac{1}{p!} A_n^p = \frac{n(n-1) \dots (n-p+1)}{p!}.$$

On remarquera que pour $p = 0$ les deux premières expressions donnent bien 1 comme résultat et que $1 = \text{card}(\{\emptyset\})$.

Démonstration :

Notons $\mathcal{F}_p(E)$ l'ensemble des parties à p éléments de E . On considère l'application $f: \mathcal{J}([1, p], E) \rightarrow \mathcal{F}_p(E)$, $\alpha \mapsto \text{Im}(\alpha)$; le principe des bergers donne

$$A_n^p = \text{card}(\mathcal{J}([1, p], E)) = \sum_{A \in \mathcal{F}_p(E)} \text{card}(f^{-1}(A)).$$

Fixons $A \in \mathcal{F}_p(E)$; l'application $\alpha \mapsto \alpha|_A$ est bijective de $f^{-1}(A)$ sur l'ensemble des bijections de $[1, p]$ dans A . Donc $\text{card}(f^{-1}(A)) = \text{card}(\mathcal{B}) = p!$ (corollaire du théorème III.4.1); d'où $A_n^p = \sum_{A \in \mathcal{F}_p(E)} p! = p! \text{card}(\mathcal{F}_p(E))$. ■

Le nombre de p -parties de E , qui vaut donc $\frac{n!}{p! (n-p)!}$ se note

$\binom{n}{p}$ ou C_n^p . On l'appelle parfois nombre de combinaisons (de n objets pris p à p). Puisque $\binom{n}{p} = \text{card}(\mathcal{F}_p(\llbracket 1, n \rrbracket))$, c'est un entier, ce qui n'apparaît de façon évidente dans aucune des écritures

$$\frac{n!}{p!(n-p)!}, \frac{A_n^p}{p!}, \frac{n(n-1)\dots(n-p+1)}{p!}.$$

Voici quelques remarques élémentaires sur les nombres $\binom{n}{p}$:

- On a : $\binom{n}{p} = \binom{n}{n-p}$, ce qui se voit sur l'expression symétrique $\frac{n!}{p!(n-p)!}$, et peut se retrouver directement à l'aide de la bijection $A \mapsto \llbracket 1, n \rrbracket \setminus A$, de $\mathcal{F}_p(\llbracket 1, n \rrbracket)$ sur $\mathcal{F}_{n-p}(\llbracket 1, n \rrbracket)$.

- Pour $0 \leq p \leq n$: $\binom{n}{p+1} = \frac{n-p}{p+1} \binom{n}{p}$, ce qui montre que pour n fixé la suite $\binom{n}{p}_{p \in \llbracket 0, n \rrbracket}$ est croissante tant que $2p+1 \leq n$ et décroissante ensuite, le maximum étant atteint pour $p = \frac{n}{2}$ si n est pair, pour $p = \frac{n-1}{2}$ et pour $p = \frac{n+1}{2}$ si n est impair.

- Pour $1 \leq p \leq n$:

$$\boxed{\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}}$$

Cette relation qui se vérifie aisément par le calcul s'appelle **relation de Pascal** ⁽¹⁾. Elle est remarquable à plus d'un titre : elle prouve que la famille $\binom{n}{p}$ ($n \in \mathbb{N}$, $p \in \mathbb{N}$) est la seule famille $u_{n,p}$ de rationnels définie à partir de $u_{n,0} = 1$ pour tout $n \in \mathbb{N}$ et de $u_{0,p} = 0$ pour tout $p \in \mathbb{N}^*$ par la relation de récurrence

$$(\forall p \geq 1, \forall n \geq 1) \quad u_{n,p} = u_{n-1,p} + u_{n-1,p-1}$$

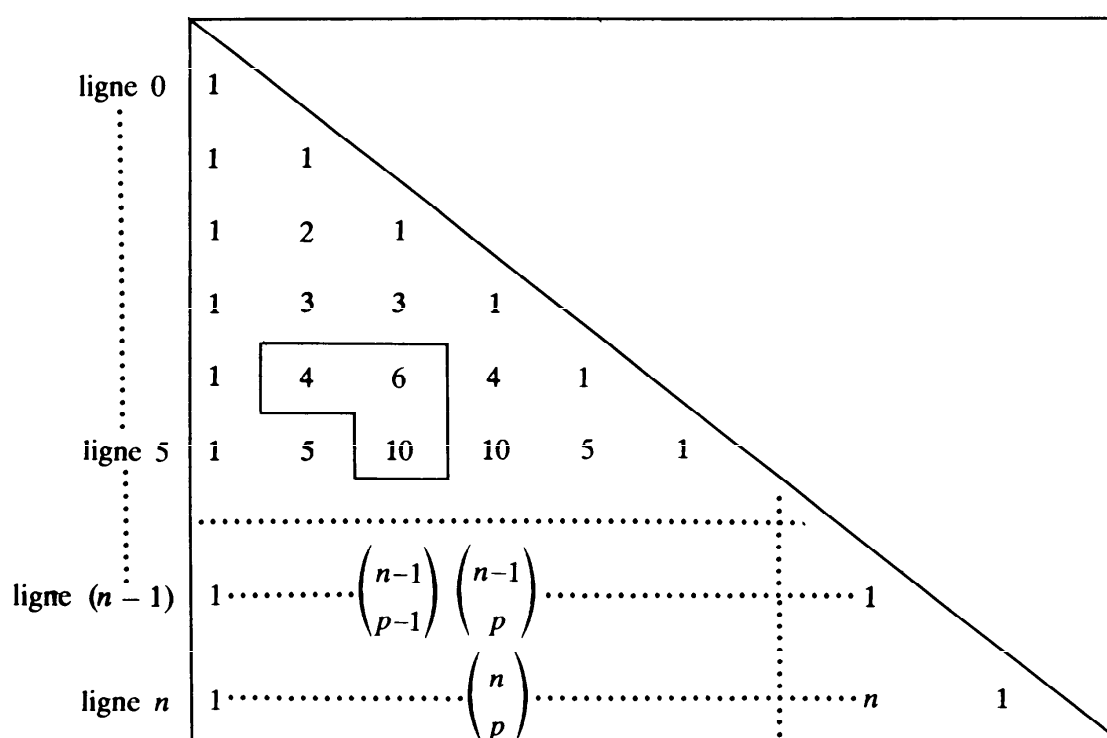
(on remarquera que $u_{n,p} = 0$ chaque fois que $p > n$, et que tous les $u_{n,p}$ sont évidemment dans \mathbb{N}).

Pour $1 \leq p \leq n$ on peut prouver directement la relation de Pascal comme suit : si E est un ensemble de cardinal $n \geq 1$ et si $n \geq p \geq 1$, on sépare les p -parties de E en deux sous-ensembles : d'une part celles qui contiennent un

⁽¹⁾ *Pascal* (Blaise) (1623-1662), mathématicien et physicien français ; a écrit un *Traité du triangle arithmétique* en 1654. Il est surtout célèbre comme écrivain et philosophe (*les Pensées*).

élément déterminé $a \in E$, d'autre part celles qui ne contiennent pas a . Les premières sont au nombre de $\binom{n-1}{p-1}$, les autres, de $\binom{n-1}{p}$, d'où le résultat.

La relation de Pascal permet de retrouver les $\binom{n}{p}$, ligne après ligne, à l'aide du **triangle de Pascal**, dont la construction apparaît sur le schéma ci-contre, dans lequel on n'a pas fait figurer tous les termes nuls qui rempliraient la partie droite du schéma.



- Enfin, $\sum_{p=0}^n \binom{n}{p} = 2^n$, comme il résulte immédiatement du fait que les $(\mathcal{F}_p(\llbracket 1, n \rrbracket))_{0 \leq p \leq n}$ forment un partage de $\llbracket 1, n \rrbracket$, et du corollaire du théorème III.4.2. On remarquera que la famille des $(u_{n,p})_{p \in \mathbb{N}}$ est (pour n fixé), à support fini, et que $\sum_{p \in \mathbb{N}} u_{n,p} = 2^n$.

Combinaisons avec répétition (ou p -choix)

THÉORÈME III.4.5

Soit E un ensemble fini à n éléments, $n \geq 1$, et $p \in \mathbb{N}$.
 (I) Le nombre des applications $u : E \rightarrow \llbracket 0, p \rrbracket$ telles que

$$\sum_{x \in E} u(x) \leq p \quad \text{est} \quad \binom{n+p}{p} = \binom{n+p}{n}$$

$$\left\| \begin{array}{l} \text{(II) Le nombre des applications } u : E \rightarrow \llbracket 0, p \rrbracket \text{ telles que} \\ \sum_{x \in E} u(x) = p \text{ est } \binom{n+p-1}{n-1} = \binom{n+p-1}{p}. \end{array} \right.$$

En prenant $E = \llbracket 1, n \rrbracket$, l'assertion (I) signifie que le nombre des n -uplets $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$ tels que $\alpha_1 + \alpha_2 + \dots + \alpha_n \leq p$ est $\binom{n+p}{p}$, l'assertion (II) que le nombre de ces points tels que $\alpha_1 + \alpha_2 + \dots + \alpha_n = p$ est $\binom{n+p-1}{n-1}$.

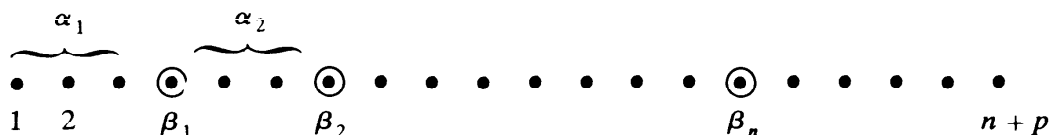
Démonstration du théorème III.4.5 :

Observons d'abord que (II) se déduit de (I) grâce à la relation de Pascal $\binom{n+p}{n} - \binom{n+p-1}{n} = \binom{n+p-1}{n-1}$. Pour prouver (I), on peut supposer que $E = \llbracket 1, n \rrbracket$. Soit \mathcal{E} l'ensemble des $\alpha \in \mathcal{F}(\llbracket 1, n \rrbracket, \llbracket 0, p \rrbracket)$ telles que $\sum_{i=1}^n \alpha(i) \leq p$, et \mathcal{F} l'ensemble des applications *strictement croissantes* de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, n+p \rrbracket$. Une application $\beta \in \mathcal{F}$ se définit de manière unique par son image, d'où $\text{card}(\mathcal{F}) = \binom{n+p}{n}$. A chaque $u \in \mathcal{E}$, associons $v = \Phi(u) \in \mathcal{F}$ définie par $v(x) = x + \sum_{i=1}^x u(i)$ ($x \in \llbracket 1, n \rrbracket$). A chaque $v \in \mathcal{F}$, associons $u = \Psi(v) \in \mathcal{E}$ définie par : $u(1) = v(1) - 1$, $u(x) = v(x) - v(x-1) - 1$ pour $x \geq 2$. On vérifie que $\Phi(u)$ (resp. $\Psi(v)$) est bien un élément de \mathcal{F} (resp. de \mathcal{E}), et que $\Psi \circ \Phi = \text{Id}_{\mathcal{E}}$, $\Phi \circ \Psi = \text{Id}_{\mathcal{F}}$, donc Φ et Ψ sont bijectives, et $\text{card}(\mathcal{E}) = \text{card}(\mathcal{F}) = \binom{n+p}{n}$. ■

Les applications Φ et Ψ de cette preuve formalisent la très simple idée suivante : représentons les points de $\llbracket 1, n+p \rrbracket$ par une échelle horizontale ; pour dénombrer les $\alpha \in \mathcal{E}$, on les écrit $\alpha = (\alpha(1), \alpha(2), \dots, \alpha(n))$, et on remplace α par

$$(\alpha(1) + 1, \alpha(1) + \alpha(2) + 2, \dots, \alpha(1) + \dots + \alpha(n) + n) = (\beta_1, \beta_2, \dots, \beta_n).$$

La suite $(\beta_1, \beta_2, \dots, \beta_n)$ est à valeurs dans $\llbracket 1, n+p \rrbracket$ et croît strictement, elle est définie de manière unique par son image $\{\beta_1, \beta_2, \dots, \beta_n\}$.



qui est une partie à n éléments de $\llbracket 1, n+p \rrbracket$: il s'agit donc de compter ces parties, d'où (I). Pour avoir (II) on tient compte de la contrainte supplémentaire $\beta_n = n+p$, et il ne reste plus qu'à compter les parties $\{\beta_1, \beta_2, \dots, \beta_{n-1}\}$ de $\llbracket 1, n+p-1 \rrbracket$.

La formule (II) permet de résoudre le problème du nombre de rangements distincts de p boules identiques dans n cases numérotées de 1 à n ; $\beta_1, \beta_2, \dots, \beta_{n-1}$ fixent les emplacements des cloisons entre les cases et on retrouve ainsi de façon imagée le nombre de solutions dans \mathbb{N}^n de l'équation $x_1 + \dots + x_n =$

$\binom{n+p-1}{n-1}$ est parfois désigné par K_n^p et appelé *nombre de combinaisons avec répétition de n objets p à p* . Ce vocabulaire peut s'expliquer ainsi : pour chaque $i \in \llbracket 1, n \rrbracket$, soit un stock *illimité* S_i d'objets *identiques* étiquetés i (par exemple, des boules de couleur i). On forme tous les échantillons possibles de taille p en puisant p objets dans la réunion des stocks S_i ; soit α_i le nombre d'objets de S_i présents dans l'échantillon α . Deux échantillons α et β seront dits *équivalents* ssi $\alpha_i = \beta_i$ pour tout i ; alors K_n^p est le nombre des classes d'équivalence d'échantillons. (Les S_i sont infinis, mais K_n^p est bien fini...)

Exemple 1 : Le nombre des monômes $X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$ de degré total $\leq p$ est $\binom{n+p}{n}$; celui des monômes $X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$ de degré total p est $\binom{n+p-1}{n-1}$.

Exemple 2 : Le nombre de n -uples $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^{*n}$ tels que $\alpha_1 + \alpha_2 + \dots + \alpha_n = p$ (avec nécessairement $p \geq n$ si l'on exige qu'il y ait effectivement des solutions) est $\binom{p-1}{n-1}$, comme on le voit immédiatement en posant $\alpha'_i = \alpha_i - 1$, ce qui ramène le problème au nombre de n -uples $(\alpha'_1, \dots, \alpha'_n) \in \mathbb{N}^n$ tels que

$$\alpha'_1 + \alpha'_2 + \dots + \alpha'_n = p - n.$$

Exemple 3 : Sommes de puissances données d'entiers consécutifs.

Pour $k \in \mathbb{N}$ et $\alpha \in \mathbb{C}$ commençons par généraliser la notation $\binom{n}{p}$ en posant *par définition* :

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} = \frac{1}{k!} \prod_{i=0}^{k-1} (\alpha-i) \quad \text{et} \quad \binom{\alpha}{0} = 1.$$

Il résulte immédiatement de cette définition que $\binom{\alpha}{0} = 1$, et que :

- si $\alpha \in \llbracket 0, k-1 \rrbracket$, $\binom{\alpha}{k} = 0$ pour $k \in \mathbb{N}^*$;
- si $\alpha \in \mathbb{C}$ et $k \in \mathbb{N}^*$ on a :

$$(3) \quad \binom{\alpha}{k} = \binom{\alpha-1}{k} + \binom{\alpha-1}{k-1};$$

- si $\alpha \in \mathbb{C}$, $\alpha \neq -1$ et $k \in \mathbb{N}$ on a :

$$(4) \quad \binom{\alpha}{k} = \frac{k+1}{\alpha+1} \binom{\alpha+1}{k+1}.$$

Cela dit, proposons-nous d'évaluer, pour $p \in \mathbb{N}$ et $n \in \mathbb{N}^*$ la somme

$$1^p + 2^p + \dots + n^p = \sum_{k=1}^n k^p,$$

que nous noterons $\mathcal{B}(n, p)$. Les premiers résultats

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}, \quad 1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6},$$

$$1^3 + 2^3 + \dots + n^3 = (1 + 2 + \dots + n)^2 = \frac{n^2(n+1)^2}{4}$$

sont sans doute bien connus du lecteur (sinon il peut vérifier par récurrence l'exactitude de ces formules).

Le passage de $\mathcal{B}(n, p)$ à $\mathcal{B}(n, p+1)$ n'étant pas évident, on peut essayer un moyen détourné, l'idée simple étant que la somme $\sum_{k=1}^n \binom{k}{p}$ est beaucoup plus facile à évaluer que la somme $\sum_{k=1}^n k^p$. Il suffit donc de savoir exprimer k^p comme « combinaison linéaire » des $\binom{k}{i}$. De manière précise montrons l'existence d'une suite double $S_{i,p}$ (i et $p \in \mathbb{N}$) telle que $S_{1,1} = 1$ et que :

$$(\forall i \in \mathbb{N}^*, \forall p \in \mathbb{N}^*) \quad n^p = \sum_{i=1}^p S_{i,p} \binom{n}{i}.$$

En effet, en supposant les $S_{i,p}$ définis pour $p \in \mathbb{N}^*$ fixé, on a

$$\begin{aligned} n^{p+1} &= n^p(n+1-1) = \sum_{i=1}^p S_{i,p} \left[(n+1) \binom{n}{i} - \binom{n}{i} \right] \\ &= \sum_{i=1}^p S_{i,p} \left[(i+1) \binom{n+1}{i+1} - \binom{n}{i} \right] \\ &= \sum_{i=1}^p S_{i,p} \left[(i+1) \binom{n}{i+1} + i \binom{n}{i} \right] \end{aligned}$$

(on a utilisé les relations (4) et (3)), d'où :

$$n^{p+1} = \sum_{i=1}^{p+1} S_{i,p+1} \binom{n}{i} \quad \text{avec} \quad S_{1,p+1} = S_{1,p} = 1,$$

$$S_{p+1,p+1} = (p+1)S_{p,p} \quad \text{et} \quad S_{i,p+1} = i(S_{i,p} + S_{i-1,p}),$$

pour $2 \leq i \leq p-1$, ce qui définit bien les $S_{i,p+1}$, donc par récurrence, tous les termes de la suite double (on notera en particulier que tous les $S_{i,p}$ sont des entiers naturels, et que $S_{p,p} = p!$ pour tout p). Ay

$k^p = \sum_{i=1}^p S_{i,p} \binom{k}{i}$, il ne reste plus qu'à sommer de $k = 1$ à n pour obtenir

$$\mathcal{B}(n, p) = \sum_{i=1}^p S_{i,p} \left(\sum_{k=1}^n \binom{k}{i} \right).$$

Or

$$\sum_{k=1}^n \binom{k}{i} = \sum_{k=1}^n \left[\binom{k+1}{i+1} - \binom{k}{i+1} \right] = \binom{n+1}{i+1}$$

par utilisation de (3) et réduction, d'où finalement

$$\boxed{\mathcal{B}(n, p) = \sum_{i=1}^p S_{i,p} \binom{n+1}{i+1}},$$

qui résout le problème posé à condition de connaître le triangle de Pascal pour les $\binom{n+1}{i+1}$, et la suite double $S_{i,p}$, qui se construit par récurrence un peu à la manière du triangle de Pascal. L'expression ci-dessus montre en particulier que, pour $p \in \mathbb{N}^*$ fixé, la somme $1^p + 2^p + \dots + k^p + \dots + n^p$ est une expression polynomiale en n , de degré $p+1$, et dont le terme dominant est $\frac{n^{p+1}}{p+1}$, résultat déjà suggéré par les cas $p = 1, 2$ ou 3 .

Exercice 1 : Soit E un ensemble qui est l'union disjointe de deux ensembles non vides Y et Z . Montrer que l'application : $\mathcal{P}(E) \rightarrow \mathcal{P}(Y) \times \mathcal{P}(Z)$, $A \mapsto (A \cap Y, A \cap Z)$ est bijective. En déduire, pour $p, q, r \in \mathbb{N}$: $\sum_{i=0}^p \binom{p}{i} \binom{q}{r-i} = \binom{p+q}{r}$. Calculer en particulier $\sum_{i=0}^n (C_n^i)^2$ et en déduire $\sum_{i=0}^n \frac{(2n)!}{(i!)^2 [(n-i)!]^2}$.

Exercice 2 : Pour $n \in \mathbb{N}^*$ montrer que $1! + 2! + \dots + n! = (n+1)! - 1$.

Exercice 3 : Montrer en utilisant la relation de Pascal que $\sum_{i=0}^k \binom{p+i}{p} = \binom{p+k+1}{p+1}$.

Exercice 4 : Soit $\alpha \in \mathbb{C}$, n et $r \in \mathbb{N}$. Montrer que $\sum_{i=0}^n \binom{\alpha-i}{r} = \binom{\alpha+1}{r+1} - \binom{\alpha-n}{r+1}$. Montrer également que $\sum_{i=0}^n (-1)^i \binom{\alpha}{i} = (-1)^n \binom{\alpha-1}{n}$.

Exercice 5 : Soit p, q, n des entiers tels que $0 \leq q < p \leq n$. Etablir la relation

$$\sum_{k=q+1}^{n-p+q+1} \binom{n-k}{p-(q+1)} \binom{k-1}{q} = \binom{n}{p}.$$

En déduire : $\sum_{k=0}^n 2^k \binom{2n-k}{n} = 2^{2n}$.

Indication. On ordonne les p -parties de $\llbracket 1, n \rrbracket$ sous la forme $i_1 < i_2 < \dots <$

par A_k celles pour lesquelles $i_{q+1} = k$. Pour la fin, transformer en somme double en remplaçant 2^k par $\sum_{i=0}^k \binom{k}{i}$.

Exercice 6 : On considère un ensemble E muni d'une loi de composition \top non présumée associative. Pour $n \in \mathbb{N}$, $n \geq 2$, et $x \in E$ on cherche à donner un sens au composé de n termes égaux à x . Pour cela on utilise des parenthèses de façon à n'avoir à opérer que sur 2 termes contigus. On désigne par \mathcal{C}_n le nombre de manières différentes dont on peut placer les parenthèses (*nombre de Catalan*). On convient que $\mathcal{C}_1 = 1$. Montrer que $\mathcal{C}_n = \sum_{k=1}^{n-1} \mathcal{C}_k \mathcal{C}_{n-k}$.

Vérifier alors par récurrence que $\mathcal{C}_n = \frac{1}{n} \binom{2n-2}{n-1}$.

Remarque. Nous verrons au chapitre sur les séries formelles comment on arrive facilement à établir ce résultat dont l'origine peut sembler mystérieuse.

Exercice 7 : Vérifier que $\frac{1}{2^{2n}} \binom{2n}{n}$ peut s'écrire sous la forme $(-1)^n \binom{-1/2}{n}$ pour $n \in \mathbb{N}$, et que $\frac{1}{n} \binom{2n-2}{n-1} \times \frac{1}{2^{2n-1}}$ peut s'écrire $(-1)^{n-1} \binom{1/2}{n}$. Montrer également que si $a \in \mathbb{C}$, $k \in \mathbb{N}$, $\binom{-a}{k}$ peut s'écrire $(-1)^k \binom{a+k-1}{k}$. En déduire, dans le cas particulier où $a \in \mathbb{N}^*$ les identités $\binom{a}{k} - \binom{a}{k-1} + \dots \mp \binom{a}{1} \pm 1 = \binom{a-1}{k}$ et (cf. exercice 1)

$$\sum_{i=0}^p (-1)^i \binom{p}{i} \binom{q-i}{r} = \binom{q-p}{q-r}.$$

Exercice 8 : Démontrer, pour $q \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$:

$$\sum_{k=0}^{n-1} \frac{1}{2^k} \binom{q+k-1}{k} = \frac{1}{2^{n-1}} \sum_{k=0}^{n-1} \binom{q+n-1}{k}.$$

Cas particulier où $q = n$.

Exercice 9 (formule du crible) : Soit E_1, E_2, \dots, E_n des ensembles finis. Démontrer :

$$\sum_{k=0}^n (-1)^k \sum_{I \subset \llbracket 1, n \rrbracket, \text{card}(I)=k} \text{card} \left(\bigcap_{i \in I} E_i \right) = 0.$$

Exercice 10 : Pour n et $p \in \mathbb{N}^*$, calculer le nombre de n -upies $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{Z}^n$ tels que $|\alpha_1| + |\alpha_2| + \dots + |\alpha_n| = p$.

Exercice 11 : Pour n et $p \in \mathbb{N}^*$, calculer le nombre de n -uples $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{Z}^n$ tels que

$$(\forall i) |\alpha_i| \leq p \quad \text{et} \quad \sup_{i \in \llbracket 1, n \rrbracket} |\alpha_i| = p.$$

Exercice 12 : Notons en abrégé $1^p + 2^p + \dots + n^p = \mathcal{B}_p$. Calculer \mathcal{B}_4 et \mathcal{B}_5 . Vérifier que

$$\begin{aligned} 3 \mathcal{B}_5 &= \mathcal{B}_1^2 (4 \mathcal{B}_1 - 1); & 4(\mathcal{B}_1 + \mathcal{B}_3) &= n(n+1)(n^2 + n + 2); \\ 6(\mathcal{B}_3 + \mathcal{B}_5) &= n^2(n+1)^2(n^2 + n + 1); & \mathcal{B}_5 + \mathcal{B}_7 &= \mathcal{B}_1^3 n(n+1) = 2 \mathcal{B}_3^2. \end{aligned}$$

Si $p, q, r, m \in \mathbb{N}^*$, prouver que $(\forall n \in \mathbb{N}^*) (\mathcal{B}_p)^m = (\mathcal{B}_q)^r$ ne peut être réalisé que si et seulement si $(p, q) = (1, 3)$ et $(m, r) = (2, 1)$.

Exercice 13 (familles spernériennes) : Soit E un ensemble fini à n éléments et (A_1, A_2, \dots, A_h) une famille de parties de E . On dit qu'une telle famille possède la propriété S si, pour tout couple (i, j) tels que $i \in \llbracket 1, h \rrbracket$, $j \in \llbracket 1, h \rrbracket$ et $i \neq j$, la relation $A_i \subset A_j$ n'est jamais vérifiée.

a) Soit k fixé $\in \llbracket 1, n \rrbracket$. Montrer que la famille des k -parties de E (pris quelconque) possède la propriété S .

b) On prend $h = 2$. Combien y a-t-il de couples (A_1, A_2) qui possèdent la propriété S ?
 c) On appelle chaîne de E une famille (C_1, C_2, \dots, C_n) de parties de E vérifiant : $(\forall i \in \llbracket 1, n \rrbracket) \text{ card } C_i = i$ et $C_i \subset C_{i+1}$ pour $i \in \llbracket 1, n-1 \rrbracket$. Combien y a-t-il de chaînes dans E ? Montrer que si une famille $\mathcal{A} = (A_1, A_2, \dots, A_h)$ de parties de E possède la propriété S , pour toute chaîne $\mathcal{C} = (C_1, C_2, \dots, C_n)$ on a $\text{card}(\mathcal{A} \cap \mathcal{C}) \leq 1$.

d) Soit $\mathcal{A} = (A_1, A_2, \dots, A_h)$ une famille de parties de E possédant la propriété S . Soit $\Gamma_{\mathcal{A}}$ l'ensemble des chaînes \mathcal{C} de E telles que $\text{card}(\mathcal{A} \cap \mathcal{C}) = 1$. On définit l'application $\varphi : \Gamma_{\mathcal{A}} \rightarrow \mathcal{A}$ qui, à toute chaîne $\mathcal{C} \in \Gamma_{\mathcal{A}}$, associe $\varphi(\mathcal{C})$, le seul élément commun à \mathcal{A} et à \mathcal{C} . Montrer que φ est surjective et calculer, en fonction de $\text{card } \mathcal{A}$, le cardinal de $\varphi^{-1}(\{A\})$, c'est-à-dire le nombre de chaînes dont l'image par φ est A .

e) Montrer que $\sum_{i=1}^h \frac{1}{\binom{n}{\text{card } A_i}} \leq 1$ et en déduire que $h \leq \sup_{p \in \llbracket 0, n \rrbracket} \binom{n}{p}$, ce dernier

nombre étant noté $\omega(n)$. Vérifier que $\omega(2n-1) = \frac{1}{2} \omega(2n)$. Existe-t-il une famille $(A_1, A_2, \dots, A_{\omega(n)})$ de parties de E possédant la propriété S ?

f) On pose $E = \llbracket 1, n \rrbracket$ et on donne n nombres réels a_1, a_2, \dots, a_n , tous > 0 (sans être nécessairement distincts). On définit l'application f de $\mathcal{P}(E)$ dans \mathbb{R} par : $f(A) = \sum_{i \in A} a_i$;

$f(\emptyset) = 0$. Montrer que, pour tout $t \in \mathbb{R}$, le nombre $g(t)$ de parties A de E telles que $f(A) = t$ est nécessairement $\leq \omega(n)$.

Exercice 14 (points sur un quadrillage) : Dans un plan euclidien rapporté à un repère orthonormé on considère, pour $n \in \mathbb{N}^*$, l'ensemble E_n de tous les points du plan de coordonnées $(\frac{i}{n}, \frac{j}{n})$ où $i \in \llbracket 0, n \rrbracket$ et $j \in \llbracket 0, n \rrbracket$.

- I. a) Nombre de points de E_n . Nombre de paires de points de E_n .
 b) Nombre de carrés dont les sommets appartiennent à E_n et dont les côtés sont parallèles aux axes. Calculer la valeur moyenne de leur périmètre (et sa limite quand $n \rightarrow +\infty$).
 c) Nombre de rectangles dont les sommets appartiennent à E_n et les côtés sont parallèles aux axes.
 d) Trouver le nombre de droites du plan passant par au moins deux points de E_n et dont la pente appartient à \mathbb{N}^* (on distinguera les cas : n pair, et n impair, et on vérifiera le résultat pour $n = 3$ et pour $n = 4$).

II. a) P étant un point quelconque du plan et M un point quelconque de E_n , calculer la somme des carrés des aires des triangles OMP obtenus quand M décrit E_n .

b) On place ensuite P en chacun des points de E_n . Calculer la somme S de toutes les sommes S obtenues au a).

c) Pour M et P décrivant E_n , évaluer la valeur moyenne μ du carré de l'aire des triangles OMP obtenus (et sa limite μ' quand $n \rightarrow +\infty$. Le nombre μ' peut-il être un nombre décimal ?).

III. M étant un point quelconque de E_n , on relie O à M par une chaîne de segments à supports parallèles aux axes, dont les extrémités sont deux points de E_n , tels que si l'on parcourt ce « chemin » de O à M , l'abscisse et l'ordonnée du point courant ne décroissent jamais strictement.

- a) Nombre de chemins allant de A à M . Retrouver ainsi $\binom{k}{i} = \binom{k-1}{i} + \binom{k-1}{i-1}$.
 b) On affecte chaque sommet M de E_n d'un coefficient égal au nombre de chemins qui le relie à l'origine O . Evaluer la somme des coefficients ainsi affectés aux sommets situés :
 b₁) sur une même parallèle à la seconde bissectrice coupant le segment $[0, 1]$ de Ox noté OA ;
 b₂) à l'intérieur ou sur le pourtour du triangle OAC (C étant le point de coordonnées $(0, 1)$) ;
 b₃) sur une même parallèle à Oy ;
 b₄) à l'intérieur ou sur le pourtour du carré $OABC$.

c) Parmi les chemins reliant O à B , combien y en a-t-il qui empruntent longueur $\frac{1}{n}$ fixé à l'avance ?

d) Parmi les chemins reliant O à B , combien y en a-t-il qui restent strictement en dessous de la diagonale OB ?

Indication : Compter les chemins OB qui touchent ou traversent cette diagonale symétrisant, éventuellement, le début de leur parcours jusqu'au premier point p sur OB .

Exercice 15 : Soit, pour $k \in \mathbb{N}$, le produit $A_k = \prod_{i=1}^p (k+i)$. Calculer simplement $\sum_{k=0}^{n-1} A_k$.

Indication : Que vaut $\frac{A_k}{k!}$? (cf. exercice 3).

Autre méthode. On pose $f(k) = \prod_{i=0}^p (k+i)$. Que vaut $f(k+1) - f(k)$?

Exercice 16 (extrait du problème de Saint-Cloud M' 1985) : Pour tout entier $n > 0$ on note A_n l'ensemble des n -uplets $a = (a_1, a_2, \dots, a_n)$ tels que $a_i = +1$ ou -1 , pour tout entier $i \in \llbracket 1, n \rrbracket$. On pose $S_i(a) = \sum_{j=1}^i a_j$.

1° a) Quelles sont les valeurs possibles de $S_n(a)$?

b) Calculer pour tout $p \in \mathbb{Z}$ le nombre $N_{n,p}$ d'éléments $a \in A_n$ tels que $S_n(a) = p$.

2° a) Montrer par récurrence sur n que, pour tout entier $p > 0$, le nombre d'éléments $a \in A_n$, tels que $S_n(a) = p$ et $S_i(a) > 0$ pour tout entier $i \in \llbracket 1, n \rrbracket$ est égal à $\frac{p}{n} N_{n,p}$.

b) En déduire le nombre d'éléments $a \in A_{2n}$, tels que $S_i(a) > 0$ pour tout entier $i \in \llbracket 1, 2n-1 \rrbracket$ et $S_{2n}(a) = 0$ (on remarquera que $S_{2n-1}(a) = 1$).

c) Retrouver le résultat du b) sans utiliser a) en utilisant une symétrie partielle.

3° a) Soit h_n le nombre d'éléments $a \in A_{2n}$, tels que $S_i(a) \neq 0$ pour tout indice $i \in \llbracket 1, 2n-1 \rrbracket$ et $S_{2n}(a) = 0$. Vérifier que si l'on pose $g_n = N_{2n,0}$ (avec la convention $g_0 = 1$), h_n est combinaison linéaire de g_n et g_{n-1} , à coefficients constants.

b) Soit B_{2n} l'ensemble : $\{a \in A_{2n}, S_i(a) \neq 0 \text{ pour tout indice } i \in \llbracket 1, 2n \rrbracket\}$. Calculer $\text{card } B_{2n}$.

4° Montrer que $g_n = \sum_{r=1}^n h_r g_{n-r}$ (considérer le premier indice i tel que $S_i(a) = 0$).

5° Soit C_{2n} l'ensemble : $\{a \in A_{2n}, S_i(a) \geq 0 \text{ pour tout } i \in \llbracket 1, 2n \rrbracket\}$. En étudiant l'application qui à tout élément $a \in C_{2n}$ associe $T(a) = (+1, a_1, \dots, a_{2n-1})$, montrer que C_{2n} a g_n éléments.

6° Posons, pour tout $a \in A_{2n}$ et tout entier $i \in \llbracket 1, 2n \rrbracket$ $u_i(a) = \frac{S_{i-1}(a) + S_i(a)}{2}$ (avec la convention $S_0(a) = 0$).

a) Montrer que $u_i(a)$ n'est jamais nul.

b) Notons $\alpha_n(a)$ le nombre d'indices $i \in \llbracket 1, 2n \rrbracket$, tels que $u_i(a)$ soit positif.

Montrer que $\alpha_n(a)$ est pair.

c) Pour tout indice $k \in \llbracket 0, n \rrbracket$, on note $\pi_{k,n}$ le nombre d'éléments $a \in A_{2n}$, tels que $\alpha_n(a) = 2k$.

Montrer que

$$\pi_{k,n} = \frac{1}{2} \sum_{r=1}^k h_r \pi_{k-r,n-r} + \frac{1}{2} \sum_{r=1}^{n-k} h_r \pi_{k,n-r}, \quad \text{si } k \in \llbracket 1, n-1 \rrbracket.$$

7° Déduire des résultats des trois questions précédentes que $\pi_{k,n} = g_k g_{n-k}$.

§ III.5 FORMULE DU BINÔME

On donne ici un anneau A ; on rappelle que la notation $m \cdot x$ (où $m \in \mathbb{Z}$ et $x \in A$) désigne l'élément $(m \cdot 1_A) x$ de A . On écrira $m \cdot x = mx$.

THÉORÈME III.5.1

Si a et b sont deux éléments **permutables** de A , on a, pour $n \in \mathbb{N}^*$, la **formule du binôme**

(I)

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$$

Démonstration :

Pour $p \in \llbracket 0, n \rrbracket$, nous noterons \mathcal{F}_p l'ensemble des p -parties de $\llbracket 1, n \rrbracket$. En posant $a = c_1$, $b = c_2$, la formule (3) du § III.2 nous donne :

$$(a + b)^n = \sum_{(i_1, i_2, \dots, i_n) \in \{1, 2\}^n} c_{i_1} c_{i_2} \dots c_{i_n}.$$

Fixons $p \in \llbracket 0, n \rrbracket$ et désignons par \mathcal{E}_p l'ensemble des $(i_1, i_2, \dots, i_n) \in \{1, 2\}^n$ pour lesquels le nombre des $k \in \llbracket 1, n \rrbracket$ vérifiant $i_k = 1$ est égal à p . Les (\mathcal{E}_p) forment une partition de $\{1, 2\}^n$, d'où

$$(a + b)^n = \sum_{p=0}^n \left(\sum_{(i_1, i_2, \dots, i_n) \in \mathcal{E}_p} c_{i_1} c_{i_2} \dots c_{i_n} \right).$$

Mais, a et b étant **permutables**, si $(i_1, i_2, \dots, i_n) \in \mathcal{E}_p$ on a toujours $c_{i_1} c_{i_2} \dots c_{i_n} = a^p b^{n-p}$ (un nombre fini de permutations des lettres a et b consécutives permettant de placer les p facteurs a en tête). Par ailleurs il est clair que \mathcal{E}_p est équipotent à l'ensemble des p -parties de $\llbracket 1, n \rrbracket$, donc $\text{card}(\mathcal{E}_p) = \binom{n}{p} = \binom{n}{n-p}$, d'où

$$(a + b)^n = \sum_{p=0}^n \text{card}(\mathcal{E}_p) a^p b^{n-p} = \sum_{p=0}^n \binom{n}{n-p} a^p b^{n-p}$$

qui équivaut à (I). ■

Formule du multinôme

Pour $n \in \mathbb{N}^*$, $p \in \mathbb{N}$, nous noterons $\mathcal{T}_{n,p}$ l'ensemble des n -uples de \mathbb{N}^n , $(\alpha_1, \alpha_2, \dots, \alpha_n)$ tels que $\alpha_1 + \alpha_2 + \dots + \alpha_n = p$. Pour

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathcal{T}_{n,p},$$

on pose : $\binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n} = \frac{p!}{\alpha_1! \alpha_2! \dots \alpha_n!}$.

LEMME 1

|| Pour $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathcal{T}_{n,p}$, le rationnel $\binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n}$ est un entier naturel ; c'est le nombre d'applications $f: \llbracket 1, p \rrbracket \rightarrow \llbracket 1, n \rrbracket$ pour lesquelles $\text{card}(f^{-1}(k)) = \alpha_k$ pour tout $k \in \llbracket 1, n \rrbracket$.

Démonstration :

Il y a une bijection évidente entre l'ensemble \mathcal{A}_α des applications $f: \llbracket 1, p \rrbracket \rightarrow \llbracket 1, n \rrbracket$ de l'énoncé, et l'ensemble \mathcal{A}'_α des suites (E_1, E_2, \dots, E_n) de parties de $\llbracket 1, p \rrbracket$ telles que

$$\text{card}(E_1) = \alpha_1, \dots, \text{card}(E_n) = \alpha_n.$$

Pour dénombrer \mathcal{A}'_α , on observe qu'on peut choisir E_1 de $\binom{p}{\alpha_1}$ manières ; qu'une fois E_1 choisi, on peut choisir arbitrairement E_2 dans $\llbracket 1, p \rrbracket \setminus E_1$, ce qui peut se faire de $\binom{p - \alpha_1}{\alpha_2}$ manières, etc. (nous laissons au lecteur le soin d'écrire une récurrence correcte). Le nombre total de choix possibles pour (E_1, E_2, \dots, E_n) est donc

$$\binom{p}{\alpha_1} \times \binom{p - \alpha_1}{\alpha_2} \times \dots \times \binom{p - \alpha_1 - \dots - \alpha_{n-1}}{\alpha_n} = \binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n}. \quad \blacksquare$$

THÉORÈME III.5.2

|| Soit a_1, a_2, \dots, a_n des éléments deux à deux permutables de A , et soit $p \in \mathbb{N}^*$. Alors :

$$(a_1 + a_2 + \dots + a_n)^p = \sum_{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n, \alpha_1 + \alpha_2 + \dots + \alpha_n = p} \binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n} a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}$$

|| (formule du multinôme).

Démonstration :

La relation (3) du § III.2 donne :

$$(a_1 + a_2 + \dots + a_n)^p = \sum_{(i_1, i_2, \dots, i_p) \in \llbracket 1, n \rrbracket^p} a_{i_1} a_{i_2} \dots a_{i_p}.$$

Avec les notations du lemme 1, les ensembles $(\mathcal{A}_\alpha)_{\alpha \in \mathcal{T}_{n,p}}$ forment une partition de $\llbracket 1, n \rrbracket^p = \mathcal{F}(\llbracket 1, p \rrbracket, \llbracket 1, n \rrbracket)$. Donc :

$$(a_1 + a_2 + \cdots + a_n)^p = \sum_{\alpha \in \mathcal{T}_{n,p}} \left(\sum_{\varphi \in \mathcal{A}_\alpha} a_{\varphi(1)} a_{\varphi(2)} \cdots a_{\varphi(p)} \right).$$

Mais, si $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathcal{T}_{n,p}$ est donné, le fait que les a_i sont deux à deux *permutables* permet de réordonner les facteurs, en sorte que pour tout $\varphi \in \mathcal{A}_\alpha$, $a_{\varphi(1)} a_{\varphi(2)} \cdots a_{\varphi(p)} = a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_n^{\alpha_n}$. Donc

$$\begin{aligned} (a_1 + a_2 + \cdots + a_n)^p &= \sum_{\alpha \in \mathcal{T}_{n,p}} \text{card}(\mathcal{A}_\alpha) a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_n^{\alpha_n} = \\ &= \sum_{\alpha \in \mathcal{T}_{n,p}} \binom{p}{\alpha_1, \alpha_2, \dots, \alpha_n} a_1^{\alpha_1} \cdots a_n^{\alpha_n}. \quad \blacksquare \end{aligned}$$

Une analyse plus poussée de la formule du multinôme nécessiterait la connaissance des groupes de permutation, que nous étudierons plus loin. Bien entendu, si l'anneau A est *commutatif*, les formules du binôme et du multinôme s'appliquent à *tous* les éléments de l'anneau.

Il importe de savoir écrire rapidement la formule du multinôme pour de petites valeurs de p , et aussi de n . Ainsi, on retiendra par cœur, sous les hypothèses du théorème III.5.2 :

$$(a_1 + a_2 + \cdots + a_n)^2 = \sum_{i=1}^n a_i^2 + 2 \sum_{1 \leq i < j \leq n} a_i a_j \quad \text{qu'on peut énoncer : « le carré d'une somme est égal à la somme des carrés plus la somme de tous les doubles produits »}.$$

De même

$$(a_1 + a_2 + \cdots + a_n)^3 = \sum_{i=1}^n a_i^3 + 3 \sum_{(i,j) \in \llbracket 1, n \rrbracket^2, i \neq j} a_i^2 a_j + 6 \sum_{1 \leq i < j < k \leq n} a_i a_j a_k$$

et aussi

$$\begin{aligned} (a + b + c)^3 &= a^3 + b^3 + c^3 + 3 \sum a^2 b + 6 abc = \\ &= a^3 + b^3 + c^3 + 3(b+c)(c+a)(a+b) \end{aligned}$$

pour a, b, c deux à deux *permutables*.

Exercice 1 : Démontrer les égalités valables pour n entier, $n \geq 2$:

$$1 - \binom{n}{1} + \binom{n}{2} - \cdots = 0; \quad \binom{n}{1} + 2 \binom{n}{2} + 3 \binom{n}{3} + \cdots = n 2^{n-1};$$

$$\binom{n}{1} - 2 \binom{n}{2} + 3 \binom{n}{3} - \cdots = 0;$$

$$2 \cdot 1 \binom{n}{2} + 3 \cdot 2 \binom{n}{3} + 4 \cdot 3 \binom{n}{4} + \cdots = n(n-1) 2^{n-2}.$$

Exercice 2 : Pour n et $p \in \mathbb{N}^*$, si t est élément d'un anneau A ,

$$\sum_{k \in \mathbb{N}} \binom{n}{k} \binom{n-k}{p-k} t^k = \binom{n}{p} (1+t)^p.$$

On rappelle que la somme qui est au premier membre ne renferme qu'un nombre fini de termes non nuls.

Exercice 3 : On pose

$$f(x) = \sum_{k=0}^n \binom{n}{k} \frac{x^{k+1}}{k+1}, \quad g(x) = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \frac{x^k}{k} \quad \text{où } x \in \mathbb{R}.$$

Grâce à la dérivation, calculer les valeurs de $f(1)$ et de $g(1)$.

Exercice 4 : Développer, dans un anneau commutatif :

$$(x_1 + x_2 + x_3)^p \quad \text{où } p \in \mathbb{N}^*; \quad (x_1 + x_2 + \dots + x_n)^5.$$

Exercice 5 : Vérifier, dans un anneau commutatif, les identités suivantes :

$$a^3 + b^3 + c^3 - 3abc = (a+b+c)(a^2 + b^2 + c^2 - bc - ca - ab)$$

$$(a+b+c)^3 = 3(a+b+c)(a^2 + b^2 + c^2) - 2(a^3 + b^3 + c^3) + 6abc$$

$$a^4 + b^4 + c^4 - 2(b^2c^2 + c^2a^2 + a^2b^2) = (a+b+c)(a+b-c)(a-b+c)(a-b-c).$$

Exercice 6 : Démontrer par récurrence, en utilisant la formule du binôme, que :

$$\binom{n}{1} \frac{1}{1} - \binom{n}{2} \frac{1}{2} + \dots + (-1)^{n-1} \binom{n}{n} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}.$$

Essayer de trouver une solution plus élégante à partir de $\sum_{k=0}^{n-1} (1-t)^k$.

Exercice 7 : Vérifier, pour $a \in \mathbb{Q}^*$, la formule de Reyley (1825) :

$$a = \left[\frac{a^6 + 45a^4 - 81a^2 + 27}{6a(a^2 + 3)^2} \right]^3 + \left[\frac{-a^4 + 30a^2 - 9}{6a(a^2 + 3)} \right]^3 + \left[\frac{-6a^3 + 18a}{(a^2 + 3)^2} \right]^3.$$

(N.B. : Cette formule montre que tout nombre rationnel peut s'écrire comme la somme de trois cubes de nombres rationnels.)

Exercice 8 : Dans un anneau commutatif, établir les identités :

$$x^n + y^n = (x+y)^n - n(x+y)^{n-2}xy + \dots + (-1)^k n \frac{(n-k-1)!}{k!(n-2k)!} (x+y)^{n-2k}(xy)^k + \dots, \text{ et}$$

$$(x+y)^n = x^n + \binom{n}{1} y(x+z)^{n-1} + \binom{n}{2} y(y-2z)(x+2z)^{n-2} + \dots \\ + \binom{n}{k} y(y-kz)^{k-1}(x+kz)^{n-k} + \dots + y(y-nz)^{n-1}.$$

Exercice 9 : Dans un anneau commutatif, démontrer l'identité de Lagrange

$$\left(\sum_{k=1}^n x_k y_k \right)^2 + \sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i)^2 = \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{j=1}^n y_j^2 \right)$$

et en déduire l'inégalité de Cauchy-Schwarz

$$\left(\sum_{k=1}^n x_k y_k \right)^2 \leq \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{j=1}^n y_j^2 \right)$$

pour des nombres réels x_k et y_k , $k \in \llbracket 1, n \rrbracket$, $n \in \mathbb{N}^*$.

Exercice 10 : Dans un anneau commutatif, prouver l'identité :

$$2 \sum_{k=1}^n a_k (a_1 + a_2 + \dots + a_k) = \left(\sum_{k=1}^n a_k \right)^2 + \sum_{k=1}^n a_k^2.$$

Exercice 11 : On dit qu'un élément x d'un anneau est *nilpotent* s'il existe un entier $n \geq 1$ tel que $x^n = 0$ (l'anneau n'est pas supposé commutatif). Montrer que si x est nilpotent, alors $1_A - x$ est inversible (cf. la formule (6) du § III.2). Montrer que si deux éléments nilpotents x et y de A sont permutables, alors $x + y$ est nilpotent, ainsi que xy .

Soit a un élément fixé dans A . On considère l'application $u : k \rightarrow k, x \mapsto ax - xa$. Montrer que si a est nilpotent, il existe un entier q tel que $u^{<q>} = 0$, c'est-à-dire $(\forall x \in A) u \circ u \dots \circ u(x) = 0$.

Indication : Commencer par examiner les cas $a^2 = 0$ ou $a^3 = 0$.

Exercice 12 : Soit A un anneau. On suppose que \mathbb{Q} est un sous-anneau de A , ce qui permet de donner un sens au produit d'un rationnel quelconque par un élément quelconque de A . Soit x un élément nilpotent de A (cf. exercice 11). On définit $\exp x = 1 + \frac{x}{1!} + \dots + \frac{x^k}{k!} + \dots$ (il s'agit en réalité d'une somme finie). Montrer que si x et y sont deux nilpotents *permutables*, alors on a : $\exp(x + y) = \exp x \times \exp y$. Montrer aussi que $1 - \exp x$ est nilpotent.

§ III.6 SOUS-GROUPES ADDITIFS DE \mathbb{Z} . APPLICATION AUX GROUPES

THÉORÈME III.6.1 (Division euclidienne dans \mathbb{Z})

|| Si $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$, il existe au moins un couple $(q, r) \in \mathbb{Z}^2$ tel que
|| $a = bq + r$ et $|r| < |b|$.

Démonstration :

Il suffit de procéder à la division euclidienne de $|a|$ par $|b|$ dans \mathbb{N} : $|a| = |b| q_1 + r_1$ avec $r_1 < |b|$. Si l'on pose

$$|a| = \varepsilon_1 a \quad \text{et} \quad |b| = \varepsilon_2 b \quad (\varepsilon_1 \text{ et } \varepsilon_2 \in \{-1, +1\})$$

alors $a = b(\varepsilon_1 \varepsilon_2 q_1) + \varepsilon_1 r_1$ et le couple $(q, r) = (\varepsilon_1 \varepsilon_2 q_1, \varepsilon_1 r_1)$ convient. ■

Remarque 1 : Si $a = bq + r$ avec $-|b| < r < 0$, on peut écrire $a = b(q - \varepsilon_2) + (|b| + r)$, c'est-à-dire $a = bq' + r'$ avec $0 < r' < |b|$. Dans le théorème III.6.1, on peut donc écrire $0 \leq r < |b|$.

Remarque 2 : Le lecteur s'étonnera peut-être qu'on n'impose pas l'unicité du couple (q, r) . Il est vrai que le plus souvent $b > 0$; on convient alors d'imposer une condition supplémentaire sur le reste, par exemple $0 \leq r < b$, ou encore $\frac{-b}{2} < r \leq \frac{b}{2}$ et dans ces conditions il y a unicité de (q, r) . Mais dans l'étude des anneaux euclidiens généraux (c'est-à-dire ceux où l'on peut définir une division euclidienne), l'unicité est rarement satisfaite et n'est heureusement pas indispensable (voir par exemple au chapitre des nombres complexes les exercices sur les entiers de Gauss) pour montrer que tout anneau euclidien est principal (cf. la définition III.7.4), ce qui lui confère des propriétés arithmétiques tout à fait analogues à celle

Donnons-nous $a \in \mathbb{Z}$; l'ensemble $\{\lambda a\}_{\lambda \in \mathbb{Z}}$ des multiples entiers de a , que l'on note $a\mathbb{Z}$, est un sous-groupe du groupe additif de \mathbb{Z} .

On a : $a\mathbb{Z} = (-a)\mathbb{Z}$; de plus, si $b \in \mathbb{Z}$ est tel que $b\mathbb{Z} = a\mathbb{Z}$, et si $a \neq 0$, $b = \lambda a$ et $a = \mu b$ avec $\lambda, \mu \in \mathbb{Z}$, d'où $a = \lambda \mu a$, $(1 - \lambda \mu)a = 0$, donc $\lambda \mu = 1$ car \mathbb{Z} est intègre, donc $\lambda = \mu = 1$ ou $\lambda = \mu = -1$ (cf. exemple 3, § II.5). Par suite, pour tout $a \in \mathbb{Z}$, il existe un et un seul $\alpha \in \mathbb{N}$ tel que $a\mathbb{Z} = \alpha\mathbb{Z}$: c'est $\alpha = |a|$, et les seuls $\beta \in \mathbb{Z}$ tels que $a\mathbb{Z} = \beta\mathbb{Z}$ sont α et $-\alpha$. Le naturel α s'appelle *générateur positif* de $a\mathbb{Z}$.

THÉOREME III.6.2

|| L'application $\alpha \mapsto \alpha\mathbb{Z}$ est une bijection de \mathbb{N} sur l'ensemble des sous-groupes additifs de \mathbb{Z} ; en particulier, pour tout sous-groupe G de $(\mathbb{Z}, +)$, il existe $\alpha \in \mathbb{N}$ unique tel que $G = \alpha\mathbb{Z}$.

Démonstration :

Il suffit de prouver la seconde assertion, évidente si $G = \{0\}$. Soit donc un sous-groupe quelconque $G \neq \{0\}$ de $(\mathbb{Z}, +)$. Si $k \in G \setminus \{0\}$, k et $-k$ sont dans G , donc $G \cap \mathbb{N}^*$ est non vide. Notons α le plus petit élément de $G \cap \mathbb{N}^*$ (partie de l'ensemble bien ordonné \mathbb{N}). Comme G est un sous-groupe additif de \mathbb{Z} , on a d'abord $\alpha\mathbb{Z} \subset G$. Soit alors $x \in G$. Il est possible de trouver $(q, r) \in \mathbb{Z}^2$ tel que $x = \alpha q + r$, $0 \leq |r| < \alpha$. Alors $r = x - \alpha q \in G$ puisque G est un sous-groupe et que $\alpha q \in G$. Donc $|r| \in G$, et cependant $|r| \notin G \cap \mathbb{N}^*$ car $|r| < \alpha$. Nécessairement $|r| = 0$, donc $r = 0$, donc $x = \alpha q = q\alpha \in \alpha\mathbb{Z}$, d'où $G \subset \alpha\mathbb{Z}$. ■

Ordre d'un élément dans un groupe.

Soit G un groupe arbitraire, noté multiplicativement. Si $x \in G$, nous avons défini au § II.5 l'homomorphisme de groupes $\varphi_x : (\mathbb{Z}, +) \rightarrow G$, $m \mapsto x^m$. Nous avons vu que le noyau $\text{Ker}(\varphi_x)$ est un sous-groupe de $(\mathbb{Z}, +)$. Nous savons maintenant qu'il existe un, et un seul, $\alpha \in \mathbb{N}$ tel que $\text{Ker}(\varphi_x) = \alpha\mathbb{Z}$.

DÉFINITION III.6.1

Si x est élément d'un groupe G , on dit que x est **sans torsion**, ou **d'ordre infini**, ssi $\text{Ker}(\varphi_x) = \{0\}$. On dit que x est de « **de torsion** », ou **d'ordre fini**, ssi $\text{Ker}(\varphi_x) \neq \{0\}$. Dans ce dernier cas, l'entier $\alpha \in \mathbb{N}^*$ tel que $\text{Ker}(\varphi_x) = \alpha\mathbb{Z}$ s'appelle **l'ordre de x** .

D'après ce qui précède, si x est sans torsion, l'application $\varphi_x : \mathbb{Z} \rightarrow G$ est injective et c'est un isomorphisme du groupe $(\mathbb{Z}, +)$ sur le sous-groupe de G (dit **engendré par x**) formé des $\{x^m\}_{m \in \mathbb{Z}}$, et égal à $\text{Im}(\varphi_x)$.

Si x est de torsion, et si α est son ordre, alors α est l'unique élément de \mathbb{Z}_+^* possédant les propriétés suivantes :

$$x^\alpha = e_G \quad \text{et} \quad (\forall k \in \mathbb{Z}) \quad x^k = e_G \Leftrightarrow k \in \alpha\mathbb{Z} :$$

THÉORÈME III.6.3

Si x est un élément de torsion de G , d'ordre α , alors le sous-groupe engendré par x est fini ; et l'application :

$$[[0, \alpha - 1]] \rightarrow \text{Im}(\varphi_x), k \mapsto x^k$$

est bijective, donc $\text{card}(\text{Im}(\varphi_x)) = \alpha$.

Démonstration :

Si $x^k = e_G$ et $k \in [[0, \alpha - 1]]$, alors

$$k \in [[0, \alpha - 1]] \cap \alpha\mathbb{Z} = \{0_{\mathbb{Z}}\},$$

donc $k = 0$. Donc si k_1 et $k_2 \in [[0, \alpha - 1]]$, $k_1 \leq k_2$ sont tels que $x^{k_1} = x^{k_2}$, alors $x^{k_2 - k_1} = e_G$, d'où l'injectivité.

Soit maintenant $k \in \mathbb{Z}$. Par division euclidienne $k = \alpha q + r$, avec $0 \leq r < \alpha$, d'où

$$x^k = x^{\alpha q + r} = x^{\alpha q} x^r = (x^\alpha)^q x^r = e_G x^r = x^r$$

d'où la surjectivité. ■

Remarque 3 : Si G est noté additivement il faut prendre garde que, dans ce qui précède, pour $k \in \mathbb{Z}$, $\varphi_x(k)$ s'écrit kx et non x^k .

DÉFINITION III.6.2

Un groupe G est dit **monogène** ssi il existe $x \in G$ tel que G soit engendré par x (i.e. $G = \text{Im}(\varphi_x)$).

Il y a deux types seulement de groupes monogènes : ceux qui sont *infinis*, isomorphes à $(\mathbb{Z}, +)$ et ceux qui sont *finis* qu'on appelle **groupes cycliques**.

§ III.7 NOTION D'IDÉAL D'UN ANNEAU COMMUTATIF

Considérons un sous-groupe additif G de \mathbb{Z} . Nous savons que $G = \alpha\mathbb{Z}$, avec $\alpha \in \mathbb{N}$. On constate alors immédiatement qu'en plus de sa stabilité pour l'addition et la soustraction provenant du fait que c'est un sous-groupe, G possède la propriété suivante : pour tout $\lambda \in \mathbb{Z}$ et tout :

$\lambda x \in G$. On résume toutes ces propriétés en disant que G est un idéal de l'anneau \mathbb{Z} . De manière générale, on pose :

DÉFINITION III.7.1

⌋ Dans un anneau **commutatif** A , on appelle **idéal** tout sous-ensemble \mathfrak{a} de A tel que :

⌋ (I) \mathfrak{a} est un sous-groupe additif de A

⌋ (II) $(\forall (\lambda, x) \in A \times \mathfrak{a}), \lambda x \in \mathfrak{a}$.

$\{0_A\}$ est un idéal de A , appelé *idéal nul* ; A est un idéal de A , appelé *idéal unité*.

Pour qu'un idéal \mathfrak{a} de A soit égal à A , il faut et il suffit qu'il contienne 1_A ; ou aussi, qu'il contienne au moins un élément inversible de A ; en particulier, si A est un *corps commutatif*, il n'y a dans A que deux idéaux : $\{0_A\}$ et A .

On l'a vu ci-dessus, dans l'anneau \mathbb{Z} , tout sous-groupe additif est un idéal. Mais cette propriété cesse d'être vraie en général : par exemple dans l'anneau \mathbb{Z}^2 défini au chapitre II, § 5, le sous-groupe formé de $\{a, a\}_{a \in \mathbb{Z}}$ n'est pas un idéal. La proposition suivante est de démonstration immédiate :

PROPOSITION III.7.1

⌋ Soit A un anneau commutatif.

⌋ (I) Si $(\mathfrak{a}_i)_{i \in I}$ est une famille d'idéaux de A , alors $\bigcap_{i \in I} \mathfrak{a}_i$ est un idéal de A .

⌋ (II) Si B est un anneau commutatif et si $\rho : A \rightarrow B$ est un homomorphisme d'anneaux, pour tout idéal \mathfrak{b} de B , l'image réciproque $\rho^{-1}(\mathfrak{b})$ est un idéal de A .

En particulier, l'image réciproque $\rho^{-1}(0_B)$ de $\{0_B\}$ par l'homomorphisme d'anneaux commutatifs $\rho : A \rightarrow B$ est un idéal de A , noté $\text{Ker}(\rho)$ et appelé **noyau** de ρ . Un homomorphisme d'anneaux est en particulier un homomorphisme entre les groupes additifs de ces anneaux, par conséquent on déduit du théorème II.4.3 le suivant :

PROPOSITION III.7.2

⌋ Soit $\rho : A \rightarrow B$ un homomorphisme d'anneaux commutatifs. Pour que ρ soit injectif, il faut et il suffit que $\text{Ker}(\rho) = \{0_A\}$.

Combinaisons A-linéaires.

Soit $a = (a_i)_{i \in I}$ une famille finie ou infinie d'éléments de l'anneau commutatif A ; notons $A^{(I)}$ l'ensemble des familles à support

d'éléments de A (si I est fini, $A^{(I)} = A^I = \mathcal{F}(I, A)$). A chaque élément $\lambda = (\lambda_i)_{i \in I}$ de $A^{(I)}$, associons l'élément $\mathcal{C}_a(\lambda) = \sum_{i \in I} \lambda_i a_i$: un tel élément

s'appelle une **combinaison A -linéaire des a_i** . Il est important de remarquer qu'il s'agit d'une combinaison A -linéaire finie. Si l'on considère alors l'ensemble $\mathcal{C}(a)$ de toutes ces combinaisons A -linéaires $\mathcal{C}_a(\lambda)$, il est immédiat que $\mathcal{C}(a)$ est un idéal de A , et que chaque a_i appartient à cet idéal, puisque $a_i = 1_A \times a_i = \sum_{j \in I} \delta_{i,j} a_j$, où $\delta_{i,j} = 0$ pour $j \neq i$ et $\delta_{i,j} = 1_A$ pour $j = i$.

Réciproquement, supposant qu'un idéal \mathfrak{a} de A contienne tous les a_i : par les axiomes (I) et (II) de la définition III.7.1, on voit que $\mathcal{C}(a) \subset \mathfrak{a}$. Donc l'intersection de la famille des idéaux de A qui contiennent tous les a_i n'est autre que l'idéal $\mathcal{C}(a)$. On pose :

DÉFINITION III.7.2

*Si $(a_i)_{i \in I}$ est une famille d'éléments de l'anneau commutatif A , on appelle **idéal engendré par les a_i** l'idéal formé par toutes les combinaisons A -linéaires des a_i .*

Si I est fini, égal à $\llbracket 1, n \rrbracket$, l'idéal engendré par (a_1, a_2, \dots, a_n) est simplement l'ensemble de toutes les sommes $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n$, où le n -uplet $(\lambda_1, \lambda_2, \dots, \lambda_n)$ parcourt A^n .

Si E est une partie de A , l'idéal engendré par E est l'ensemble des sommes finies $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_p a_p$, où $p \in \mathbb{N}^*$ est arbitraire, où les $a_i (i \in \llbracket 1, p \rrbracket)$ sont arbitraires dans E et où les $\lambda_i (i \in \llbracket 1, p \rrbracket)$ sont arbitraires dans A .

Idéaux de type fini, idéaux principaux.

Soit \mathfrak{a} un idéal de l'anneau commutatif A . Les parties de \mathfrak{a} (resp. les familles d'éléments de \mathfrak{a}) telles que \mathfrak{a} soit l'idéal engendré par cette partie (resp. cette famille) sont dites **génératrices de l'idéal \mathfrak{a}** .

DÉFINITION III.7.3

*Un idéal \mathfrak{a} d'un anneau commutatif A est dit **de type fini** ssi il peut être engendré par une famille finie (ou, ce qui revient au même, par une partie finie).
L'idéal \mathfrak{a} est dit **principal** ssi il peut être engendré par un seul élément.*

Si $a \in A$, l'idéal engendré par a est l'ensemble $\{\lambda a\}_{\lambda \in A}$ qu'on peut appeler l'ensemble des multiples de a . On le note aA : c'est un idéal principal de A . Observons que, pour tout élément inversible $\varepsilon \in A$, on a : $aA = (\varepsilon a)A$.

Notons encore \mathcal{U}_A le groupe multiplicatif des éléments inversibles de l'anneau commutatif A . Sur A , la relation binaire \mathcal{R} définie par :

$$(x \mathcal{R} y \text{ ssi } \exists \varepsilon \in \mathcal{U}_A | y = \varepsilon x),$$

est d'équivalence ; x et y sont dits **associés** ssi $x \mathcal{R} y$. On vient de voir que deux éléments associés de A engendrent le même idéal principal. La réciproque n'est pas vraie en général, mais on a :

PROPOSITION III.7.3

|| Si l'anneau A est **intègre**, deux éléments x et y de A sont **associés** ssi $xA = yA$.

Démonstration :

On sait déjà que si x et y sont associés, $xA = yA$. Supposons donc $xA = yA$. Cela signifie qu'il existe u et v dans A tels que $y = ux$ et $x = vy$, d'où $y = uvx$ et $x = vux$. Si $x = y = 0_A$, ils sont évidemment associés. Si $x \neq 0_A$, de $x = vux$ on déduit $x(1_A - uv) = 0_A$, d'où $uv = 1_A$ puisque A est intègre, donc u et v sont tous deux inversibles, et x et y sont bien associés. ■

En raison de la proposition III.7.3, la notion d'idéal principal prend un intérêt tout particulier dans les anneaux intègres : si l'on suppose l'anneau A intègre, désignons par \mathcal{A} l'ensemble des classes d'éléments associés de A . A chaque idéal principal \mathfrak{a} de A , associons l'ensemble $\Gamma(\mathfrak{a})$ des $a \in A$ tels que $\mathfrak{a} = aA$. D'après la proposition III.7.3 $\Gamma(\mathfrak{a}) \in \mathcal{A}$, et d'après tout ce qui précède, on voit que l'application : $\mathfrak{a} \mapsto \Gamma(\mathfrak{a})$, de l'ensemble \mathcal{D}_A des idéaux principaux de A , dans l'ensemble \mathcal{A} des classes d'éléments associés de A , est bijective.

On a $\Gamma(\{0_A\}) = \{0_A\}$, $\Gamma(A) = \mathcal{U}_A$.

On a vu que les idéaux de l'anneau \mathbb{Z} sont ses sous-groupes additifs. Le théorème III.6.2 prouve que ces idéaux sont tous principaux. On pose :

DÉFINITION III.7.4

Un anneau commutatif est dit **principal** ssi

- (I) il est **intègre**, et :
- (II) tous ses idéaux sont **principaux**.

Exemple 1 : On vient de voir que l'anneau \mathbb{Z} est principal.

Exemple 2 : Tout corps commutatif est un anneau principal.

Somme d'idéaux.

Dans l'anneau commutatif A , considérons une famille d'idéaux $(\mathfrak{a}_i)_{i \in I}$; en raison de l'axiome (II) de la définition III.7.1, l'idéal en

réunion des \mathfrak{a}_i se réduit à l'ensemble des sommes $\sum_{i \in I} x_i$, où $(x_i)_{i \in I}$ est une famille à support fini arbitraire telle que $(\forall i \in I, x_i \in \mathfrak{a}_i)$. Pour cette raison, l'idéal engendré par l'ensemble $\bigcup_{i \in I} \mathfrak{a}_i$ s'appelle **somme des idéaux** \mathfrak{a}_i , et se note $\sum_{i \in I} \mathfrak{a}_i$.

Si I est *fini*, égal à $\llbracket 1, n \rrbracket$, on le note aussi $\mathfrak{a}_1 + \mathfrak{a}_2 + \dots + \mathfrak{a}_n$: c'est l'image de l'application

$$\mathfrak{a}_1 \times \mathfrak{a}_2 \times \dots \times \mathfrak{a}_n \rightarrow A, \quad (x_1, x_2, \dots, x_n) \mapsto x_1 + x_2 + \dots + x_n.$$

Il faut prendre garde qu'en général $\bigcup_{i \in I} \mathfrak{a}_i$ n'est pas un idéal, il ne faudra donc pas le confondre avec $\sum_{i \in I} \mathfrak{a}_i$, qui en est un. On peut considérer en particulier la somme de *deux* idéaux, ce qui définit une loi de composition $(\mathfrak{a}, \mathfrak{b}) \mapsto \mathfrak{a} + \mathfrak{b}$ sur l'ensemble des idéaux de A . Cette opération est évidemment commutative et associative. De plus, on notera que $\mathfrak{a} + \mathfrak{a} = \mathfrak{a}$ et $\mathfrak{a} + A = A$ pour tout idéal \mathfrak{a} de A .

Exercice 1 : Soit A un anneau commutatif non nul. On appelle **nilradical** de A l'ensemble des éléments *nilpotents* de A , c'est-à-dire l'ensemble $\mathcal{N}(A)$ formé des $x \in A$ tels que $x^n = 0$ pour au moins un $n \in \mathbb{N}^*$. Montrer que $\mathcal{N}(A)$ est un idéal de A .

Exercice 2 : Soit \mathfrak{a} un idéal d'un anneau commutatif A . On appelle **racine** de \mathfrak{a} (que l'on peut noter $\sqrt{\mathfrak{a}}$) l'ensemble des $x \in A$ tels que $x^n \in \mathfrak{a}$ pour au moins un $n \in \mathbb{N}^*$. Montrer que $\sqrt{\mathfrak{a}}$ est un idéal de A . Que se passe-t-il si $\mathfrak{a} = \{0_A\}$? Montrer que : $\sqrt{\sqrt{\mathfrak{a}}} = \sqrt{\mathfrak{a}}$ et que, si \mathfrak{a} et \mathfrak{b} sont des idéaux de A , on a :

$$\sqrt{\mathfrak{a}} \cap \sqrt{\mathfrak{b}} = \sqrt{\mathfrak{a} \cap \mathfrak{b}}; \quad \sqrt{\mathfrak{a}} + \sqrt{\mathfrak{b}} \subset \sqrt{\mathfrak{a} + \mathfrak{b}}; \quad \sqrt{\mathfrak{a} + \mathfrak{b}} = \sqrt{\sqrt{\mathfrak{a}} + \sqrt{\mathfrak{b}}}.$$

Exercice 3 : Dans un anneau commutatif non nul A , un idéal \mathfrak{a} est dit **maximal** ssi : $\mathfrak{a} \neq A$, et les seuls idéaux de A contenant \mathfrak{a} sont \mathfrak{a} et A . On donne un ensemble *fini* non vide, un corps commutatif K et on considère l'anneau K^E .

a) Pour chaque $a \in E$, l'ensemble $\mathfrak{M}_a = \{f \in K^E \mid f(a) = 0\}$ est un idéal maximal de K^E .

b) Montrer que $a \mapsto \mathfrak{M}_a$ est une bijection de E sur l'ensemble des idéaux maximaux de K^E .

c) Montrer que cette propriété tombe en défaut si E n'est plus supposé fini :

Exercice 4 : Donner un anneau commutatif A , si \mathfrak{a} et \mathfrak{b} sont deux idéaux, on définit le **produit** de \mathfrak{a} et \mathfrak{b} , que l'on note $\mathfrak{a}\mathfrak{b}$: c'est l'idéal \mathfrak{J} de A formé de toutes les sommes possibles

$\sum_{i=1}^n a_i b_i$, où $n \in \mathbb{N}^*$, $(a_1, \dots, a_n) \in \mathfrak{a}^n$ et $(b_1, \dots, b_n) \in \mathfrak{b}^n$, sont arbitraires. Vérifier que \mathfrak{J} est

bien un idéal et qu'on a les propriétés suivantes : a) la loi $(\mathfrak{a}, \mathfrak{b}) \mapsto \mathfrak{a}\mathfrak{b}$ est associative, commutative, et distributive par rapport à la somme d'idéaux : $\mathfrak{a}(\mathfrak{b} + \mathfrak{c}) = \mathfrak{a}\mathfrak{b} + \mathfrak{a}\mathfrak{c}$;

b) $\mathfrak{a}\mathfrak{b} \subset \mathfrak{a} \cap \mathfrak{b}$; $\sqrt{\mathfrak{a}} \cap \sqrt{\mathfrak{b}} = \sqrt{\mathfrak{a}\mathfrak{b}} = \sqrt{\mathfrak{a} \cap \mathfrak{b}}$ (cf. exercice 2).

Exercice 5 : Soit A un anneau commutatif non nul. Deux idéaux $\mathfrak{a}, \mathfrak{b}$ de A sont dits **étrangers** ssi $\mathfrak{a} + \mathfrak{b} = A$. Montrer : si \mathfrak{a} et \mathfrak{b} sont étrangers, alors $(\forall (m, n) \in \mathbb{N}^2)$, \mathfrak{a}^m et \mathfrak{b}^n

et qu'on a : $a \bar{b} = a \cap \bar{b}$. Montrer que si \bar{b} est étranger à chacun des idéaux a_i pour $1 \leq i \leq n$, alors \bar{b} est étranger au produit (cf. exercice 4) $a_1 a_2 \dots a_n$. On fait maintenant l'hypothèse : $n \geq 2$, a_1, a_2, \dots, a_n sont des idéaux de A , et pour tout $i \in \llbracket 1, n \rrbracket$, a_i est étranger à $\bigcap_{j \neq i} a_j$. Montrer :

$$(\forall p \in \mathbb{N}^*), \quad \bigcap_{i=1}^p a_i^p = (a_1 a_2 \dots a_n)^p = \left(\bigcap_{i=1}^p a_i \right)^p.$$

Montrer enfin que deux idéaux maximaux et distincts dans A sont étrangers.

Exercice 6 : Soit G un sous-groupe additif de \mathbb{Q} . Montrer qu'il n'y a que deux cas possibles : ou bien il existe $a \in \mathbb{Q}_+$ tel que $G = a\mathbb{Z} = \{ma\}_{m \in \mathbb{Z}}$; ou bien G est **partout dense** dans \mathbb{Q} , c'est-à-dire rencontre tout intervalle $] \alpha, \beta [$ où $\alpha \in \mathbb{Q}$, $\beta \in \mathbb{Q}$, $\alpha < \beta$.

Exercice 7 : Pour qu'un sous-anneau A de \mathbb{Q} soit *partout dense* (cf. exercice 6), il faut et il suffit que $A \cap]0, 1[\neq \emptyset$.

Exercice 8 : Dans un anneau commutatif non nul A , un idéal \mathfrak{p} est dit **premier** s'il vérifie les conditions suivantes : (I) $\mathfrak{p} \neq A$; (II) $(\forall (x, y) \in A^2), (xy \in \mathfrak{p}) \Rightarrow (x \in \mathfrak{p} \text{ ou } y \in \mathfrak{p})$.

a) Si \mathfrak{p} est un idéal premier de A et si a_1, a_2, \dots, a_n sont des idéaux de A tels que $a_1 a_2 \dots a_n \subset \mathfrak{p}$, alors il existe i tel que $a_i \subset \mathfrak{p}$.

b) Soit E une partie de A , stable pour l'addition et la soustraction et la multiplication, et $\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_n$ des idéaux premiers de A tels que $E \subset \bigcup_{i=1}^n \mathfrak{p}_i$. Montrer qu'il existe i tel que $E \subset \mathfrak{p}_i$.

c) Montrer que tout idéal maximal de A est premier. (N.B. : la réciproque est fausse.)

Exercice 9 : On dit qu'un idéal \mathfrak{I} d'un anneau commutatif A est **primaire** ssi,

$$(\forall (x, y) \in A^2)(xy \in \mathfrak{I} \text{ et } x \notin \mathfrak{I}) \Rightarrow (\exists n \in \mathbb{N}^* \text{ tel que } y^n \in \mathfrak{I}).$$

Montrer que la racine (cf. exercice 2) d'un idéal primaire est un idéal premier. Si \mathfrak{M} est un idéal maximal de A , montrer que les puissances \mathfrak{M}^p ($p \in \mathbb{N}^*$) de \mathfrak{M} sont des idéaux primaires ayant \mathfrak{M} pour racine.

Exercice 10 : Soit A l'ensemble des fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ de la forme

$$x \mapsto A_0(x) + A_1(x) \sin x + \dots + A_n(x) \sin^n x$$

où $n \in \mathbb{N}$ est arbitraire et où les $A_i(x)$ sont des fonctions polynomiales arbitraires à coefficients dans \mathbb{R} . Vérifier que A est un sous-anneau intègre de l'anneau $\mathcal{F}(\mathbb{R}, \mathbb{R})$, mais que ce n'est pas un anneau principal.

Exercice 11 : Soit A l'ensemble des nombres réels de la forme $x + y\sqrt{10}$ où x et $y \in \mathbb{Z}$. Montrer que A est un anneau commutatif. Vérifier qu'il est intègre mais qu'il n'est pas principal (par exemple l'idéal engendré par 2 et $\sqrt{10}$ n'est pas principal).

Chapitre IV

NOTIONS D'ARITHMÉTIQUE

§ IV.1 CONGRUENCES DANS \mathbb{Z} , ANNEAUX $\mathbb{Z}/n\mathbb{Z}$

Divisibilité

DÉFINITION IV.1.1

Soit A un anneau **intègre**. On dit que l'**élément** $a \in A$ **divise** l'**élément** $b \in A$ (ou que a est **diviseur** de b ou que b est **divisible par** a , et l'on écrit $a|b$) ssi b est **multiple de** a , c'est-à-dire : il existe $c \in A$ tel que $b = ac$.

Dire que a divise b signifie donc : $b \in aA$ (cf. § III.7), ce qui équivaut à $bA \subset aA$. La relation $a|b$ est **réflexive** ($a|a$ pour tout a) et **transitive** ($a|b$ et $b|c$ entraînent $a|c$). On dit que c'est un **préordre**, mais ce n'est pas une relation d'ordre à cause du fait qu'on peut avoir à la fois $a|b$ et $b|a$ avec $b \neq a$. Cela se produit si, et seulement si, a et b sont **associés** dans A , c'est-à-dire $aA = bA$ (cf. la preuve de la proposition III.7.3). Toutefois il importe de remarquer que la relation $a|b$ de l'anneau \mathbb{Z} induit sur le sous-ensemble \mathbb{N} une relation d'ordre.

Notons aussi que si $a|b$ et $a \neq 0$, il y a **unicité** du $c \in A$ tel que $a = bc$. On le note $c = a/b$ ou $\frac{a}{b}$ et on l'appelle **quotient exact** de a par b .

Notons enfin que si K est le corps des fractions de l'anneau intègre A , le vocabulaire utilisé dans la définition IV.1.1 peut s'étendre à deux éléments a et $b \in K$: on dit que a divise b s'il existe $c \in A$ tel que $b = ac$.

Congruences dans \mathbb{Z}

Soit n un entier relatif non nul. La relation binaire \mathcal{R} définie sur \mathbb{Z} par : $x \mathcal{R} y$ ssi n divise $x - y$, est d'**équivalence**.

DÉFINITION IV.1.2

Deux entiers relatifs x et y sont dits **congrus modulo n** ssi $x - y$ est multiple de n . S'il en est ainsi on écrit

(I)

$$x \equiv y \pmod{n}$$

ce qui se lit : « x et y sont congrus modulo n ». Une relation du type (I) s'appelle une **congruence modulo n** ; l'entier n s'appelle le **module** de la congruence.

Il est clair que la congruence modulo n est la même chose que la congruence modulo $(-n)$, et que $x \equiv y \pmod{n}$ équivaut à $x - y \in n\mathbb{Z}$. En particulier si $n = 0$ la congruence modulo 0 n'est autre que l'égalité dans \mathbb{Z} .

Voici quelques propriétés élémentaires des congruences dont la démonstration est laissée au lecteur :

(CG1) Pour a, b, a', b' dans \mathbb{Z} :

$$[a \equiv a' \pmod{n} \text{ et } b \equiv b' \pmod{n}] \Rightarrow [a + b \equiv a' + b' \pmod{n}],$$

qui exprime la compatibilité de la congruence mod (n) avec l'addition.

(CG2) Pour a, b, a', b' dans \mathbb{Z} :

$$[a \equiv a' \pmod{n} \text{ et } b \equiv b' \pmod{n}] \Rightarrow [ab \equiv a'b' \pmod{n}].$$

C'est la compatibilité de la congruence mod (n) avec la multiplication, qui se démontre en écrivant $a' = a + \lambda n$ ($\lambda \in \mathbb{Z}$), $b' = b + \mu n$ ($\mu \in \mathbb{Z}$), d'où $aa' = ab + \nu n$.

(CG3) Pour tout $k \in \mathbb{N}$, $[a \equiv b \pmod{n}] \Rightarrow [a^k \equiv b^k \pmod{n}]$ qui se démontre par récurrence à partir de (CG2).

(CG4) Pour tout $\lambda \in \mathbb{N}^*$ $[a \equiv b \pmod{n}] \Leftrightarrow [\lambda a \equiv \lambda b \pmod{\lambda n}]$.

(CG5) Si $a \equiv b \pmod{n}$ alors $a \equiv b \pmod{n'}$ pour tout diviseur n' de n .

(CG6) Si $d \neq 0$ divise a, b et n , alors $a \equiv b \pmod{n}$ implique : $\frac{a}{d} \equiv \frac{b}{d} \pmod{\left(\frac{n}{d}\right)}$, mais attention ! si $d \neq 0$ divise seulement a et b , la congruence $a \equiv b \pmod{n}$ n'entraîne pas, en général, $\frac{a}{d} \equiv \frac{b}{d} \pmod{n}$.

Soit γ une classe de congruence modulo n , ($n \neq 0$), et a un élément fixe de γ . Alors par définition $\gamma = \{a + \lambda n\}_{\lambda \in \mathbb{Z}}$. D'après l'étude de la division euclidienne dans \mathbb{Z} (cf. § III.6), on a, en supposant dorénavant que $n \geq 1$, compte tenu de la remarque qui suit la définition IV.1.2

PROPOSITION IV.1.1

|| Dans chaque classe de congruence γ modulo n , il y a un et un seul élément de $\llbracket 0, n-1 \rrbracket$, autrement dit $\gamma \cap \llbracket 0, n-1 \rrbracket$ est un singleton.

On convient de désigner par $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes de congruence mod (n) dans \mathbb{Z} . Si $a \in \mathbb{Z}$, la classe $\gamma_n(a)$ est l'ensemble $\{a + \lambda n\}_{\lambda \in \mathbb{Z}}$ noté $a + n\mathbb{Z}$. Si $\gamma \in \mathbb{Z}/n\mathbb{Z}$, l'unique élément de $\gamma \cap \llbracket 0, n-1 \rrbracket$ s'appelle son **reste mod (n)** . De même si $a \in \mathbb{Z}$, l'unique $r \in \llbracket 0, n-1 \rrbracket$ tel que $a \equiv r \pmod{n}$ est appelé son reste mod (n) . On déduit immédiatement de la proposition IV.1.1 :

THÉORÈME IV.1.1

|| L'entier $n \geq 1$ étant donné, l'application

$$\gamma_n: \llbracket 0, n-1 \rrbracket \rightarrow \mathbb{Z}/n\mathbb{Z}, \quad r \mapsto r + n\mathbb{Z}$$

|| est bijective. En particulier, l'ensemble $\mathbb{Z}/n\mathbb{Z}$ est fini, de cardinal n .

Si $n = 1$, $\mathbb{Z}/n\mathbb{Z}$ est réduit à un singleton, ce qui présente peu d'intérêt.

On utilise souvent des **systèmes complets mod (n)** : par définition, un tel système est une partie E de \mathbb{Z} telle que chaque classe mod (n) rencontre E suivant un singleton, c'est-à-dire telle que l'application $E \rightarrow \mathbb{Z}/n\mathbb{Z}, x \mapsto x + n\mathbb{Z}$ soit bijective. Par exemple, pour tout $a \in \mathbb{Z}$, $\llbracket a, a+n-1 \rrbracket$ est un système complet, mais on en utilise parfois de plus subtils.

Montrons sur des exemples comment on calcule sur des congruences :

Exemple 1 : Soit à démontrer que $2^{32} + 1 \equiv 0 \pmod{641}$ (Euler) ⁽¹⁾.

On remarque : $641 = 128 \times 5 + 1 = 1 + 5 \times 2^7$ et aussi : $641 = 2^4 + 5^4$.
Donc $5 \times 2^7 \equiv -1 \pmod{641}$ et $5^4 \equiv -2^4 \pmod{641}$. (CG3) donne

$$(5 \times 2^7)^4 \equiv (-1)^4 \equiv 1 \pmod{641},$$

c'est-à-dire $5^4 \times 2^{28} \equiv 1 \pmod{641}$

(CG2) donne alors $-2^4 \times 2^{28} \equiv 1 \pmod{641}$

d'où enfin $2^{32} + 1 \equiv 0 \pmod{641}$.

Exemple 2 : Montrer que, pour tout $n \in \mathbb{N}$, $3^{2n+1} + 2^{n+2} \equiv 0 \pmod{7}$. Comme $3^6 \equiv 1 \pmod{7}$ et $2^3 \equiv 1 \pmod{7}$ il suffit de raisonner sur la classe de $n \pmod{3}$, ce qui ne laisse que trois cas à examiner, d'où le tableau (dont

⁽¹⁾ Euler (Leonhard), né à Bâle en 1707, mort à Saint-Petersbourg en 1783, est le plus grand mathématicien suisse. Il fut aussi physicien et ingénieur et laisse une œuvre

la dernière ligne répond à la question posée) :

reste de n	mod (3)	0	1	2
reste de 2^n	mod (7)	1	2	4
reste de 3^{2n}	mod (7)	1	2	4
reste de 2^{n+2}	mod (7)	4	1	2
reste de 3^{2n+1}	mod (7)	3	6	5
reste de $3^{2n+1} + 2^{n+2}$	mod (7)	0	0	0

Dans cet exemple 2 on aurait pu remarquer directement que :

$$3^{2n+1} + 2^{n+2} \equiv 3 \times (3^2)^n + 4 \times 2^n \equiv 3 \times 2^n + 4 \times 2^n \equiv 0 \pmod{7}.$$

Structure d'anneau de $\mathbb{Z}/n\mathbb{Z}$

Dans ce qui suit l'entier $n \geq 2$ est supposé fixé une fois pour toutes. Nous désignerons par $Y_n: \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ l'application canonique. La compatibilité de la congruence mod (n) avec l'addition et la multiplication de \mathbb{Z} signifient : si $x \in \mathbb{Z}$ et $y \in \mathbb{Z}$, les éléments $Y_n(x+y)$ et $Y_n(xy)$ ne dépendent que de $Y_n(x)$ et de $Y_n(y)$.

Puisque Y_n est surjective, cela permet donc de définir deux lois de composition, que nous noterons provisoirement \oplus et \otimes , sur l'ensemble $\mathbb{Z}/n\mathbb{Z}$ en posant : si $X \in \mathbb{Z}/n\mathbb{Z}$ et $Y \in \mathbb{Z}/n\mathbb{Z}$,

- (1) $X \oplus Y =$ l'élément S de $\mathbb{Z}/n\mathbb{Z}$
tel que $\forall x \in X, \forall y \in Y, S = Y_n(x+y)$
- (2) $X \otimes Y =$ l'élément P de $\mathbb{Z}/n\mathbb{Z}$
tel que $\forall x \in X, \forall y \in Y, P = Y_n(xy)$.

Ces deux lois de composition sur $\mathbb{Z}/n\mathbb{Z}$ sont appelées respectivement **addition** et **multiplication** et se notent simplement avec les signes $+$ et \times (le second étant d'ailleurs souvent omis), sans que cela puisse créer d'ambiguïté car le contexte indique toujours clairement dans quel ensemble (soit \mathbb{Z} , soit $\mathbb{Z}/n\mathbb{Z}$) on opère.

Si $x \in \mathbb{Z}$, pour abréger, nous écrirons \bar{x} au lieu de $Y_n(x)$ pour la classe de $x \pmod{n}$.

Il est immédiat de vérifier que l'addition et la multiplication que nous venons de définir munissent $\mathbb{Z}/n\mathbb{Z}$ d'une **structure d'anneau commutatif**.

A titre d'exemple, soit à prouver que l'addition est associative :

Si X, Y et Z sont pris dans $\mathbb{Z}/n\mathbb{Z}$, choisissons $x \in X, y \in Y, z \in Z$, c'est-à-dire x, y et z dans \mathbb{Z} tels que $\bar{x} = X, \bar{y} = Y, \bar{z} = Z$. Alors

$$(X + Y) + Z = \overline{(x+y)} + \bar{z} = \overline{(x+y) + z}$$

(d'après (1)) et $X + (Y + Z) = \bar{x} + \overline{(y+z)} = \overline{x + (y+z)}$ (toujours

Puisque l'addition dans \mathbb{Z} est associative $(x + y) + z = x + (y + z)$, ce qui entraîne bien $(x + z) + z = x + (y + z)$, d'où l'égalité de $(X + Y) + Z$ avec $X + (Y + Z)$. Les autres propriétés se démontrent de façon aussi élémentaire.

L'élément nul de l'anneau $\mathbb{Z}/n\mathbb{Z}$ ainsi défini est $\bar{0} = Y_n(0)$; l'élément unité est $\bar{1} = Y_n(1)$. On a $\bar{1} \neq \bar{0}$ car $n \geq 2$.

Les lois d'addition et de multiplication données par (1) et (2) sont appelées **lois quotient (de l'addition et de la multiplication de \mathbb{Z}) par l'idéal $n\mathbb{Z}$** .

Par définition même, on a, pour x et y dans \mathbb{Z} :

$$Y_n(x + y) = Y_n(x) + Y_n(y), \quad Y_n(xy) = Y_n(x) Y_n(y)$$

et aussi : $Y_n(1) = \bar{1} = 1_{\mathbb{Z}/n\mathbb{Z}}$. Autrement dit, Y_n est un **homomorphisme d'anneaux** de \mathbb{Z} dans $\mathbb{Z}/n\mathbb{Z}$. Son image est $\mathbb{Z}/n\mathbb{Z}$ car Y_n est surjectif, comme application canonique d'un ensemble dans un quotient de cet ensemble par une relation d'équivalence. Si $(x, y) \in \mathbb{Z}^2$, les définitions prouvent que

$$(Y_n(x) = Y_n(y)) \Leftrightarrow (x \equiv y \text{ mod } (n)),$$

$$\text{c'est-à-dire : } (Y_n(x) = Y_n(y)) \Leftrightarrow (x - y \in n\mathbb{Z}).$$

En particulier, $Y_n(x) = \bar{0} \Leftrightarrow x \in n\mathbb{Z}$, autrement dit le noyau de l'homomorphisme Y_n est l'idéal $n\mathbb{Z}$ de \mathbb{Z} .

Résumons les résultats obtenus :

THÉORÈME IV.1.2

Soit n un entier ≥ 2 . Muni des lois quotient de l'anneau \mathbb{Z} par l'idéal $n\mathbb{Z}$, l'ensemble $\mathbb{Z}/n\mathbb{Z}$ devient un **anneau commutatif** ; avec cette structure d'anneau, l'application canonique $Y_n : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ est un **homomorphisme d'anneaux**, surjectif, de noyau $n\mathbb{Z}$. Enfin l'anneau $\mathbb{Z}/n\mathbb{Z}$ est **fini**, de cardinal n .

Pour $n = 1$, ce théorème reste vrai, mais l'anneau obtenu \mathbb{Z}/\mathbb{Z} est l'anneau nul.

Pour $n \geq 2$, en revanche, l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'est pas nul et on a : $\bar{1} \neq \bar{0}$.

Désormais les ensembles $\mathbb{Z}/n\mathbb{Z}$ seront systématiquement munis de la structure d'anneau ainsi construite.

Exemple 3 : Tables de l'anneau $\mathbb{Z}/2\mathbb{Z}$ et de l'anneau $\mathbb{Z}/3\mathbb{Z}$

$\mathbb{Z}/2\mathbb{Z}$				$\mathbb{Z}/3\mathbb{Z}$			
addition		multiplication		addition		multiplication	
+	$\bar{0}$ $\bar{1}$	\times	$\bar{0}$ $\bar{1}$	+	$\bar{0}$ $\bar{1}$ $\bar{2}$	\times	$\bar{0}$ $\bar{1}$ $\bar{2}$
$\bar{0}$	$\bar{0}$ $\bar{1}$	$\bar{0}$	$\bar{0}$ $\bar{0}$	$\bar{0}$	$\bar{0}$ $\bar{1}$ $\bar{2}$	$\bar{0}$	$\bar{0}$ $\bar{0}$ $\bar{0}$
$\bar{1}$	$\bar{1}$ $\bar{0}$	$\bar{1}$	$\bar{0}$ $\bar{1}$	$\bar{1}$	$\bar{1}$ $\bar{2}$ $\bar{0}$	$\bar{1}$	$\bar{0}$ $\bar{1}$ $\bar{2}$
				$\bar{2}$	$\bar{2}$ $\bar{0}$ $\bar{1}$	$\bar{2}$	$\bar{0}$ $\bar{2}$ $\bar{1}$

On constate que ces deux anneaux sont des corps. Dans le second, en notant $-\bar{1}$ la classe $\bar{2}$, on constate que le groupe multiplicatif $(\mathbb{Z}/3\mathbb{Z} \setminus \{0\}, \times)$ est isomorphe au groupe additif $(\mathbb{Z}/2\mathbb{Z}, +)$. On reconnaît aussi la règle des signes.

Exemple 4 : Tables de l'anneau $\mathbb{Z}/4\mathbb{Z}$

addition					multiplication				
+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	\times	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{0}$

On constate que l'anneau $\mathbb{Z}/4\mathbb{Z}$ n'est pas intègre : $\bar{2} \times \bar{2} = \bar{0}$ alors que $\bar{2} \neq \bar{0}$. Si l'on voulait avoir un groupe multiplicatif, il ne faudrait garder que les éléments $\bar{1}$ et $\bar{3} = -\bar{1}$.

Exercice 1 : Dans l'anneau $\mathbb{Z}/6\mathbb{Z}$, résoudre les équations suivantes :

$$X^5 - \bar{1} = \bar{0}; \quad X^5 - \bar{2} = \bar{0}; \quad X^4 + X^3 + \bar{2} = \bar{0}.$$

Exercice 2 : Calculer le reste mod (41) de l'entier $(51\,200)^{2^{100}}$, le reste mod (31) de $(1\,986)^{10\,000}$ et le reste mod (17) de $(1\,035\,125)^{5\,642}$.

Exercice 3 : Vérifier que $10^6 \equiv 1 \pmod{7}$ et en déduire

$$\sum_{k=1}^{10} 10^{10^k} \equiv 5 \pmod{7}.$$

Exercice 4 : Vérifier que $10^3 \equiv -1 \pmod{7}$ et en déduire un critère de divisibilité par 7.

Exercice 5 : Pour quels entiers $n \in \mathbb{N}^*$ a-t-on :

- $3 \times 5^{2n+1} + 2^{3n+1} \equiv 0 \pmod{7}$
- $2^{10n-7} + 3^{5n-2} \equiv 2 \pmod{11}$
- $2^{2n} + 2^n + 1 \equiv 0 \pmod{21}$

- d) $n \times 7^{n+1} - (n+1)7^n - 1 \equiv 0 \pmod{17}$
 e) $n^{12} - n^8 - n^4 + 1 \equiv 0 \pmod{512}$?

Exercice 6 : Montrer que $(\forall n \in \mathbb{N}^*)$:

- a) $2^{2n}(2^{2n+1} - 1) \equiv 1 \pmod{9}$
 b) $9^{2n+1} + 8^{n+2} \equiv 0 \pmod{73}$
 c) $16^n - 15n - 1 \equiv 0 \pmod{225}$.

Exercice 7 : Résoudre les équations

- a) $X^2 + \overline{1} = \overline{0}$ dans $\mathbb{Z}/65\mathbb{Z}$ b) $X^2 + \overline{2} = \overline{0}$ dans $\mathbb{Z}/33\mathbb{Z}$.

Exercice 8 : Par quel nombre faut-il multiplier le nombre A qui s'écrit en base dix 12 345 679 pour que le produit soit formé de neuf chiffres égaux ? (on pourra comparer $10A$ et A . De manière analogue $37 \times 3 = 111, \dots$).

Exercice 9 : Démontrer : $(\forall n \in \mathbb{N}^*)$

- a) $10^{6n} + 10^{3n} - 2 \equiv 0 \pmod{111}$
 b) $7^{2n+1} - 48n - 7 \equiv 0 \pmod{288}$

Exercice 10 : Montrer que : $a^3 + b^3 + c^3 \equiv 0 \pmod{7} \Rightarrow abc \equiv 0 \pmod{7}$.

Exercice 11 : Démontrer : $(\forall n \in \mathbb{N})$

- a) $4^{2n} + 2^{2n} + 1 \equiv 0 \pmod{7}$ (Stifel) b) $2^{2n} + 15n - 1 \equiv 0 \pmod{9}$

Exercice 12 : Si $3^m + 1 \equiv 0 \pmod{10}$, alors $3^{m+4n} + 1 \equiv 0 \pmod{10}$ (m et n dans \mathbb{N} ainsi que $m - 4n$).

Exercice 13 : Démontrer : $10^{3n} - 1 \equiv 0 \pmod{3^{n+2}}$.

Exercice 14 : Si 9 divise $a^3 + b^3 + c^3$, alors 3 divise a ou b ou c .

Exercice 15 : Pour quelles valeurs de $n \in \mathbb{Z}$ l'entier $n^2 + (n+1)^2 + (n+3)^2$ est-il multiple de 10 ?

Exercice 16 : Montrer que 121 ne divise jamais $n^2 + 3n + 5$; que 95 ne divise jamais $4n^2 + 1$.

Exercice 17 : On désigne par $E(x)$ le plus grand entier inférieur ou égal au réel x . Montrer que : $(\forall n \in \mathbb{N}^*) \quad 2^{n+1} \mid E((1 + \sqrt{3})^{2n+1})$.

Exercice 18 : Quels sont les entiers $n \geq 1$ qui sont divisibles par tous les entiers $\leq \sqrt{n}$?

Exercice 19 : Etude des anneaux de Boole finis.

1) Soit E un ensemble fini non vide. Montrer que l'anneau $\mathcal{F}(E, \mathbb{Z}/2\mathbb{Z}) = \mathcal{B}_0(E)$ des fonctions de E dans $\mathbb{Z}/2\mathbb{Z}$ est en bijection naturelle avec l'ensemble des parties de E , et que l'addition et la multiplication de $\mathcal{B}_0(E)$ correspondent respectivement aux lois $(X, Y) \mapsto (X \cup Y) \setminus (X \cap Y)$ noté aussi $X \triangle Y$ et $(X, Y) \mapsto X \cap Y$ de $\mathcal{P}(E)$.

2) Soit B un sous-anneau de $B_0(E)$ tel que

$$(\forall x \in E, \forall y \in E) \quad (x \neq y) \Rightarrow (\exists f \in B \mid f(x) \neq f(y)).$$

Montrer que $B = B_0(E)$.

3) Si A est un anneau commutatif fini, montrer que tout idéal \mathfrak{a} de A autre que A est contenu dans un idéal maximal de A (cf. les exercices du § III.7).

4) Soit A un anneau tel que $(\forall x \in A) \quad x^2 = x$.

a) Montrer qu'un tel anneau est nécessairement commutatif et que

$$(\forall x \in A) \quad x + x = 0.$$

b) Soit \mathcal{A} l'ensemble des idéaux maximaux de A . Prouver que $\bigcap_{I \in \mathcal{A}} I = \{0\}$.

Indication : pour $x \in A \setminus \{0\}$, $1 - x$ n'est pas inversible. Considérer u contenant $1 - x$.

c) On définit $\varphi : A \rightarrow \mathcal{F}(\mathcal{M}, \mathbb{Z}/2\mathbb{Z})$ par $(\forall x \in A) \varphi(x) = \tilde{x}$, où $\tilde{x}(I) = \bar{0}$ si $x \in I$ et $\tilde{x}(I) = \bar{1}$ si $x \notin I$. Montrer que φ est un isomorphisme d'anneaux.

5) Trouver alors tous les sous-anneaux de $\mathcal{B}_0(E)$.

(En calcul des probabilités on dit qu'une famille de parties de E constitue un **clan** si elle contient E et si elle est stable pour la réunion et la complémentation. Cet exercice permet de construire tous les clans d'événements dans le cas où l'univers E est fini, et montre en particulier qu'un clan a pour cardinal une puissance de 2).

Exercice 20 : Pour $n \in \mathbb{N}$, le nombre $\sum_{k=0}^n 2^{3k} C_{2n+1}^{2k+1}$ n'est jamais divisible par 5.

Exercice 21 : Soit A la somme des chiffres de 4444^{4444} et B la somme des chiffres de A . Trouver la somme des chiffres de B ; la numération est la numération décimale. (*Olympiades, 1975.*)

§ IV.2 ARITHMÉTIQUE DANS \mathbb{Z} ET \mathbb{N}

Plus petit commun multiple (ppcm)

Donnons-nous $n \in \mathbb{N}^*$ et une suite finie (a_1, a_2, \dots, a_n) d'éléments de \mathbb{Z} . Les multiples communs aux nombres a_1, a_2, \dots, a_n sont évidemment les éléments de $a_1 \mathbb{Z} \cap a_2 \mathbb{Z} \cap \dots \cap a_n \mathbb{Z} = \bigcap_{i=1}^n a_i \mathbb{Z}$.

Cet ensemble, intersection d'idéaux de \mathbb{Z} , est donc un idéal de \mathbb{Z} (cf. proposition III.7.1). Il existe donc un unique $\mu \in \mathbb{N}$ tel que $\bigcap_{i=1}^n a_i \mathbb{Z} = \mu \mathbb{Z}$. l'entier μ possède donc, et c'est le seul entier naturel à la posséder, la propriété suivante :

(P₁) $\left\{ \begin{array}{l} \bullet \text{ Pour chaque } i, (i \in \llbracket 1, n \rrbracket), \mu \text{ est multiple de } a_i. \\ \bullet \text{ Tout entier } \lambda \text{ qui est multiple de chaque } a_i \text{ est multiple de } \mu. \end{array} \right.$

En d'autres termes, *l'ensemble des multiples communs aux a_i est l'ensemble des multiples de μ .*

DÉFINITION IV.2.1

$\left. \begin{array}{l} \} \text{ L'entier } \mu \text{ qui est le seul naturel à posséder la propriété (P}_1\text{) ci-} \\ \} \text{ dessus s'appelle le } \mathbf{\text{plus petit commun multiple}} \text{ des } a_i ; \text{ on le note} \\ \} \text{ ppcm } ((a_i)_{1 \leq i \leq n}). \end{array} \right\}$

Dans \mathbb{Z} , il y a exactement deux éléments (un seul si $\mu = 0$) qui vérifient (P₁). Ce sont μ et $-\mu$. On les appelle les ppcm des (a_i) dans \mathbb{Z} .

L'application $\mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ (à valeurs dans \mathbb{N}), $(x, y) \mapsto \text{ppcm}(x, y)$ (noté aussi $x \wedge y$) est une loi de composition sur \mathbb{Z} dont on vérifie immédiatement les propriétés suivantes : elle est **associative** et **commutative** (comme l'intersection des idéaux correspondants) ;

$$(\forall (a, b, c) \in \mathbb{Z}^3), \quad \text{ppcm}(ab, ac) = |a| \text{ppcm}(b, c),$$

De plus :

PROPOSITION IV.2.1

$$\left\| \begin{array}{l} \text{Si } a_1, a_2, \dots, a_n \in \mathbb{Z}, \text{ on a : } \text{ppcm}(a_1, a_2, \dots, a_n) = \bigwedge_{i=1}^n a_i. \quad ({}^1) \end{array} \right.$$

Démonstration (abrégée) :

Soit $\mu = \text{ppcm}(a_1, a_2, \dots, a_n)$, et $\mu' = \bigwedge_{i=1}^n a_i$;
alors μ' est multiple de chaque a_i , donc de μ ; réciproquement, par
récurrence sur k on voit que μ est multiple de $\bigwedge_{i=1}^k a_i$ pour $1 \leq k \leq n$, donc μ
est multiple de μ' , d'où finalement $\mu = \mu'$. ■

On notera que si l'un des a_i est nul, alors $\text{ppcm}(a_1, a_2, \dots, a_n) = 0$, ce qui fait apparaître 0 comme multiple de tout entier, ce qu'il est en effet.

Plus grand commun diviseur (pgcd)

Considérons à nouveau $n \in \mathbb{N}^*$ et une suite finie (a_1, a_2, \dots, a_n) d'éléments de \mathbb{Z} . Il existe un unique $d \in \mathbb{N}$ tel que l'idéal

$$a_1 \mathbb{Z} + a_2 \mathbb{Z} + \dots + a_n \mathbb{Z} = \sum_{i=1}^n a_i \mathbb{Z}$$

soit égal à $d\mathbb{Z}$. On sait de plus que pour $(a, b) \in \mathbb{Z}^2$, l'inclusion $a\mathbb{Z} \subset b\mathbb{Z}$ équivaut à la relation : b divise a (cf. début du § IV.1). Puisque l'idéal $\sum_{i=1}^n a_i \mathbb{Z} = d\mathbb{Z}$ est, au sens de l'inclusion, le plus petit idéal α de \mathbb{Z} contenant chaque a_i , on voit que l'entier d est le seul entier naturel à posséder la propriété suivante :

$$(P_2) \left\{ \begin{array}{l} \bullet \text{ Pour chaque } i, (i \in \llbracket 1, n \rrbracket), d \text{ divise } a_i. \\ \bullet \text{ Tout entier } k \in \mathbb{Z} \text{ qui divise chaque } a_i \text{ est un diviseur de } d. \end{array} \right.$$

En d'autres termes, l'ensemble des diviseurs communs aux a_i est l'ensemble des diviseurs de d .

(¹) La notation $\bigwedge_{i=1}^n a_i$ représente ici le *composé itéré* des a_i pour la loi de composition

\wedge défini 6 lignes plus haut. (Voir Définition III.1.1)

DÉFINITION IV.2.2

$\left\{ \begin{array}{l} \text{L'entier } d \text{ qui est le seul naturel à posséder la propriété } (P_2) \text{ ci-dessus} \\ \text{s'appelle le plus grand commun diviseur des } a_i ; \text{ on le note} \\ \text{pgcd } ((a_i)_{1 \leq i \leq n}). \end{array} \right.$

Les entiers d et $-d$ sont les seuls éléments de \mathbb{Z} vérifiant (P_2) ; on les appelle les pgcd dans \mathbb{Z} des (a_i) .

L'application $\mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ (à valeurs dans \mathbb{Z}), $(x, y) \mapsto \text{pgcd}(x, y)$ (noté aussi $x \vee y$) est une loi de composition sur \mathbb{Z} dont on vérifie immédiatement les propriétés suivantes : elle est **associative** et **commutative** (comme la somme des idéaux correspondants) et :

$$(\forall (a, b, c) \in \mathbb{Z}^3) \quad \text{pgcd}(ab, ac) = |a| \text{pgcd}(b, c).$$

De plus :

PROPOSITION IV.2.2

$$\left\| \begin{array}{l} \text{Si } a_1, a_2, \dots, a_n \in \mathbb{Z}, \text{ on a : } \text{pgcd}(a_1, a_2, \dots, a_n) = \bigvee_{i=1}^n a_i. \end{array} \right. \quad (1)$$

Démonstration :

Soit $d = \text{pgcd}((a_i)_{1 \leq i \leq n})$, $d' = \bigvee_{i=1}^n a_i$; alors d' divise chaque a_i , donc divise d ; réciproquement, par récurrence sur k , on voit que d divise $\bigvee_{k=1}^k a_i$ ($k \in \llbracket 1, n \rrbracket$). Donc d divise d' , et enfin $d = d'$. ■

Il est utile, pour la suite, de se rappeler que $d = \text{pgcd}(a_1, a_2, \dots, a_n)$ est un élément de l'idéal somme $a_1 \mathbb{Z} + a_2 \mathbb{Z} + \dots + a_n \mathbb{Z}$, c'est-à-dire :

$$\exists (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{Z}^n \mid d = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n.$$

On notera que si l'un des a_i est tel que $|a_i| = 1$, alors $\text{pgcd}(a_1, a_2, \dots, a_n) = 1$.

Algorithme d'Euclide

La proposition IV.2.2 montre que, dès qu'on sait calculer le pgcd de deux éléments quelconques de \mathbb{Z} , on peut facilement connaître le pgcd de (a_1, a_2, \dots, a_n) en appliquant $n - 1$ fois la technique de calcul pour $\text{pgcd}(a, b)$. Supposons donc $a \in \mathbb{Z}$ et $b \in \mathbb{Z}$ et cherchons leur pgcd. On remarque d'abord que $\text{pgcd}(a, b) = \text{pgcd}(|a|, |b|)$, ce qui permet de limiter la recherche à $\text{pgcd}(a, b)$ avec a et b dans \mathbb{N} . Si $b = 0$, il est clair que $\text{pgcd}(a, b) = a$.

Si $b \neq 0$, $\text{pgcd}(a, b) = \text{pgcd}(b, r)$ où r est le reste de la division euclidienne de a par b . Même si $a < b$ cette première étape est intéressante puisqu'elle a permis de placer le nombre le plus petit en seconde position dans le couple. En posant $b = r_0$ nous écrirons dans tous les cas $a = r_0 q_1 + r_1$ avec $r_1 \in \llbracket 0, r_0 - 1 \rrbracket$. Ici deux

(1) $\bigvee_{i=1}^n a_i$ représente le composé itéré des a_i pour la loi \vee introduite 6 lig

cas se présentent : ou bien $r_1 = 0$, et alors $a = bq_1$, et dans ce cas le pgcd de (a, b) est évidemment b (ce cas ne se présente que si $a \geq b$) ; ou bien $r_1 \neq 0$, et alors on peut effectuer la division euclidienne de r_0 par r_1 : $r_0 = r_1 q_2 + r_2$ et l'on a

$$\text{pgcd}(a, b) = \text{pgcd}(r_0, r_1) = \text{pgcd}(r_1, r_2)$$

où $r_2 \in \llbracket 0, r_1 - 1 \rrbracket$. Supposons ainsi définis les couples (q_i, r_i) pour $i \in \llbracket 1, k \rrbracket$. Si l'on tombe sur un reste r_k nul on arrête les divisions successives. Si $r_k \neq 0$ on pose $r_{k-1} = r_k q_{k+1} + r_{k+1}$ avec $r_{k+1} \in \llbracket 0, r_k - 1 \rrbracket$. On définit ainsi par récurrence deux suites $(r_i)_{i \in \mathbb{N}}$, $(q_i)_{i \in \mathbb{N}}$, étant entendu que si $r_k = 0$, on pose $r_i = 0$ pour $i \geq k$. La suite $(r_i)_{i \in \mathbb{N}}$ est *décroissante*. Elle ne peut décroître *strictement* indéfiniment (\mathbb{N} est bien ordonné), donc l'ensemble \mathcal{E} des $i \in \mathbb{N}$ tels que $r_i = 0$ est non vide. Si k_0 est le *minimum* de \mathcal{E} , on a par construction :

$$(3) \quad a = r_0 q_1 + r_1, \quad r_0 = r_1 q_2 + r_2, \dots, r_{k-2} = r_{k_0-1} q_{k_0}$$

avec

$$r_{k_0-1} \geq 1 \quad \text{et} \quad r_0 > r_1 > r_2 > \dots > r_{k_0-1} > 0.$$

La suite (3) est par définition la liste des **divisions successives** (de a par $b = r_0$).

D'après la remarque ci-dessus, les ensembles des diviseurs communs aux entiers des couples

$$(a, r_0), (r_0, r_1), \dots, (r_{k_0-2}, r_{k_0-1})$$

sont tous égaux. Le dernier d'entre eux est l'ensemble des diviseurs de r_{k_0-1} puisque $r_{k_0-1} \mid r_{k_0-2}$.

Donc le pgcd de a et b est r_{k_0-1} : c'est le **dernier reste non nul** de la suite des divisions successives. L'algorithme des divisions successives (3) est appelé **algorithme d'Euclide**. Il se prête parfaitement au calcul informatisé. (On remarque qu'on peut augmenter la rapidité de l'algorithme en remplaçant, à l'étape $r_{i-1} = r_i q_{i+1} + r_{i+1}$, le nombre r_{i+1} par $\text{Min}(r_{i+1}, r_i - r_{i+1})$).

Entiers premiers entre eux

DÉFINITION IV.2.3

Des entiers a_1, a_2, \dots, a_n (éléments de \mathbb{Z}) sont dits **premiers entre eux** ssi leur pgcd est égal à 1, ce qui signifie : les seuls diviseurs communs aux a_i sont 1 et -1 .

THÉORÈME IV.2.1 (Bachet ⁽¹⁾, Bezout ⁽²⁾)

Pour que les entiers a_1, a_2, \dots, a_n de \mathbb{Z} soient premiers entre eux, il faut et il suffit qu'il existe $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{Z}^n$ tel que $\sum_{i=1}^n \lambda_i a_i = 1$.

⁽¹⁾ Claude Gaspar Bachet de Méziriac (1581-1638) est l'auteur de *Problèmes plaisans et délectables* (éditions en 1612 et 1624) où il résout pour la première fois les équations indéterminées du 1^{er} degré.

⁽²⁾ Etienne Bezout (1730-1783) a étendu les résultats connus dans \mathbb{Z} à l'anneau $\mathbb{R}[X]$.

Démonstration :

Nous avons déjà vu après la proposition IV.2.2 que cette condition est nécessaire. Réciproquement si on a trouvé

$$(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{Z}^n \text{ tel que } \sum_{i=1}^n \lambda_i a_i = 1,$$

il est clair que tout diviseur commun aux a_i divise aussi $\sum_{i=1}^n \lambda_i a_i = 1$, donc les seuls diviseurs communs possibles aux a_i sont 1 et -1 . ■

Toute relation entre les a_i du type $\sum_{i=1}^n \lambda_i a_i = 1$ s'appelle une **relation de Bezout** entre les a_i : il en existe ssi les a_i sont premiers entre eux, et s'il en est ainsi, il y a *une infinité* de telles relations (cf. exercice n° 15).

Remarque 1 : Pour $n \geq 3$ il ne faut pas confondre les propriétés « a_1, a_2, \dots, a_n sont premiers entre eux » et « les $(a_i)_{1 \leq i \leq n}$ sont premiers entre eux deux à deux ». Si la seconde entraîne la première, la première n'entraîne pas la deuxième. On peut éviter la confusion en disant dans le premier cas que « les $(a_i)_{1 \leq i \leq n}$ sont premiers entre eux *dans leur ensemble* (ou globalement) ».

Remarque 2 : Si les éléments a et b de \mathbb{Z} sont premiers entre eux, la méthode des divisions successives permet de trouver une relation de Bezout entre a et b .

Le théorème IV.2.1 est fondamental en Arithmétique. Nous en développons ci-après les principales conséquences :

THÉOREME IV.2.2 (Théorème de Gauss ⁽¹⁾)

|| Soit a, b, c dans \mathbb{Z} . Si a divise bc et si a est premier avec b , alors a divise c .

Démonstration :

Soit une relation de Bézout : $\lambda a + \mu b = 1$; d'autre part soit $q \in \mathbb{Z}$ tel que $bc = aq$. Alors :

$$c = \lambda ac + \mu bc ; \quad c = \lambda ac + \mu aq = a (\lambda c + \mu q)$$

d'où $a|c$. ■

Autre démonstration :

$\text{pgcd}(a, b) = 1 \Rightarrow \text{pgcd}(ac, bc) = |c|$. Comme a divise ac et bc , a divise leur pgcd, d'où $a|c$.

⁽¹⁾ Carl Friedrich Gauss (1777-1855), physicien et astronome allemand. célèbre par sa précocité et son génie qui l'avait fait surnommer le prince des mathématiciens

THÉORÈME IV.2.3

|| Soit a, b_1, b_2, \dots, b_n dans \mathbb{Z} avec $n \geq 2$; si a est premier avec chaque b_i ($1 \leq i \leq n$), alors a est premier avec le produit $B = b_1 b_2 \dots b_n$.

Démonstration :

Par récurrence sur n ; si $n = 2$, soit

$$\lambda_i a + \mu_i b_i = 1 \quad (i = 1, 2)$$

des relations de Bezout. En les multipliant membre à membre on obtient $ua + vb_1 b_2 = 1$ avec

$$u = \lambda_1 \lambda_2 a + \lambda_1 \mu_2 b_2 + \lambda_2 \mu_1 b_1 \quad \text{et} \quad v = \mu_1 \mu_2.$$

Donc $\text{pgcd}(a, b_1 b_2) = 1$.

Supposons la propriété vraie à l'ordre $n \geq 2$, montrons-la à l'ordre $n + 1$: d'après l'hypothèse de récurrence, a et $b_1 b_2 \dots b_n$ sont premiers entre eux et, puisque a et b_{n+1} le sont aussi, l'étude du cas $n = 2$ prouve que a et $b_1 b_2 \dots b_{n+1}$ sont premiers entre eux. ■

THÉORÈME IV.2.4

|| Si les éléments a_i de \mathbb{Z} ($1 \leq i \leq n, n \geq 2$) sont deux à deux premiers entre eux, alors $\text{ppcm}(a_1, a_2, \dots, a_n) = \left| \prod_{i=1}^n a_i \right|$.

Démonstration :

Si $n = 2$, soit $\mu = \text{ppcm}(a_1, a_2)$ et q_1, q_2 les éléments de \mathbb{Z} tels que $\mu = a_1 q_1 = a_2 q_2$. D'après le théorème de Gauss $a_1 | q_2$, d'où $q_2 = a_1 u$ avec $u \in \mathbb{Z}$, et par suite $\mu = a_1 a_2 u$ est multiple de $a_1 a_2$. Mais $a_1 a_2$ est multiple commun à a_1, a_2 , donc multiple de μ , d'où $\text{ppcm}(a_1, a_2) = |a_1 a_2|$. Supposons la propriété vraie à l'ordre $n \geq 2$, prouvons-la à l'ordre $n + 1$: d'après le théorème IV.2.3, a_{n+1} est premier avec $a_1 a_2 \dots a_n$. Donc

$$\text{ppcm}(a_1 a_2 \dots a_n, a_{n+1}) = |a_1 a_2 \dots a_n| |a_{n+1}| = \left| \prod_{i=1}^{n+1} a_i \right|.$$

Par l'hypothèse de récurrence $\text{ppcm}(a_1, a_2, \dots, a_n) = |a_1 a_2 \dots a_n|$. D'après l'associativité du ppcm, il s'ensuit bien :

$$\text{ppcm}(a_1, a_2, \dots, a_n, a_{n+1}) = \left| \prod_{i=1}^{n+1} a_i \right|.$$

Le théorème est donc prouvé par récurrence sur n . ■

COROLLAIRE

|| Si des $a_i \in \mathbb{Z}$ ($1 \leq i \leq n$, $n \geq 2$) sont deux à deux premiers entre eux et si chaque a_i divise l'entier $\mu \in \mathbb{Z}$, alors $a_1 a_2 \dots a_n$ divise μ .

THÉORÈME IV.2.5

|| Soit $d \in \mathbb{N}^*$ un diviseur commun aux entiers a_1, a_2, \dots, a_n ($a_i \in \mathbb{Z}$).
 || Pour que d soit le pgcd de (a_1, a_2, \dots, a_n) , il faut et il suffit que les entiers $(a_i/d)_{1 \leq i \leq n}$ soient premiers entre eux.

Démonstration :

La condition est suffisante : en effet, soit une relation de Bezout

$$\lambda_1 a_1/d + \lambda_2 a_2/d + \dots + \lambda_n a_n/d = 1 \quad (\lambda_i \in \mathbb{Z}).$$

Alors $\sum_{i=1}^n \lambda_i a_i = d$ et $d \in a_1 \mathbb{Z} + a_2 \mathbb{Z} + \dots + a_n \mathbb{Z}$, donc d est multiple de pgcd (a_1, a_2, \dots, a_n) et comme d divise chaque a_i , d divise pgcd (a_1, \dots, a_n) d'où $d = \text{pgcd}(a_1, \dots, a_n)$.

La condition est nécessaire : car si $d = \text{pgcd}(a_1, \dots, a_n)$, on a

$$(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{Z}^n \quad \text{tel que} \quad d = \sum_{i=1}^n \lambda_i a_i ;$$

d'où $\sum_{i=1}^n \lambda_i \frac{a_i}{d} = 1$, donc les $\left(\frac{a_i}{d}\right)$ sont premiers entre eux. ■

Calcul du ppcm

Soit a et b dans \mathbb{Z}^* ; notons $\mu = \text{ppcm}(a, b)$, $d = \text{pgcd}(a, b)$. Alors $a = da'$, $b = db'$ avec $\text{pgcd}(a', b') = 1$, d'où $\text{ppcm}(a', b') = a' b'$ (d'après le théorème IV.2.4).

On en déduit $\text{ppcm}(a, b) = d \text{ppcm}(a', b') = da' b'$, c'est-à-dire

(4)

$\text{ppcm}(a, b) \times \text{pgcd}(a, b) = |ab|$

Compte tenu de ce qui précède, le calcul de $\text{ppcm}(a_1, a_2, \dots, a_n)$ est, tout comme le calcul de $\text{pgcd}(a_1, a_2, \dots, a_n)$, ramené de proche en proche à celui du pgcd de deux entiers.

Forme irréductible d'un rationnel

THÉORÈME IV.2.6

Soit $r \in \mathbb{Q}^*$; il existe un et un seul couple $(a, b) \in \mathbb{Z} \times \mathbb{N}^*$ tel que $r = \frac{a}{b}$ avec $\text{pgcd}(a, b) = 1$. Notant (a, b) ce couple, l'ensemble des couples $(c, d) \in \mathbb{Z} \times \mathbb{Z}^*$ tels que $\frac{c}{d} = r$ est l'ensemble des couples $(\rho a, \rho b)_{\rho \in \mathbb{Z}^*}$.

Démonstration :

Existence : Soit $(c_0, d_0) \in \mathbb{Z} \times \mathbb{Z}^*$ tel que $r = \frac{c_0}{d_0}$. Notons δ le pgcd de c_0 et d_0 , d'où $c_0 = \delta c'_0$, $d_0 = \delta d'_0$ avec $\text{pgcd}(c'_0, d'_0) = 1$. Alors si $a = \varepsilon c'_0$, $b = \varepsilon d'_0$ (où $\varepsilon = \frac{d'_0}{|d'_0|}$), on a : $\text{pgcd}(a, b) = 1$ et $r = \frac{a}{b}$ répond à la question.

Unicité. De $r = \frac{a}{b}$, on déduit évidemment $r = \frac{\rho a}{\rho b}$ pour tout $\rho \in \mathbb{Z}^*$. Réciproquement, si $r = \frac{c}{d}$, $c \in \mathbb{Z}$ et $d \in \mathbb{Z}^*$, alors $bc = da$. Le théorème de Gauss montre que $b|d$, d'où $d = \rho b$, d'où $bc = \rho ab$, d'où $c = \rho a$ avec $\rho \in \mathbb{Z}^*$. ■

Par définition le couple (a, b) défini dans l'énoncé du théorème IV.2.6 s'appelle le **représentant irréductible** du rationnel r .

Exercice 1 : Montrer que pour $n \in \mathbb{N}$, la fraction $\frac{n^3 + n}{2n^2 + 1}$ est irréductible.

Exercice 2 : Soit a, b, c dans \mathbb{N} avec b et c premiers entre eux. Démontrer : $(a^b - 1)(a^c - 1)$ divise $(a - 1)(a^{bc} - 1)$.

Exercice 3 : Soit r_1, r_2, \dots, r_n dans \mathbb{Q}_+^* ;

a) l'ensemble $r_1 \mathbb{Z} + r_2 \mathbb{Z} + \dots + r_n \mathbb{Z} = \{\lambda_1 r_1 + \dots + \lambda_n r_n\}_{(\lambda_i) \in \mathbb{Z}^n}$ est un sous-groupe additif de \mathbb{Q} , que nous notons G .

b) Montrer qu'il existe $g \in \mathbb{Q}_+^*$ tel que $G = g\mathbb{Z}$ et expliquer comment on peut trouver g connaissant les représentants irréductibles des r_i .

Application numérique : $n = 3$, $r_1 = \frac{24}{13}$, $r_2 = \frac{25}{126}$, $r_3 = \frac{147}{14}$.

Exercice 4 : Trouver deux entiers connaissant leur produit (resp. leur quotient exact) et leur ppcm.

Exercice 5 : Trouver l'ensemble $\{(m, n) \in \mathbb{N}^2 \mid 2^m - 3^n = 1\}$.

Exercice 6 : Prouver que pour tout $n \in \mathbb{Z}$, la fraction $\frac{15n^2 + 8n + 6}{30n^2 + 21n + 13}$ est irréductible.

Exercice 7 : Trouver deux entiers a et b tels que $a^2 + b^2 = 85\,113$ et ppcm

Exercice 8 : Si a et b sont premiers entre eux, trouver $\text{pgcd}(a^3 - b^3, (a - b)^3)$.

Exercice 9 : Si un entier N est à la fois la puissance m -ième exacte d'un entier et la puissance n -ième exacte d'un entier avec m et n premiers entre eux, alors N est une puissance (mn) -ième exacte.

Exercice 10 : Soit $n \geq 2$ et $(a_i, b_i) \in \mathbb{N}^* \times \mathbb{N}^*$ tel que $a_i \vee b_i = 1$ pour tout $i \in \llbracket 1, n \rrbracket$. On suppose les b_i premiers entre eux deux à deux sans être tous égaux à 1. Prouver que $\frac{a_1}{b_1} + \frac{a_2}{b_2} + \dots + \frac{a_n}{b_n} \notin \mathbb{N}$.

Exercice 11 : Soit $P \in \mathbb{N}^*$ multiple commun à a_1, a_2, \dots, a_n ($a_i \in \mathbb{N}^*, n \geq 2$). On pose $P_i = \frac{P}{a_i}$ et $D = \text{pgcd}(P_1, P_2, \dots, P_n)$. Montrer : $\text{ppcm}(a_1, a_2, \dots, a_n) = \frac{P}{D}$.

Exercice 12 : Soit $a_1, a_2, \dots, a_n \in \mathbb{N}^*$ ($n \geq 2$) ; $\mu = \text{ppcm}(a_1, a_2, \dots, a_n)$; $M = a_1 a_2 \dots a_n$; $M_i = \frac{M}{a_i}$ ($1 \leq i \leq n$) ; $\Delta = \text{pgcd}(M_1, M_2, \dots, M_n)$.

a) Prouver : $M = \mu \Delta$.

b) Soit $R = \text{ppcm}(M_1, \dots, M_n)$ et $D = \text{pgcd}(a_1, \dots, a_n)$. Prouver : $M = RD$.

Exercice 13 (théorème chinois).

Soit $m_1, m_2, \dots, m_n \in \mathbb{N}^*$ ($n \geq 2$), deux à deux premiers entre eux. On pose $M = m_1 m_2 \dots m_n$ et $M_i = \frac{M}{m_i}$ pour tout $i \in \llbracket 1, n \rrbracket$.

a) Pour tout i , $M_i \vee m_i = 1$. En déduire l'existence de $\mu_i \in \mathbb{Z}$ tel que $M_i \mu_i \equiv 1 \pmod{(m_i)}$.

b) Soit $a_1, a_2, \dots, a_n \in \mathbb{Z}$. On pose $x = \sum_{i=1}^n a_i M_i \mu_i$. Montrer :

$$(1) \quad (\forall i \in \llbracket 1, n \rrbracket) \quad x \equiv a_i \pmod{(m_i)}.$$

Cela prouve que le système de congruences (1) admet une solution.

c) Prouver : $\{y \in \mathbb{Z} \mid \forall i, y \equiv a_i \pmod{(m_i)}\}$ est la classe de congruence de $x \pmod{(M)}$.

Exercice 14 : Soit a, b, n des entiers avec $n \geq 3$. On étudie l'équation (1) $\bar{a}U = \bar{b}$ à l'inconnue $U \in \mathbb{Z}/n\mathbb{Z}$.

a) Si $a \vee n = 1$ montrer que l'équation (1) admet une solution unique.

b) On suppose $a \vee n = d \geq 2$. Discuter l'équation (1) si $d = n$. Puis, si $d \in \llbracket 2, n-1 \rrbracket$ montrer que (1) n'a de solution que si $d \mid b$.

c) On suppose $d \mid b$ et $d \in \llbracket 2, n-1 \rrbracket$. Soit $\alpha = \frac{a}{d}$, $\beta = \frac{b}{d}$, $\nu = \frac{n}{d}$. Soit X_0 l'unique solution dans $\mathbb{Z}/\nu\mathbb{Z}$ de l'équation (2) $\bar{\alpha}X = \bar{\beta}$. Soit Y_0 l'ensemble des classes de congruence mod (n) des éléments de X_0 . Démontrer que Y_0 est l'ensemble des solutions de (1) et que $\text{card}(Y_0) = \alpha$.

Application : Résoudre (1) pour $a = 36$, $b = 54$ et $n = 42$.

Exercice 15 : Soit a et b premiers entre eux dans \mathbb{Z} ($b \neq 0$).

a) Trouver l'ensemble des relations de Bezout $\lambda a + \mu b = 1$ ($(\lambda, \mu) \in \mathbb{Z}^2$) si on en connaît une : $\lambda_0 a + \mu_0 b = 1$. Prouver qu'il en existe une et une seule : $\alpha a + \beta b = 1$ telle que $|\alpha| < |b|$ et $|\beta| < |a|$.

b) Montrer comment l'algorithme des divisions successives de a par b permet justement de trouver (α, β) .

Exercice 16 :

a) Soit m et n deux naturels ≥ 2 . On pose $d = m \vee n$, $\mu = \frac{m}{d}$, $\nu = \frac{n}{d}$. On suppose $d \in \llbracket 2, m-1 \rrbracket \cap \llbracket 2, n-1 \rrbracket$; soit $M = \text{ppcm}(m, n) = m \wedge n$. On note F l'application : $\mathbb{Z} \rightarrow \llbracket 0, m-1 \rrbracket \times \llbracket 0, n-1 \rrbracket$, $x \mapsto (\text{reste de } x \pmod{(m)}, \text{reste de } x \pmod{(n)})$ et \mathcal{F} l'image de F . Montrer que $\Phi : \llbracket 0, M-1 \rrbracket \rightarrow \mathcal{F}$, $x \mapsto F(x)$ est bijective, et que si $(r, s) \in \mathcal{F}$ on a : $r - s \equiv 0 \pmod{(d)}$.

b) Si $\alpha \in \llbracket 0, d-1 \rrbracket$, soit

$$\mathcal{G}_\alpha = \{ (r, s) \in \llbracket 0, m-1 \rrbracket \times \llbracket 0, n-1 \rrbracket \mid r-s \equiv \alpha \pmod{d} \}.$$

Montrer que les \mathcal{G}_α sont tous non vides et de même cardinal, qui vaut M , et que $\mathcal{F} = \mathcal{G}_0$.

c) Dédurre de cette étude une discussion complète du système de congruences simultanées $x \equiv a \pmod{m}$, $x \equiv b \pmod{n}$ à l'inconnue x .

Exercice 17 : Pour $a \in \mathbb{N}^*$, trouver le reste de la division de $(a^2 + (a-1)^2)^2$ par $4a^2$. Réponse : $(2a-1)^2$.

Exercice 18 : Soit a et b deux entiers premiers entre eux ≥ 2 . On pose $S = \{ax + by \mid x \in \mathbb{N} \text{ et } y \in \mathbb{N}\}$. Montrer qu'il existe un entier m_0 , le plus petit possible, tel que $(\forall m \in \mathbb{N}) (m \geq m_0) \Rightarrow (m \in S)$. Quelle est la valeur de m_0 ?

Exercice 19 : On définit les entiers a_n et b_n par $(1 + \sqrt{2})^n = a_n + b_n \sqrt{2}$ ($n \in \mathbb{Z}$). Montrer que a_n et b_n sont premiers entre eux.

Exercice 20 : On pose $F_n = 2^{2^n} + 1$ ($n \in \mathbb{N}$). Montrer que si $m \neq n$, alors $F_m \vee F_n = 1$. Prouver aussi que F_n divise $2^{F_n} - 2$.

Exercice 21 : Résoudre, dans \mathbb{N}^* , $x \wedge y - x \vee y = 243$; $x \wedge y - 3x \vee y = 135$.

Exercice 22 : Montrer : $\text{ppcm}(1, 2, 3, \dots, 2n) = \text{ppcm}(n+1, n+2, \dots, 2n)$.

Exercice 23 : Soit $A \subset \mathbb{N}$ et deux naturels p et q premiers entre eux. Montrer que les 3 conditions :

$$(1) \quad 0 \in A \quad (2) \quad (n \in A) \Rightarrow (n+p \in A) \quad (3) \quad (n \geq q, n \in A) \Rightarrow (n-q \in A)$$

impliquent que $A = \mathbb{N}$.

Exercice 24 : Démontrer que tout sous-anneau du corps \mathbb{Q} est un anneau principal.

§ IV.3 ÉLÉMENTS INVERSIBLES DES ANNEAUX $\mathbb{Z}/n\mathbb{Z}$

Dans ce qui suit, nous fixons l'entier $n \geq 2$; l'homomorphisme d'anneaux canonique $Y_n : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ sera noté $x \mapsto \bar{x}$.

THÉORÈME IV.3.1

Si $\bar{a} \in \mathbb{Z}$, il y a équivalence entre les propriétés suivantes :

- (I) \bar{a} est un élément inversible de l'anneau $\mathbb{Z}/n\mathbb{Z}$
- (II) \bar{a} est un élément régulier (pour la multiplication) de $\mathbb{Z}/n\mathbb{Z}$
- (III) a est premier avec n , autrement dit $\text{pgcd}(a, n) = 1$.

Démonstration :

Dire que \bar{a} est inversible dans l'anneau $\mathbb{Z}/n\mathbb{Z}$ signifie : il existe $b \in \mathbb{Z}$ tel que $\bar{a}\bar{b} = \bar{1}$, c'est-à-dire : $ab - 1 \in n\mathbb{Z}$.

En d'autres termes, (I) équivaut à : il existe b et c dans \mathbb{Z} tels que $ab + cn = 1$, ce qui prouve déjà l'équivalence de (I) et (III). Nous savons que dans tout anneau commutatif un élément inversible est toujours régulier. Donc (I) \Rightarrow (II). Il reste à savoir ce qui se passe si \bar{a} n'est pas inversible dans $\mathbb{Z}/n\mathbb{Z}$. Dans ce cas a et n ne peuvent pas être premiers entre eux. Soit alors $d = \text{pgcd}(a, n)$, $a = da'$, $n = dn'$. Comme $d \geq 2$, $n' \in \llbracket 1, n-1 \rrbracket$ et en particulier $\bar{n}' \neq \bar{0}$. Alors $an' = a' dn' = a' n$, d'où $\bar{a}\bar{n}' = \bar{0}$, et puisque $\bar{n}' \neq \bar{0}$, cela prouve que \bar{a} n'est pas régulier, donc (II) \Rightarrow (I). ■

Ainsi dans l'anneau $\mathbb{Z}/n\mathbb{Z}$, il y a égalité entre l'ensemble \mathcal{U}_n des éléments inversibles et l'ensemble des éléments réguliers. En général, dans un anneau commutatif ces deux ensembles ne sont pas égaux ; on peut seulement dire que le premier est inclus dans le second. Le fait qu'ici tout élément inversible \bar{a} de $\mathbb{Z}/n\mathbb{Z}$ soit régulier est intéressant car cela complète nos règles de calcul sur les congruences (de CG1 à CG6 du § IV.1) par la règle

$$(\text{CG7}) : \text{si } a \nmid n, \text{ alors } b \equiv c \pmod{n} \Leftrightarrow ab \equiv ac \pmod{n}$$

qui signifie qu'on peut diviser une congruence mod (n) par a , à condition que a soit premier avec n . En fait, si $x \equiv y \pmod{n}$ est donné, on ne divise pas x et y par a , mais on les multiplie par a' , élément de $(\bar{a})^{-1}$, l'existence de a' découlant du théorème de Bezout.

Indicateur d'Euler

Le groupe des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$ sera dorénavant noté $G(n)$ au lieu de \mathcal{U}_n , son cardinal étant noté $\varphi(n)$. L'entier $\varphi(n)$ s'appelle l'**indicateur d'Euler de n** . En combinant les théorèmes IV.1.1 et IV.3.1 on obtient :

THÉORÈME IV.3.2

|| Soit \mathcal{E}_n l'ensemble des entiers $k \in \llbracket 0, n-1 \rrbracket$ qui sont **premiers avec n** ; l'application $k \mapsto \bar{k}$ définit une bijection de \mathcal{E}_n sur le groupe $G(n)$ des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$. En particulier $\varphi(n) = \text{card}(\mathcal{E}_n)$.

Par convention on pose $\varphi(1) = 1$ car dans ce cas $G(1)$ est un singleton.

Générateurs du groupe additif $\mathbb{Z}/n\mathbb{Z}$

THÉORÈME IV.3.3

|| Les éléments **inversibles** de l'anneau $\mathbb{Z}/n\mathbb{Z}$ sont les $\alpha \in \mathbb{Z}/n\mathbb{Z}$ qui **engendrent le groupe additif $\mathbb{Z}/n\mathbb{Z}$** , c'est-à-dire tels que $\mathbb{Z}/n\mathbb{Z} = \{k\alpha\}_{k \in \mathbb{Z}}$.

Démonstration :

Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$, $\alpha = \bar{a}$, où $a \in \mathbb{Z}$. Dire que α engendre le groupe additif $\mathbb{Z}/n\mathbb{Z}$ signifie que $\bar{1}$ appartient au groi

engendré par α , puisque $\bar{1}$ engendre évidemment le groupe additif (cyclique) $\mathbb{Z}/n\mathbb{Z}$.

Or $\bar{1} \in G$ équivaut à : il existe $k \in \mathbb{Z}$ tel que $k\alpha = \bar{1}$, c'est-à-dire : il existe $k, l \in \mathbb{N}$ tels que $ka = 1 + ln$, ou encore : $\text{pgcd}(a, n) = 1$, ce qui équivaut enfin à : $\alpha \in G(n)$. ■

Nous allons maintenant prouver que la fonction φ d'Euler possède, comme un certain nombre d'autres fonctions arithmétiques intéressantes, la propriété d'être une *fonction multiplicative*, ce qui se traduit par l'énoncé :

THÉORÈME IV.3.4

|| Soit m et n deux entiers **premiers entre eux** ($m \geq 2, n \geq 2$). Alors
|| $\varphi(mn) = \varphi(m) \varphi(n)$.

Démonstration :

Pour tout naturel $a \geq 2$, notons \mathcal{E}_a l'ensemble des entiers $k \in \llbracket 0, a-1 \rrbracket$ qui sont premiers avec a . Si $x \in \mathcal{E}_{mn}$, soit $f(x)$ (resp. $g(x)$) l'entier de $\llbracket 0, m-1 \rrbracket$ (resp. $\llbracket 0, n-1 \rrbracket$) tel que

$$x \equiv f(x) \pmod{m} \quad (\text{resp. } x \equiv g(x) \pmod{n}).$$

x étant premier avec mn est évidemment premier avec m , et avec n . Il est donc clair que $f(x) \in \mathcal{E}_m$ et $g(x) \in \mathcal{E}_n$. On a donc défini une application

$$F: \mathcal{E}_{mn} \rightarrow \mathcal{E}_m \times \mathcal{E}_n, x \mapsto (f(x), g(x)).$$

Il est facile de voir que F est *injective* : en effet si $F(x) = F(x')$, il s'ensuit que $x - x' \equiv 0 \pmod{m}$ et $x - x' \equiv 0 \pmod{n}$, et comme m et n sont premiers entre eux, $x - x' \equiv 0 \pmod{mn}$ (cf. la démonstration du théorème IV.2.4). Mais comme $|x - x'| < mn$, cela entraîne $x = x'$.

Montrons que F est *surjective* : si $(r, s) \in \mathcal{E}_m \times \mathcal{E}_n$, en utilisant une relation de Bezout $\lambda m + \mu n = 1$ on obtient $u = \lambda(r - s)$ et $v = \mu(r - s)$ tels que $r - s = um + vn$, d'où $r - um = s + vn$, entier que nous appelons X . Si x est l'unique élément de $\llbracket 0, mn-1 \rrbracket$ tel que $X - x \equiv 0 \pmod{mn}$, on a d'abord $x \in \mathcal{E}_{mn}$ car X étant premier avec m et n , l'est avec mn (théorème IV.2.3), puis $F(x) = (r, s)$.

Finalement F est *bijective*, d'où

$$\text{card}(\mathcal{E}_{mn}) = \text{card}(\mathcal{E}_m \times \mathcal{E}_n) = \text{card}(\mathcal{E}_m) \times \text{card}(\mathcal{E}_n)$$

et enfin $\varphi(mn) = \varphi(m) \varphi(n)$, compte tenu du théorème IV.3.2. ■

On remarque que ce théorème est valable pour tous $m, n \in \mathbb{N}^*$ avec la convention $\varphi(1) = 1$ qui suit le théorème IV.3.2 et trouve ici une meilleure justification.

THÉORÈME IV.3.5 (Théorème d'Euler)

|| Soit n un entier naturel ≥ 2 , et soit a un entier **premier avec n** . Alors
||
$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

Démonstration :

Notons toujours $x \mapsto \bar{x}$ l'homomorphisme canonique : $\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$.

Raisonnons dans le groupe multiplicatif $G(n)$ des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$. L'application : $G(n) \rightarrow G(n)$, $X \mapsto \bar{a}X$, est bien définie (car $\bar{a} \in G(n)$) et bijective (puisque $(G(n), \cdot)$ est un groupe). Soit $\gamma_1, \gamma_2, \dots, \gamma_{\varphi(n)}$ les éléments de $G(n)$. On a donc une bijection $s : \llbracket 1, \varphi(n) \rrbracket \rightarrow \llbracket 1, \varphi(n) \rrbracket$ telle que : $\bar{a}\gamma_i = \gamma_{s(i)}$ pour $1 \leq i \leq \varphi(n)$. Soit $c_i \in \mathbb{Z}$ tel que $\bar{c}_i = \gamma_i$. Alors $\bar{a}c_i \equiv c_{s(i)} \pmod{n}$ pour $1 \leq i \leq \varphi(n)$. Multipliant membre à membre ces relations, on obtient (1)

$$a^{\varphi(n)} \times C \equiv C \pmod{n},$$

avec $C = c_1 c_2 \dots c_{\varphi(n)}$.

Mais chaque c_i est premier avec n , donc C est premier avec n ; donc les deux membres de la congruence (1) peuvent être divisés par C , d'où $a^{\varphi(n)} \equiv 1 \pmod{n}$. ■

Remarque 1 : Le théorème IV.3.5 sera revu d'une autre manière dans le chapitre V consacré aux groupes.

Remarque 2 : Soit toujours a et n premiers entre eux dans \mathbb{Z} , avec $n \geq 2$. Par définition la **période de $a \pmod{n}$** est l'ordre de l'élément \bar{a} du groupe $G(n)$. Soit d cet ordre : il est tel que $\bar{a}^d = \bar{1}$ et que l'ensemble des $k \in \mathbb{Z}$ tels que $\bar{a}^k = \bar{1}$ est $d\mathbb{Z}$. Le théorème IV.3.5 signifie donc que la **période de $a \pmod{n}$** est un diviseur de $\varphi(n)$.

Exercice 1 : Les éléments inversibles de $\mathbb{Z}/2^n\mathbb{Z}$ sont de la forme $\pm \bar{5}^k$, où $k \in \mathbb{N}$.

Exercice 2 : Les éléments inversibles de $\mathbb{Z}/3^n\mathbb{Z}$ sont de la forme $\pm \bar{2}^k$, où $k \in \mathbb{N}$.

Exercice 3 : Déterminer les entiers naturels n tels que 7 divise $2^{2n} + 2^n + 1$.

Exercice 4 : Résoudre dans \mathbb{N} : $19^n \equiv 2 \pmod{7}$; $2^n \equiv 1 \pmod{9}$.

Exercice 5 (suite de Fibonacci) : soit $(u_n)_{n \in \mathbb{N}}$ la suite de naturels définie par :

$$u_0 = 0, u_1 = 1, u_{n+2} = u_{n+1} + u_n \text{ pour } n \geq 0.$$

a) Montrer que u_n est le nombre de parties non vides de $\llbracket 1, n \rrbracket$ dont deux éléments quelconques ne sont pas des entiers consécutifs.

b) Pour que $u_j = \frac{1}{2}(u_i + u_k)$, il faut et il suffit que $k = j + 1$ et $i = j - 2$.

c) $\forall k \in \mathbb{N} \quad u_{k+1}^2 = u_k u_{k+2} + (-1)^{k+1}$.

d) $\forall k \in \mathbb{N} \quad u_{2k+1} = u_k^2 + u_{k+1}^2$. Peut-on aussi linéariser $u_k^2 + u_{k+1}^2 + u_{k+2}^2$?

e) $\forall (n, p) \in \mathbb{N}^{*2} \quad u_{n+p-1} = u_{n-1} u_{p-1} + u_n u_p$.

f) Soit $\alpha = \frac{1-\sqrt{5}}{2}$ et $\beta = \frac{1+\sqrt{5}}{2}$. Montrer que $u_n = \frac{1}{\sqrt{5}}(\beta^n - \alpha^n)$.

g) Montrer qu'il existe une infinité de valeurs de u_n se terminant par quatre zéros (numération décimale).

Exercice 6 : Evaluer le nombre maximum de divisions successives dans la recherche de $a \vee b$ par l'algorithme d'Euclide.

Exercice 7 : Montrer $\varphi(n) = \sum_{k=1}^{n-1} \left[\frac{1}{n \vee k} \right]$, où φ est l'indicateur d'Euler, et où $[x]$ est la partie entière du réel x .

Exercice 8 : Soit $(a_i)_{i \in \mathbb{N}}$ une suite d'entiers en progression arithmétique. Montrer, en utilisant le théorème IV.3.5 qu'on peut extraire de cette suite une sous-suite de termes en progression géométrique. Généraliser si $a_i \in \mathbb{Q}$ pour tout i .

§ IV.4 NOMBRES PREMIERS

DÉFINITION IV.4.1

⎧ On appelle **nombre premier** tout entier naturel $p \geq 2$ dont les seuls
⎩ diviseurs dans \mathbb{N} sont 1 et p .

Exemple 1 : Les nombres 2, 3, 5, 7, 11, ..., 37, 41, ..., 19 999 999 sont premiers, mais 1, 4, 12 ne le sont pas. Le nombre de Fermat $F_n = 2^{2^n} + 1$ est premier pour $n = 0, 1, 2, 3, 4$ mais pas pour $n = 5$ (cf. exemple 1 du § IV.1). Les nombres de Mersenne $2^{19\,937} - 1$, $2^{21\,701} - 1$, $2^{132\,049} - 1$, $2^{216\,091} - 1$ sont premiers, mais il a fallu des gros ordinateurs pour l'établir (en 1950 le « record » était seulement $2^{127} - 1$).

DÉFINITION IV.4.2

⎧ Dans un anneau **intègre** A , on appelle **élément irréductible** tout
⎩ élément p **non nul**, **non inversible**, et dont les seuls diviseurs dans A
⎩ sont les éléments **inversibles** de A et les éléments **associés** à p .

Exemple 2 : Les éléments irréductibles de l'anneau intègre \mathbb{Z} sont les nombres de la forme εp , où $\varepsilon \in \{-1, +1\}$, et où p est un nombre premier. Autrement dit, $q \in \mathbb{Z}$ est irréductible dans \mathbb{Z} ssi $|q|$ est premier dans \mathbb{N} .

Il est clair que si p est un élément irréductible de l'anneau intègre A , et si $a \in A$, il n'y a que deux cas possibles : ou bien p divise a , ou bien les seuls diviseurs communs à p et a sont les éléments inversibles de A . En particulier :

THÉORÈME IV.4.1

|| Un nombre premier est premier avec tout élément de \mathbb{Z} qu'il ne divise pas. En particulier, deux nombres premiers distincts sont premiers entre eux.

COROLLAIRE

|| Soit p un nombre premier. L'entier $\binom{p}{k}$ est divisible par p pour $k \geq 1$, $k \neq p$.

Démonstration :

C'est évident pour $k > p$ puisqu'alors $\binom{p}{k} = 0$. Si $k \in \llbracket 1, p-1 \rrbracket$, on a :

$$k! \binom{p}{k} = p(p-1)\dots(p-k+1).$$

Or p , premier avec tous les facteurs de $k!$ (théorème I.4.1) est premier avec $k!$ en vertu du théorème IV.2.3. Donc, d'après le théorème de Gauss, p divise $\binom{p}{k}$. ■

THÉORÈME IV.4.2

|| *Tout entier $n \geq 2$ admet au moins un diviseur premier.*

Démonstration :

Soit E l'ensemble des diviseurs de n appartenant à $\llbracket 2, n \rrbracket$; comme $n \in E$, cet ensemble est non vide. Soit p son plus petit élément. D'abord $p \geq 2$, ensuite p divise n , enfin tout diviseur de p divise n : ce ne peut donc être que 1 ou un élément de E divisant p , c'est-à-dire p . Donc p est premier. ■

THÉORÈME IV.4.3 (Euclide)

|| *L'ensemble \mathcal{P} des nombres premiers est infini.*

Démonstration :

On sait déjà que \mathcal{P} est non vide. Si \mathcal{P} était fini, on pourrait former l'entier naturel $N = 1 + \prod_{p \in \mathcal{P}} p$; pour tout p de \mathcal{P} , on a $N \equiv 1 \pmod{p}$, donc p ne divise pas N ; cependant N admet au moins un diviseur premier q car $N \geq 2$. Or un tel q ne serait pas dans \mathcal{P} , d'où une contradiction. ■

Ce raisonnement se trouve déjà dans les *Eléments* d'Euclide (proposition 20 du livre IX). Les nombres premiers forment bien un ensemble car c'est une partie de l'ensemble préexistant \mathbb{N} , constituée d'éléments ayant une propriété particulière. Grâce au théorème IV.4.3, on peut ranger les éléments de \mathcal{P} en une suite strictement croissante, et cela de manière unique (cf. Th. II.3.9) $(p_i)_{i \in \mathbb{N}^*}$, ($p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc...). L'étude de cette suite est aussi passionnante que difficile : elle constitue par définition l'étude de la **répartition des nombres premiers**. Cette répartition est « localement » irrégulière (par exemple il est facile de trouver des intervalles de longueur arbitraire sans nombre premier, alors qu'ailleurs il y a des nombres premiers « jumeaux » p et $p+2$), mais non « aléatoire » (par exemple, le segment $\llbracket n, 2n \rrbracket$ contient toujours des nombres premiers (théorème de Bertrand)). Non seulement \mathbb{N} possède une infinité de nombres premiers, mais toute *progression arithmétique* $\{a + bn\}_{n \in \mathbb{N}}$ avec $\text{pgcd}(a, b) = 1$ possède

infinité de nombres premiers (théorème de Dirichlet). Une chose est sûre, c'est que les nombres premiers vont *en se raréfiant* : si $x \in \mathbb{R}_+$ on désigne habituellement par $\pi(x)$ le nombre de nombres premiers p tels que $p < x$. On peut prouver que $\pi(x) \sim \frac{x}{\text{Log } x}$ (théorème de Hadamard), ce qui équivaut à $p_n \sim n \text{ Log } n$ quand $n \rightarrow +\infty$.

Nombres premiers et anneaux $\mathbb{Z}/n\mathbb{Z}$

THÉOREME IV.4.4

Soit n un entier ≥ 2 . Les propriétés suivantes sont équivalentes :

- (I) n est premier
- (II) l'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre
- (III) l'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Démonstration :

Il est clair que (III) \Rightarrow (II). Montrons que (II) entraîne (III) : si l'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre, soit $X \in \mathbb{Z}/n\mathbb{Z}$, $X \neq 0_{\mathbb{Z}/n\mathbb{Z}} = \bar{0}$.

L'application $\mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$, $Y \mapsto XY$ est injective car X est régulier, donc bijective car l'ensemble $\mathbb{Z}/n\mathbb{Z}$ est fini. D'où, pour X' convenable : $XX' = 1_{\mathbb{Z}/n\mathbb{Z}}$ et $\mathbb{Z}/n\mathbb{Z}$ est bien un corps.

Faisons maintenant l'hypothèse que n est premier. Pour tout $k \in \llbracket 1, n-1 \rrbracket$ les entiers n et k sont premiers entre eux donc (cf. théorème IV.3.1) tous les éléments non nuls de $\mathbb{Z}/n\mathbb{Z}$ sont inversibles, d'où (I) \Rightarrow (II). Il reste à examiner le cas où n est non premier. Soit alors $p, q \in \llbracket 1, n-1 \rrbracket$ tels que $n = pq$; alors les classes \bar{p} et \bar{q} de p et q dans $\mathbb{Z}/n\mathbb{Z}$ sont non nulles et cependant $\bar{p}\bar{q} = \bar{0}$, donc l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre, et ce n'est pas un corps, donc (III) \Rightarrow (I). En fin de compte on a bien l'équivalence annoncée des trois propriétés. ■

Le théorème de Fermat

THÉOREME IV.4.5 (« Petit » théorème de Fermat)

Si p est un nombre premier, et si a est un entier non divisible par p , alors

$$a^{p-1} \equiv 1 \pmod{p}.$$

C'est un cas particulier du théorème IV.3.5 car si p est premier, l'indicateur d'Euler $\varphi(p)$ vaut $\text{card}(\llbracket 1, p-1 \rrbracket) = p-1$. Nous donnerons une preuve différente de ce théorème dans le prochain chapitre consacré aux groupes. A ce jour plus de CENT démonstrations de ce théorème sont connues⁽¹⁾

⁽¹⁾ Cf. Dickson, L. E., *History of the theory of numbers*, Tome 1.

Si nous interprétons le théorème IV.4.5 dans le corps $\mathbb{Z}/p\mathbb{Z}$, il signifie : $\forall X \in \mathbb{Z}/p\mathbb{Z} \setminus \{0\}, X^{p-1} = \bar{1}$, ce qui équivaut, puisque $\mathbb{Z}/p\mathbb{Z}$ est un corps à : $\forall X \in \mathbb{Z}/p\mathbb{Z}, X^p - X = \bar{0}$.

THÉOREME IV.4.6 (Théorème de Wilson)

Soit p un entier ≥ 2 . Pour que p soit premier, il faut et il suffit que

(I)
$$(p-1)! + 1 \equiv 0 \pmod{p}.$$

Démonstration (abrégée) :

Il est d'abord clair que (I) $\Rightarrow p$ premier, sinon on aurait $(p-1)! \equiv 0 \pmod{p}$. Réciproquement, si p est premier, on peut supposer $p \geq 5$ car pour $p = 2$ ou $p = 3$ (I) est vérifié. On remarque alors que les seuls $X \in \mathbb{Z}/p\mathbb{Z}$ tels que $X^2 = \bar{1}$ sont $X = \bar{1}$ et $X = -\bar{1} \neq \bar{1}$ (en effet $(X - \bar{1})(X + \bar{1}) = \bar{0}$ ssi $X = \bar{1}$ ou $X = -\bar{1}$ car il n'y a pas de diviseurs de zéro). Les éléments X de $E = \mathbb{Z}/p\mathbb{Z} \setminus \{\bar{0}, \bar{1}, -\bar{1}\}$ vérifient donc $X^2 \neq \bar{1}$, c'est-à-dire $X \neq \frac{1}{X}$. On définit alors sur E une relation d'équivalence \mathcal{R} en posant : $X \mathcal{R} Y$ ssi $(X = Y \text{ ou } XY = \bar{1})$. Chaque classe \mathcal{R} contient exactement deux éléments dont le produit est $\bar{1}$, d'où $\prod_{X \in E} X = \bar{1}$. Donc

$$\prod_{X \in \mathbb{Z}/p\mathbb{Z} \setminus \{0\}} X = (-\bar{1}) \times (\bar{1}) \times \prod_{X \in E} X = -\bar{1},$$

ce qui est exactement l'énoncé IV.4.6. ■

Exemple 3 : D'après le théorème de Fermat, le nombre $10^6 - 1$ est divisible par 7. Il est facile de le vérifier en divisant 999 999 par 7.

D'après le théorème de Wilson, le nombre $10! + 1$ est divisible par 11. On peut le vérifier en divisant 3 628 801 par 11, ou plus simplement en utilisant un critère de divisibilité par 11 fondé sur le fait que

$$10 \equiv -1 \pmod{11} \Rightarrow 10^n \equiv (-1)^n \pmod{11}.$$

Remarquons que la réciproque du théorème de Fermat est fausse (par exemple $341 \mid 2^{341} - 2$ et pourtant $341 = 11 \times 31$ n'est pas premier).

Sous-corps premier. Caractéristique d'un corps

Considérons un corps commutatif K . L'élément 1_K définit un homomorphisme d'anneaux $\psi : \mathbb{Z} \rightarrow K, n \mapsto n \cdot 1_K$. L'image A de ψ est un sous-anneau de K , et, au sens de l'inclusion, c'est le plus petit sous-anneau de K (puisque un tel sous-anneau doit contenir 1_K , donc A).

Deux cas se présentent alors :

1^{er} cas : ψ est injectif

Dans ce cas, l'anneau A est isomorphe à \mathbb{Z} ; d'après le théorème II.6.3, l'homomorphisme ψ se prolonge de manière unique en un isomorphisme $\widehat{\psi}$ du corps \mathbb{Q} dans K . L'image P de $\widehat{\psi}$ est donc un sous-corps de K isomorphe à \mathbb{Q} , et qui est donc le corps des fractions de A . Ce corps P est au sens de l'inclusion, le *plus petit sous-corps de K* , puisqu'un sous-corps L de K doit contenir A , donc tous les inverses dans K des éléments de $A \setminus \{0\}$, donc P .

Ce corps P s'appelle le **sous-corps premier de K** , et on dit, dans ce cas étudié, que **le corps K est de caractéristique 0**.

Habituellement, dans un tel corps, le sous-corps premier P est identifié à \mathbb{Q} , et K apparaît alors comme une *extension de \mathbb{Q}* .

2^e cas : ψ n'est pas injectif

Remarquons d'abord que l'assertion « ψ n'est pas injectif » signifie exactement que *l'élément 1_K est de torsion* dans le groupe additif $(K, +)$. Dans ce cas, le noyau $\text{Ker}(\psi)$ est de la forme : $\text{Ker}(\psi) = p\mathbb{Z}$, où $p \in \mathbb{N}^*$ est l'ordre de 1_K dans le groupe additif $(K, +)$.

Considérons la relation d'équivalence \mathcal{R} définie sur \mathbb{Z} par ψ ; on a $x \mathcal{R} y$ ssi $\psi(x) = \psi(y)$, c'est-à-dire, ssi $\psi(x - y) = 0_K$, ou encore : $x - y \in \text{Ker}(\psi) = p\mathbb{Z}$. Autrement dit, \mathcal{R} est la *congruence modulo p* . L'ensemble quotient \mathbb{Z}/\mathcal{R} n'est autre que $\mathbb{Z}/p\mathbb{Z}$, et la bijection naturelle $\overline{\psi}$ associée à ψ (cf. § I.5) est une bijection $\overline{\psi} : \mathbb{Z}/p\mathbb{Z} \rightarrow A$.

Du fait que ψ et $\mathbb{Z} \xrightarrow{\text{can}} \mathbb{Z}/p\mathbb{Z}$ sont des homomorphismes d'anneaux, on déduit facilement que $\overline{\psi}$ est aussi un homomorphisme d'anneaux. Donc, $\overline{\psi}$ est un isomorphisme d'anneaux.

Mais A est un anneau intègre, car c'est un sous-anneau du corps K , lui-même anneau intègre. Donc, $\mathbb{Z}/p\mathbb{Z}$ est un anneau intègre. D'après le théorème IV.4.4, il en résulte que **p est un nombre premier**, et que $\mathbb{Z}/p\mathbb{Z}$, donc aussi A , est un corps. Dans ce cas, on dit que **le corps K est de caractéristique p** , et **A est appelé le sous-corps premier de K** ; c'est encore le plus petit sous-corps de K (et même, ici, le plus petit sous-anneau de K). En résumé :

Si l'élément 1_K de K est sans torsion dans le groupe additif $(K, +)$, le corps K est de caractéristique nulle, et s'identifie à une extension de son sous-corps premier P , qui est isomorphe à \mathbb{Q} .

Si l'élément 1_K de K est de torsion dans le groupe additif $(K, +)$, son ordre dans ce groupe est un nombre premier p , appelé caractéristique de K . Dans ce cas, le sous-corps premier de K est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

Il résulte de ce qui précède que les **extensions** de \mathbb{Q} , comme \mathbb{R} , \mathbb{C} par exemple, sont les corps de caractéristique nulle.

Exemple 4 : Si K est de caractéristique $p \neq 0$, pour tous $x, y \in K$, on a : $(x + y)^p = x^p + y^p$. En effet, la formule du binôme donne :

$$(x + y)^p = x^p + y^p + \sum_{k=1}^{p-1} \binom{p}{k} x^k y^{p-k}.$$

Mais (cf. coroll. du Th. IV.4.1), pour $1 \leq k \leq p-1$, $\binom{p}{k} \equiv 0 \pmod{p}$, d'où $\binom{p}{k} \cdot 1_K = 0_K$ pour $1 \leq k \leq p-1$, d'où le résultat.

Par récurrence, il s'ensuit, si $n \in \mathbb{N}$:

$$(x + y)^{p^n} = x^{p^n} + y^{p^n}.$$

Exercice 1 : Résoudre dans le corps $\mathbb{Z}/7\mathbb{Z}$ l'équation $X^7 = \bar{2}$.

Exercice 2 : Montrer : $(\forall n \in \mathbb{N}) \quad n^7 - n \equiv 0 \pmod{42}$.

Exercice 3 : Montrer qu'il y a une infinité de nombres premiers de la forme $4n-1$ (resp. de la forme $6n-1$).

Exercice 4 : Trouver $a \in \mathbb{N}^*$ pour que $1 + a + a^2 + a^3 + a^4$ soit un carré parfait. (Réponse $a = 3$.)

Exercice 5 : Pour $n \geq 2$ le nombre $\frac{1}{4} [n^3 + (n+2)^3]$ est entier et non premier.

Exercice 6 : Pour $n \geq 2$ le nombre $n^4 + 4$ n'est pas premier. Peut-on trouver n tel que $n^4 + 4^n$ soit premier ?

Exercice 7 : En utilisant la représentation paramétrique $t \mapsto \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right)$ du cercle $x^2 + y^2 - 1 = 0$ privé du point $(-1, 0)$, trouver tous les triplets (x, y, z) de \mathbb{Z}^3 solutions de l'équation de Pythagore $x^2 + y^2 = z^2$.

Exercice 8 : Soit a et n deux entiers, $a \geq 1, n \geq 2$. On suppose $a \nmid n = 1$, et que la période de $a \pmod{n}$ est $n-1$. Montrer que n est premier.

Exercice 9 : Montrer que toute application f de $\mathbb{Z}/p\mathbb{Z}$ dans lui-même, p premier, coïncide avec une application polynomiale. Combien cela fait-il d'applications distinctes ?

Exercice 10 : Dans $\mathbb{Z}/p\mathbb{Z}$, p premier, étudier la suite (u_n) telle que $u_{n+1} = 1 - \frac{1}{u_n}$, $u_0 \neq \bar{0}$ et montrer qu'elle est périodique. En déduire le nombre de racines, dans $\mathbb{Z}/p\mathbb{Z}$, de $X^2 - X + \frac{1}{1} = \bar{0}$.

Exercice 11 : Pour $m \neq n$ entiers naturels, montrer que $2^{2^m} + 1 \nmid 2^{2^n} + 1$. En déduire que l'ensemble des nombres premiers est infini.

Exercice 12 : Soit p_n le n -ième nombre premier. Montrer que si $n \geq 3$

$$p_1 p_2 \cdots p_n \geq p_{n+1} + p_{n+2}.$$

Exercice 13 : Le corollaire du théorème IV.4.2 montre que si p est premier, alors p divise $\binom{p}{k}$ pour $k \in \llbracket 1, p-1 \rrbracket$. En déduire le théorème de Fermat : $a^p \equiv a \pmod{p}$ en utilisant les congruences \pmod{p} .

Exercice 14 : a) Si $n \in \mathbb{N}^*$, pour tout $k \in \mathbb{N}^* k(k+1)\dots(k+n-1)$ est divisible par $n!$. Le montrer sans utiliser les coefficients binomiaux.

b) Soit p un nombre premier et n un entier $\in \llbracket 1, p-1 \rrbracket$. Montrer que le nombre $\frac{1}{n!} (p-1)(p-2)\dots(p-n) + (-1)^{n-1}$ est divisible par p .

§ IV.5 DÉCOMPOSITION EN FACTEURS PREMIERS

Dans ce qui suit, on note \mathcal{P} l'ensemble des nombres premiers et $(p_n)_{n \geq 1}$ la suite de ces nombres. Soit n un entier ≥ 1 ; si nous fixons un nombre premier p , l'ensemble $\mathcal{V}_p(n)$ des $\alpha \in \mathbb{N}$ tels que p^α divise n est non vide, car il contient 0. Cet ensemble est fini, car l'application $\alpha \mapsto p^\alpha$ est strictement croissante, d'où $p^\alpha > n$ pour $\alpha \geq n$, d'où $\mathcal{V}_p(n) \subset \llbracket 0, n-1 \rrbracket$. Donc l'ensemble $\mathcal{V}_p(n)$ admet un plus grand élément.

DÉFINITION IV.5.1

*Si p est un nombre premier et si $n \in \mathbb{N}$, $n \geq 1$, on appelle **p-valuation de n** le plus grand des entiers $\alpha \in \mathbb{N}$ tels que $p^\alpha | n$. On note $v_p(n)$ cette p-valuation. On appelle **support premier de n** (et nous noterons $\mathcal{P}(n)$) l'ensemble des nombres premiers p tels que $v_p(n) \geq 1$.*

On voit que $\mathcal{P}(n)$ est non vide ssi $n \geq 2$ (cf. théorème IV.4.2). De plus, $\mathcal{P}(n)$ est toujours un ensemble fini, puisqu'il est inclus dans l'ensemble des diviseurs de n dans \mathbb{N} , lui-même inclus dans $\llbracket 1, n \rrbracket$.

Si $n \in \mathbb{N}^*$, il résulte de tout ce qui précède que le produit $\prod_{p \in \mathcal{P}} p^{v_p(n)}$ a un sens (cf. § III.3). C'est par définition l'élément $\prod_{p \in \mathcal{P}(n)} p^{v_p(n)}$.

THÉORÈME IV.5.1

Pour tout entier $n \geq 1$, on a :

$$(I) \quad \boxed{n = \prod_{p \in \mathcal{P}} p^{v_p(n)}}.$$

Démonstration :

C'est évident pour $n = 1$ car alors le support premier de n est vide. Supposons $n \geq 2$ et notons

$$\mu = \prod_{p \in \mathcal{P}} p^{v_p(n)} = \prod_{p \in \mathcal{P}(n)} p^{v_p(n)}.$$

Si $p \in \mathcal{P}(n)$, $q \in \mathcal{P}(n)$, $p \neq q$, alors p et q sont premiers entre eux, donc aussi $p^{v_p(n)}$ et $q^{v_q(n)}$ (par application répétée du théorème

chaque $p^{v_p(n)}$ divise n ; donc μ divise n (application du théorème IV.2.4). Soit q' le quotient exact $\frac{n}{\mu}$. Montrons par l'absurde que $q' = 1$. Si l'on avait $q' \neq 1$, q' admettrait un diviseur premier p_0 ; on aurait donc $p_0 \in \mathcal{P}(n)$, et $q' = kp_0$ avec $k \in \mathbb{N}^*$, d'où

$$n = q' \mu = kp_0 \times \prod_{p \in \mathcal{P}(n)} p^{v_p(n)} = kp_0 \times p_0^{v_{p_0}(n)} \times \prod_{p \in \mathcal{P}(n), p \neq p_0} p^{v_p(n)},$$

d'où : $p_0^{v_{p_0}(n)+1}$ diviserait n , ce qui est absurde à cause de la définition de $v_{p_0}(n)$. Donc $q' = 1$ et $\mu = n$. ■

La relation (I) du théorème IV.5.1 est appelé **décomposition en facteurs premiers** de l'entier naturel n . Nous allons voir que la famille $(v_p(n))_{p \in \mathcal{P}}$ est la **seule** famille $(\alpha_p)_{p \in \mathcal{P}}$ à support fini d'entiers naturels telle que : $n = \prod_{p \in \mathcal{P}} p^{\alpha_p}$.

COROLLAIRE 1

Soit $n \in \mathbb{N}$, $n \geq 2$. Soit q_1, q_2, \dots, q_k des nombres premiers distincts et $\beta_1, \beta_2, \dots, \beta_k$ dans \mathbb{N}^* tels que $n = \prod_{i=1}^k q_i^{\beta_i}$. Alors le support premier $\mathcal{P}(n)$ est $\{q_i\}_{1 \leq i \leq k}$, et $\beta_i = v_{q_i}(n)$ pour tout i .

Démonstration :

Soit E l'ensemble $\{q_1, q_2, \dots, q_k\}$. Chaque q_i divisant n , il est clair que $E \subset \mathcal{P}(n)$; réciproquement si $p \in \mathcal{P}(n)$, p divise l'un des q_i par le théorème de Gauss, donc est égal à l'un d'eux, d'où finalement $E = \mathcal{P}(n)$. Pour $i \in \llbracket 1, k \rrbracket$, il est clair que $\beta_i \leq v_{q_i}(n)$; si l'on avait $\beta_i < v_{q_i}(n)$ la relation $\prod_{p \in \mathcal{P}(n)} p^{v_p(n)} = \prod_{j=1}^k q_j^{\beta_j}$ pourrait être divisée par $q_i^{\beta_i}$ et fournirait :

$$\prod_{j=1, j \neq i}^k q_j^{\beta_j} = q_i^{v_{q_i}(n) - \beta_i} \times \prod_{p \in \mathcal{P}(n), p \neq q_i} p^{v_p(n)},$$

relation dans laquelle le membre de droite est divisible par q_i ; mais alors le théorème de Gauss montrerait que q_i divise l'un des q_j pour $j \neq i$, ce qui est absurde. Donc $v_{q_i}(n) = \beta_i$. ■

Nous laissons au lecteur les détails de la preuve du :

COROLLAIRE 2

Tout entier $n \in \mathbb{Z}^*$ s'écrit de manière unique sous la forme $n = \varepsilon \prod_{p \in \mathcal{P}} p^{\alpha_p}$, où (α_p) est une famille à support fini d'entiers naturels, et où $\varepsilon \in \{-1, 1\}$.

α_p figurant dans ce corollaire est évidemment égal à $v_p(|n|)$ pour tout p . On convient d'écrire encore $\alpha_p = v_p(n)$, même si n est < 0 .

Le théorème IV.5.1 prouve l'existence de la décomposition d'un entier naturel en facteurs premiers, le corollaire 1 peut s'interpréter en disant que la décomposition (I) ainsi trouvée est *unique à la numérotation près des facteurs*, ce qui dans la pratique laisse une grande liberté pour effectuer cette décomposition, mais le lecteur aura remarqué que grâce à la notion de support premier de n , le fait que l'ensemble \mathcal{P} soit infini n'a pas gêné les démonstrations. Nous pourrions rencontrer des anneaux commutatifs où la décomposition en « facteurs premiers » n'est pas unique, mais cela n'a rien à voir avec le fait que l'ensemble de ces « facteurs premiers » soit fini ou infini.

Application au pgcd et au ppcm

Soit a et b dans \mathbb{Z}^* ; la p -valuation définie en IV.5.1 et prolongée à \mathbb{Z}^* après le corollaire 2 jouit des propriétés suivantes :

$$(V_1) \quad v_p(ab) = v_p(a) + v_p(b) \quad \text{pour tout } p \in \mathcal{P}$$

$$(V_2) \quad a \text{ divise } b \quad \text{ssi} \quad (\forall p \in \mathcal{P}) \quad v_p(a) \leq v_p(b).$$

La propriété (V_1) est immédiate, de même que la condition nécessaire dans (V_2) . Pour la réciproque, si on suppose $v_p(a) \leq v_p(b)$ pour *tout* p , on a : $b = a\lambda$ avec $\lambda = \varepsilon \prod_{p \in \mathcal{P}} p^{v_p(b) - v_p(a)}$, donc a est bien un diviseur de b .

Soit alors des entiers (a_1, a_2, \dots, a_n) dans \mathbb{Z}^* . Il résulte de (V_1) et (V_2) :

$$(1) \quad \text{pgcd}(a_1, a_2, \dots, a_n) = \prod_{p \in \mathcal{P}} p^{\alpha_p} \quad \text{avec} \quad (\forall p \in \mathcal{P}) \quad \alpha_p = \min_{1 \leq i \leq n} (v_p(a_i))$$

$$(2) \quad \text{ppcm}(a_1, a_2, \dots, a_n) = \prod_{p \in \mathcal{P}} p^{\beta_p} \quad \text{avec} \quad (\forall p \in \mathcal{P}) \quad \beta_p = \max_{1 \leq i \leq n} (v_p(a_i)).$$

Exemple 1 : Si $a = 999\,999$ et $b = 99\,999$ la décomposition en facteurs premiers donne

$$a = 999 \times 1\,001 = 9 \times 111 \times 11 \times 91 = 3^3 \times 7 \times 11 \times 13 \times 37$$

$$\text{et} \quad b = 9 \times 11\,111 = 3^2 \times 41 \times 271$$

$$\text{soit} \quad \mathcal{P}(a) = \{3, 7, 11, 13, 37\} \quad \text{et} \quad \mathcal{P}(b) = \{3, 41, 271\}$$

$$\text{avec} \quad v_3(a) = 3, v_7(a) = 1, v_{11}(a) = 1, v_{37}(a) = 1$$

$$\text{et} \quad v_3(b) = 2, v_{41}(b) = 1, v_{271}(b) = 1.$$

Donc (1) et (2) donnent

$$\text{pgcd}(a, b) = 3^2 = 9$$

$$\text{ppcm}(a, b) = 3^3 \times 7 \times 11 \times 13 \times 37 \times 41 \times 271 = 11\,110$$

Indicateur d'Euler

Soit α un entier ≥ 1 et soit p un nombre premier ; les nombres k , éléments de $\llbracket 1, p^\alpha - 1 \rrbracket$ qui sont premiers avec p^α sont ceux qui sont non divisibles par p . Or les nombres divisibles par p dans $\llbracket 1, p^\alpha - 1 \rrbracket$ sont les multiples de $p : p, 2p, 3p, \dots, qp$ où $q = p^{\alpha-1} - 1$; ils sont au nombre de $p^{\alpha-1} - 1$; donc le nombre des $k \in \llbracket 1, p^\alpha - 1 \rrbracket$ qui sont premiers avec p est égal à

$$p^\alpha - 1 - (p^{\alpha-1} - 1) = p^\alpha - p^{\alpha-1}.$$

Ce nombre est l'indicateur d'Euler de p^α (cf. le théorème IV.3.2). On a donc $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}(p - 1)$.

Soit alors un entier quelconque $n \geq 2$. On peut le décomposer en facteurs premiers :

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)} = \prod_{p \in \mathcal{P}} p^{\alpha_p}.$$

Si $p \notin \mathcal{P}(n)$, on a convenu que $\varphi(p^{\alpha_p}) = \varphi(p^0) = \varphi(1) = 1$.

Si $p \in \mathcal{P}(n)$, on vient de voir que $\varphi(p^{\alpha_p}) = p^{\alpha_p-1}(p - 1)$.

Mais les nombres $(p^{\alpha_p})_{p \in \mathcal{P}(n)}$ sont deux à deux premiers entre eux, d'où en appliquant le théorème IV.3.4 (la multiplicativité de l'indicateur d'Euler s'étend par récurrence à un nombre fini d'entiers **deux à deux** premiers entre eux),

$$(3) \quad \varphi(n) = \prod_{p \in \mathcal{P}(n)} p^{\alpha_p-1}(p - 1),$$

que l'on peut écrire :

$$\varphi(n) = \prod_{p \in \mathcal{P}} p^{\alpha_p-1}(p - 1) = \boxed{n \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right)}.$$

Cette relation permet donc de calculer $\varphi(n)$ si l'on connaît les facteurs premiers de n . La relation (3) est due à Euler.

Exemple 2 : Avec $a = 999\,999 = 3^3 \times 7 \times 11 \times 13 \times 37$, on obtient

$$\varphi(a) = a \times \left(\frac{2}{3}\right) \times \left(\frac{6}{7}\right) \times \left(\frac{10}{11}\right) \times \left(\frac{12}{13}\right) \times \left(\frac{36}{37}\right) = 466\,560.$$

Remarque 1 : Le théorème IV.5.1 et son corollaire 1 sont souvent réunis sous le nom de « *théorème fondamental de l'Arithmétique* » car il est évident que le fait de savoir que tout naturel n peut se décomposer de manière unique en produit de *facteurs premiers* peut rendre de multiples services. Par exemple si l'on a : $n = q_1^{\alpha_1} q_2^{\alpha_2} \dots q_r^{\alpha_r}$, on sait trouver *tous les diviseurs* de n : ils sont de la forme $q_1^{\beta_1} \dots q_r^{\beta_r}$ avec $(\forall i \in \llbracket 1, r \rrbracket) \beta_i \leq \alpha_i$. Mais le lecteur s'étonnera peut-être de ne pas avoir à sa disposition un procédé *pratique*, rapide et sûr, d'obtenir une telle décomposition. Il peut, à partir d'une table de nombres premiers, essayer si les divisions succ

3, 5, 7, 11, etc... « tombent juste », quitte à s'arrêter dès que le quotient entier obtenu sera inférieur au diviseur, mais cela risque de prendre beaucoup de temps, même avec un gros ordinateur (exemple : $F_{13} = 2^{2^{13}} + 1$ est-il premier ou non ?). C'est qu'en réalité, dès qu'un nombre s'écrit avec quelques dizaines de chiffres dans un *système de numération de base b* (nous proposons quelques exercices sur ce sujet), le problème devient extrêmement difficile et les auteurs ne connaissent pas de recette miraculeuse qui s'appliquerait de façon générale.

Exercice 1 : Montrer que $p = 1\,093$ est un nombre premier et prouver que :

$$2^{1092} - 1 \equiv 0 \pmod{p^2}.$$

Indication : Calculer 3^{14} et $2^{182} \pmod{p^2}$.

Exercice 2 : Soit p un nombre premier > 17 . Alors $p^{16} - 1 \equiv 0 \pmod{16\,320}$.

Exercice 3 (Nombres de Fermat). Pour $n \in \mathbb{N}$ on pose $F_n = 2^{2^n} + 1$. Calculer F_n pour $n \leq 4$. Vérifier qu'on obtient 5 nombres premiers. En revanche F_5 n'est pas premier (cf. l'exemple 1 qui suit le théorème IV.1.1). Vérifier à l'aide d'un ordinateur que F_6 est divisible par $q = 274\,177$ qui est son plus petit diviseur autre que 1. Montrer que si $m \in \mathbb{N}^*$ et si $2^m + 1$ est premier, alors nécessairement m est une puissance de 2.

Remarque : A l'heure actuelle on ne connaît pas de nombre de Fermat premier pour $n \geq 5$. En revanche on en connaît beaucoup de « composés » (par exemple $F_5, F_6, F_7, F_8, \dots, F_{16}, F_{18}, F_{19}, F_{21}, F_{23}$, et bien d'autres : F_{23} est divisible par $167\,772\,161$, F_{1945} qui a plus de 10^{582} chiffres est divisible par $5 \times 2^{1947} + 1$ qui a 587 chiffres !).

Exercice 4 : Trouver tous les nombres premiers de la forme $n^n + 1$ qui ont moins de 300 000 chiffres (Sierpinski).

Indication : On admettra que F_6 et F_{11} sont composés.

Exercice 5 : La 2-valuation de $5^{2^n - 2} - 1$ est n pour $n \geq 2$.

Exercice 6 : De combien de façons peut-on décomposer un entier en produit de deux facteurs premiers entre eux, l'ordre étant indifférent ?

Exercice 7 : Soit $n \in \mathbb{N}$, $n \geq 2$. On décompose n en facteurs premiers :

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} \quad (r \geq 1, \alpha_i \geq 1).$$

1) On note \mathcal{D}_n l'ensemble de ses diviseurs

a) vérifier que $D(n) = \text{card } \mathcal{D}_n = (\alpha_1 + 1)(\alpha_2 + 1) \dots (\alpha_r + 1)$;

b) quel est le plus petit entier $n \in \mathbb{N}^*$ ayant 17 diviseurs ?

c) montrer que si m et n sont premiers entre eux $D(mn) = D(m)D(n)$;

d) montrer que n est carré parfait ssi $D(n)$ est impair.

2) Pour $k \in \mathbb{N}$, on désigne par $S_k(n)$ la somme des puissances k -ièmes des diviseurs de n .

a) Prouver que

$$S_k(n) = \prod_{i=1}^r \frac{p_i^{k(\alpha_i + 1)} - 1}{p_i - 1} \quad \text{pour } k \geq 1.$$

b) Vérifier que si $m \vee n = 1$, $S_k(mn) = S_k(m)S_k(n)$.

3) a) Montrer que $n^{D(n)}$ est un carré parfait, et que le produit de tous les diviseurs de n est $\sqrt{n^{D(n)}}$.

b) Trouver un entier dont le produit des diviseurs est $3^{30} \times 5^{40}$.

Exercice 8 :

a) Soit a un entier impair premier avec 3 et 5. Prouver que :

$$(a^2 - 1)(a^4 - 16)[a^2 - (2n + 1)^2] \equiv 0 \pmod{23\,040}.$$

b) Si a est impair et premier avec 5, $(a^2 - 1)(a^2 - 9)(a^2 - 49) \equiv 0 \pmod{23\,040}$.

Exercice 9 : Soit p un nombre premier ≥ 5 . On met le rationnel $1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{p-1}$ sous forme irréductible $\frac{A_p}{B_p}$; montrer que $A_p \equiv 0 \pmod{p^2}$ (théorème de Wolstenhome).

Exercice 10 : Soit n un entier ≥ 2 et p un nombre premier $\geq n$. La partie entière d'un réel x sera notée $[x]$.

a) Si $i, j \in \mathbb{N}$ et $p^{i+j} \leq n$, alors $\left[\frac{1}{p^j} \left[\frac{n}{p^i} \right] \right] = \left[\frac{n}{p^{i+j}} \right]$.

b) Si $m = \left[\frac{n}{p} \right]$, on a : $v_p(n!) = m + v_p(m!)$. En déduire que : $v_p(n!) = \sum_{k=1}^s \left[\frac{n}{p^k} \right]$ où s est le plus grand des $\lambda \in \mathbb{N}$ tel que $p^\lambda \leq n$.

c) *Application :* Décomposer $(1\,000)!$ en facteurs premiers et vérifier (sans ordinateur) que l'écriture décimale de ce nombre se termine par 249 zéros. Programmer sur ordinateur la décomposition en facteurs premiers de $10\,000!$ et vérifier que $v_7(10\,000!) = 1\,665$.

Exercice 11 : Soit m, n, k dans \mathbb{N}^* tels que $m = nk$ (m, n et $k \geq 2$)

a) prouver : $(n!)^k$ et $(k!)^n$ divisent $m!$ (on peut utiliser l'exercice 10) ;

b) prouver que $\text{ppcm}((n!)^k, (k!)^n)$ divise $m!$

Exercice 12 : On donne $m, n \in \mathbb{N}^*$, $m \geq 2$, $n \geq 2$, non premiers entre eux. On décompose $m \vee n$ en facteurs premiers : $m \vee n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ ($r \geq 1$, $\alpha_i \geq 1$) et on pose $P = p_1 p_2 \dots p_r$. Démontrer que $\varphi(mn) = P \frac{\varphi(m) \varphi(n)}{\varphi(P)}$.

Exercice 13 (nombres de Mersenne) :

a) Soit a un entier > 2 . Alors pour $n > 1$, $a^n - 1$ n'est pas premier.

b) Si $n \in \mathbb{N}$, $n \geq 2$ et si $2^n - 1$ est premier, alors n est nécessairement premier.

Note : Si p est premier, le nombre $M_p = 2^p - 1$ s'appelle *nombre de Mersenne* d'indice p . La recherche des M_p qui sont premiers est un problème toujours ouvert. S'il est facile de vérifier que $23 \mid M_{11}$ ou que $47 \mid M_{23}$, on connaît beaucoup de nombres de Mersenne premiers (par exemple $M_2, M_3, M_5, M_7, M_{127}$) et même de très grands ($M_{19\,937}, M_{216\,091}$ sont premiers).

Exercice 14 (nombres parfaits) :

$n \in \mathbb{N}^*$ s'appelle *nombre parfait* ssi la somme de ses diviseurs est égale à $2n$.

a) Montrer que si p est un nombre premier tel que $2^p - 1$ soit premier, alors $E_p = 2^{p-1}(2^p - 1)$ est un nombre parfait (E_p est le p -ième *nombre d'Euclide*). Calculer E_2, E_3, E_5, E_7 .

b) Réciproquement soit n un nombre parfait *pair*. Mettre n sous la forme $2^a \times b$, où $a \geq 1$ et où b est impair. En déduire que $D(n) = D(b) \times (2^{a+1} - 1) = 2n = 2^{a+1} \times b$, d'où $b = (2^{a+1} - 1)c$ et $D(b) = 2^{a+1}c$. Prouver que $c = 1$ et que $2^{a+1} - 1$ est premier. Conclure que les seuls nombres parfaits pairs sont les nombres d'Euclide E_p .

N.B. : Peut-être un jour trouvera-t-on un nombre parfait impair, mais il devra être très grand, à moins que l'on puisse prouver qu'il n'en existe pas !

Exercice 15 : Soit n un entier ≥ 2 , \mathcal{D}_n l'ensemble de ses diviseurs dans \mathbb{N} et $\varphi(n)$ son indicateur d'Euler.

a) Si $a \in \llbracket 1, n \rrbracket$ soit $\delta(a) = a \vee n$. Montrer que $\delta : \llbracket 1, n \rrbracket \rightarrow \mathcal{D}_n$ est surjective.

b) Soit $\Delta_d = \{a \in \llbracket 1, n \rrbracket \mid \delta(a) = d\}$ pour $d \in \mathcal{D}_n$. Prouver : $\text{card}(\Delta_d) = \frac{n}{d} \varphi\left(\frac{d}{n \vee d}\right)$

c) Montrer que les $(\Delta_d)_{d \in \mathcal{D}_n}$ forment une partition de $\llbracket 1, n \rrbracket$ et en déduire la formule d'Euler-Gauss : $\sum_{d \in \mathcal{D}_n} \varphi(d) = n$.

d) Prouver $\sum_{a \in \Delta_1} \varphi(a) = \frac{1}{2} n \varphi(n)$.

Exercice 16 : Soit deux suites d'entiers $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)$ dans \mathbb{N}^* telles que $x_1 y_1 = x_2 y_2 = \dots = x_n y_n = M$. Soit $\Delta_x = \text{pgcd}(x_1, x_2, \dots, x_n)$, $\mu_x = \text{ppcm}(x_1, x_2, \dots, x_n)$ et de même pour Δ_y et μ_y . Montrer que $\Delta_x \mu_y = \mu_x \Delta_y = M$. En déduire : le ppcm de n entiers est égal au produit de ces nombres divisé par le pgcd de leurs produits $n-1$ à $n-1$ et de même : le pgcd de n entiers est égal à leur produit divisé par le ppcm de leurs produits $n-1$ à $n-1$ (André).

Exercice 17 : On définit dans \mathbb{N}^* la fonction μ de Möbius par : $\mu(1) = 1$, $\mu(n) = 0$ si n possède au moins un diviseur qui est un carré parfait > 1 et $\mu(n) = (-1)^k$ si $n = p_1 p_2 \dots p_k$, produit de k facteurs premiers distincts.

a) Montrer que si $a \vee b = 1$, alors $\mu(ab) = \mu(a) \mu(b)$.

b) Montrer que pour tout $n \geq 2$, $\sum_{d \in \mathcal{D}_n} \mu(d) = 0$.

Exercice 18 : Prouver que si l'un des deux nombres $2^n - 1$ et $2^n + 1$ est premier ($n \geq 3$), alors l'autre est composé.

Exercice 19 : Montrer que le rationnel $(5n)! / (40^n)(n!)$ est entier.

Exercice 20 : Montrer que $(\forall m \in \mathbb{N})(\forall n \in \mathbb{N})$, le nombre $\frac{(2m)!(2n)!}{m! n! (m+n)!}$ est entier.

Exercice 21 : Soit $a_1, \dots, a_n \in \mathbb{N}^*$. Soit P_k le produit des pgcd de ces nombres pris k à k ($1 \leq k \leq n$; $P_1 = a_1 a_2 \dots a_n$).

Démontrer :
$$\text{ppcm}(a_1, a_2, \dots, a_n) = \frac{P_1 \cdot P_3 \cdot P_5 \dots}{P_2 \cdot P_4 \cdot P_6 \dots}.$$

Voici maintenant quelques exercices sur les systèmes de numération. Par convention, quand on ne précise pas le contraire, les entiers sont écrits dans le système décimal habituel.

Exercice 1 : Tout nombre impair dont l'écriture décimale ne se termine pas par 5 a un multiple qui ne s'écrit qu'avec des 1 (exemple : $n = 33$).

Exercice 2 : Le carré de 111 111 111 est $N = 12\,345\,678\,987\,654\,321$. Montrer sans calculer N^2 que son chiffre du milieu est un 2.

Exercice 3 : On écrit un nombre $N = \overline{abcd}$ en numération décimale. Trouver N tel que N , \overline{ab} et \overline{cd} soient des carrés parfaits.

Exercice 4 : Trouver un carré parfait s'écrivant avec 4 chiffres, se terminant par 9 et qui soit divisible par 147.

Exercice 5 : Trouver les six derniers chiffres d'un nombre sachant que son cube se termine par 777 777.

Exercice 6 : Montrer que dans un système de numération à base b quelconque, les nombres qui s'écrivent 121, 12 321, 1 234 321, et ainsi de suite jusqu'à épuisement des symboles représentant les chiffres sont des carrés parfaits.

Exercice 7 : Quelle que soit la base du système de numération, montrer qu'aucun des nombres 10 101, 101 010 101, 1 010 101 010 101, etc... n'est premier (Catalan).

Exercice 8 : On considère la numération de base b ; soit ν et m des entiers ≥ 1 et N un entier s'écrivant avec m chiffres. Si ν divise N et $b^m - 1$, alors ν divise tous les nombres obtenus en permutant circulairement les chiffres de N .

Exercice 9 : Soit un naturel A_n qui s'écrit $\overline{x_n x_{n-1} \dots x_2 x_1 x_0}$ ($x_n \neq 0$) en base a . On suppose $x_{n-1} \neq 0$ et on pose $A_{n-1} = \overline{x_{n-1} \dots x_1 x_0}$. Les mêmes symboles représentant dans le système de base b deux autres naturels B_n et B_{n-1} , établir l'équivalence :

$$a > b \Leftrightarrow \frac{A_{n-1}}{A_n} < \frac{B_{n-1}}{B_n}.$$

Exercice 10 : Un naturel u_k s'écrit, en base deux, avec K chiffres 1 consécutifs. Calculer $(u_k)^3$.

Exercice 11 : Montrer l'existence, pour tout entier $x \in \mathbb{Z}$, d'une suite *unique* d'entiers $\in \{-1, 0, 1\}$ tels que

$$x = x_0 + 3x_1 + 3^2x_2 + \dots + 3^kx_k \quad (k \text{ dépendant de } x).$$

On notera $\overline{x_k x_{k-1} \dots x_1 x_0}$ (écriture *base 3 symétrique* de x , -1 étant écrit $\bar{1}$). Comment obtient-on le signe de x , d'après son écriture ? l'opposé de x ?

Chapitre V

GROUPES

Dans ce chapitre nous supposons acquises les notions générales de base données aux § II.4 et III.6, c'est-à-dire : notion de *groupe*, *homomorphismes de groupes*, *sous-groupes*, *noyau et image* d'un homomorphisme, *ordre d'un élément* dans un groupe, notion de *groupe cyclique*. Nous concentrerons notre étude sur les groupes de permutation et la notion de groupe opérant sur un ensemble.

Nous utiliserons les notations suivantes : si G et G' sont deux groupes $\text{Hom}(G, G')$ désigne l'ensemble des *homomorphismes de groupes* de G dans G' ; $\text{Is}(G, G')$ désigne l'ensemble des *isomorphismes de groupes* de G sur G' , s'il en existe ; et $\text{Aut}(G)$ l'ensemble des *automorphismes* du groupe G .

Dans le cas particulier où G' est *abélien*, une façon naturelle de définir une loi de composition interne dans $\text{Hom}(G, G')$ consiste, comme nous le ferions s'il s'agissait des applications d'un ensemble quelconque I dans G' , à poser :

$$fg : G \rightarrow G', \quad x \mapsto (fg)(x) = f(x)g(x) \quad \text{pour tout } x \in G.$$

Il est clair que fg est bien un élément de $\text{Hom}(G, G')$ et que l'opération $(f, g) \mapsto fg$ dote $\text{Hom}(G, G')$ d'une structure de groupe dont l'élément neutre est l'homomorphisme défini par $e(x) = e_{G'}$ pour tout $x \in G$; et l'inverse de $f \in \text{Hom}(G, G')$ est défini par $f^{-1}(x) = f(x^{-1})$ pour tout $x \in G$.

Rappelons enfin que si G est *abélien*, son image par un homomorphisme de groupes de G dans H est un sous-groupe abélien de H et que ce sera sûrement le cas si G est un groupe *monogène* puisque tout groupe monogène est nécessairement abélien.

§ V.1 GÉNÉRATION DE GROUPES

Soit G un groupe, noté multiplicativement. Nous appellerons **séquence** de G toute **suite finie** d'éléments de G , i.e. une applicat

$\llbracket 1, n \rrbracket \rightarrow G, k \mapsto a_k$, où $n \in \mathbb{N}^*$. Le **composé** d'une telle séquence (a_1, a_2, \dots, a_n) est le composé itéré $a_1 a_2 \dots a_n = \prod_{i=1}^n a_i$ défini au § III.1, où

l'ordre des a_i joue un rôle essentiel. On se gardera donc bien d'utiliser la notation $\prod_{\lambda \in \Lambda} a_\lambda$ sauf dans le cas où $(a_\lambda)_{\lambda \in \Lambda}$ est une famille finie (ou à

support fini) d'éléments de G **deux à deux permutables** (cf. § III.1 et III.3). Rappelons que l'existence du symétrique pour tout élément de G entraîne que tout élément de G est *régulier* et aussi que, si a et b sont donnés dans G , chacune des équations (1) $ax = b$ et (2) $xa = b$ à l'inconnue $x \in G$ admet une solution et une seule, à savoir $a^{-1}b$ pour (1) et ba^{-1} pour (2).

L'application $G \rightarrow G, x \mapsto x^{-1}$ est involutive (donc bijective), mais ce n'est *pas* un homomorphisme de groupes, sauf si G est abélien.

En vertu de ce qui précède, le symétrique du composé $a_1 a_2 \dots a_n$ d'une séquence $(a_i)_{1 \leq i \leq n}$ dans G est $a_n^{-1} a_{n-1}^{-1} \dots a_1^{-1}$.

THÉORÈME V.1.1

Soit E une partie non vide du groupe G ; notons E^{-1} l'image de E par l'application $G \rightarrow G, x \mapsto x^{-1}$ ($E^{-1} = \{x^{-1} \mid x \in E\}$). Alors l'ensemble Γ des composés des séquences d'éléments de $E \cup E^{-1}$ est un sous-groupe de G . Ce sous-groupe est l'intersection de tous les sous-groupes de G qui contiennent E ; en d'autres termes, c'est, pour l'inclusion, le plus petit sous-groupe de G contenant E .

Démonstration :

Il est clair que tout sous-groupe de G qui contient E contient nécessairement Γ . Il reste seulement à prouver que Γ est un sous-groupe de G . Or, Γ contient le neutre e_G , car si a est un élément de l'ensemble non vide E , $aa^{-1} = e_G \in \Gamma$; si $x = a_1 a_2 \dots a_n$ et $y = b_1 b_2 \dots b_p$ appartiennent à Γ , avec $a_i \in E \cup E^{-1}$ et $b_j \in E \cup E^{-1}$ pour tous i et j , on a $xy = c_1 c_2 \dots c_{n+p}$ avec $c_i = a_i$ pour $1 \leq i \leq n$ et $c_i = b_{i-n}$ pour $n+1 \leq i \leq n+p$, d'où $xy \in \Gamma$. Enfin si $x = a_1 a_2 \dots a_n$ appartient à Γ ($a_i \in E \cup E^{-1}$ pour tout i), $x^{-1} = a_n^{-1} \dots a_2^{-1} a_1^{-1}$ est lui aussi élément de Γ puisque chaque a_i^{-1} est dans $E \cup E^{-1}$. Finalement, Γ est non vide, stable pour la loi de groupe de G et stable par l'opération $x \mapsto x^{-1}$, donc c'est un sous-groupe de G . ■

DÉFINITION V.1.1

*Avec les notations du théorème V.1.1, le groupe Γ s'appelle **sous-groupe engendré par E dans G** . Nous le noterons $\text{Gr}(E)$. On dit aussi que E **engendre** le sous-groupe $\text{Gr}(E)$, ou est une **partie génératrice** du sous-groupe Γ .*

Remarque 1 : Soit $(a_i)_{i \in I}$ une famille d'éléments de G . Le groupe engendré par l'ensemble $\{a_i\}_{i \in I}$ est appelé **groupe engendré par la famille** $(a_i)_{i \in I}$. Nous le noterons $\text{Gr}((a_i)_{i \in I})$.

Remarque 2 : Avec les notations du théorème V.1.1, $E \cup E^{-1}$ est l'ensemble des x^α pour $x \in E$ et $\alpha \in \{-1, 1\}$. Donc $\text{Gr}(E)$ est l'ensemble de tous les produits $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$, où $n \in \mathbb{N}^*$ et où, pour chaque n , les x_i ($1 \leq i \leq n$) appartiennent à E , et les α_i à $\{-1, 1\}$. En regroupant les facteurs de ces produits par associativité, on voit que $\text{Gr}(E)$ est aussi l'ensemble union de e_G et des éléments $y_1^{\beta_1} y_2^{\beta_2} \dots y_n^{\beta_n}$, où n parcourt \mathbb{N}^* , où les β_i sont dans \mathbb{Z}^* et où les y_i sont dans E , choisis tels que pour tout $i < n$, $y_{i+1} \neq y_i$ et $y_{i+1} \neq y_i^{-1}$.

DÉFINITION V.1.2

Un groupe G est dit de type fini ssi il admet au moins une partie génératrice finie.

Par exemple, si l'ensemble G est fini, il est évident que le groupe G est de type fini, mais bien entendu G peut être de type fini (exemple : G monogène) sans être fini (exemple : $(\mathbb{Z}, +)$). Il existe aussi des groupes de type fini qui ne sont pas monogènes. En voici un exemple simple, à 4 éléments, $G = \{e, a, b, c\}$ avec la loi donnée par la table de multiplication ci-contre, dont le lecteur vérifiera qu'il s'agit bien d'une loi de groupe (c'est le groupe de Klein ⁽¹⁾ noté généralement V_4).

\times	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Groupe produit. Somme directe de groupes abéliens

Soit $(G_i)_{i \in I}$ une famille non vide de groupes, tous notés multiplicativement. Munissons l'ensemble produit $G = \prod_{i \in I} G_i$ de la loi de composition interne ainsi

définie : (3) si $x = (x_i)_{i \in I}$ et $y = (y_i)_{i \in I}$ appartiennent à G , alors $x \cdot y = (x_i y_i)_{i \in I}$. On vérifie immédiatement que G muni de cette loi, devient un groupe.

⁽¹⁾ Félix Klein (1849-1925), mathématicien allemand qui a introduit la théorie des groupes en géométrie.

DÉFINITION V.1.3

Etant donnée une famille non vide $(G_i)_{i \in I}$ de groupes, on appelle **groupe produit** des G_i , et on note $\prod_{i \in I} G_i$, le groupe G obtenu en munissant l'ensemble produit $\prod_{i \in I} G_i$ de la loi interne définie par (3).

Avec ces notations, chaque projection $p_i : G \rightarrow G_i$, $(x_j)_{j \in I} \mapsto x_i$ est un homomorphisme surjectif de groupes, dont le noyau est isomorphe à $\prod_{j \neq i} G_j$ si I n'est pas réduit à un élément.

Lorsque tous les G_i sont égaux à un même groupe Γ , le lecteur aura reconnu, en le groupe produit $\prod_{i \in I} G_i$, le groupe Γ^I de l'exemple 7, § II.4. Revenant au cas général, soit H un groupe. Pour tout homomorphisme de groupes $f : H \rightarrow G$, $p_i \circ f$ est, pour chaque $i \in I$, un homomorphisme de groupes de H dans G_i , donc $(p_i \circ f)_{i \in I}$ est élément de l'ensemble produit $\prod_{i \in I} \text{Hom}(H, G_i)$. Réciproquement, soit $(\varphi_i)_{i \in I}$ un élément de cet ensemble ; pour $x \in H$, posons $\varphi(x) = (\varphi_i(x))_{i \in I}$, $\varphi(x)$ élément de G . Alors $\varphi : H \rightarrow G$ est un homomorphisme de groupes et $\varphi_i = p_i \circ \varphi$ pour tout i . On en déduit aisément :

THÉORÈME V.1.2

Soit $(G_i)_{i \in I}$ une famille non vide de groupes, $G = \prod_{i \in I} G_i$ le groupe produit et $p_i : G \rightarrow G_i$ la i -ième projection ($i \in I$). Alors, pour tout groupe H , l'application $\text{Hom}(H, G) \rightarrow \prod_{i \in I} \text{Hom}(H, G_i)$, $f \mapsto (p_i \circ f)_{i \in I}$ est bijective.

Donnons-nous à présent une famille non vide $(G_i)_{i \in I}$ de groupes **abéliens**, notés additivement. L'élément nul de G_i sera noté 0_i , et les autres notations du théorème V.1.2 seront inchangées. Introduisons les applications

$$\psi_i : G_i \rightarrow G, x \mapsto (u_j)_{j \in I} \text{ tel que } u_i = x, u_j = 0_j \text{ si } j \neq i.$$

On voit que ψ_i est un *homomorphisme injectif de groupes*. Notons $\widehat{G}_i = \text{Im}(\psi_i)$; $\left(\widehat{G}_i \subset G = \prod_{j \in I} G_j \right)$: le groupe engendré dans G par $\bigcup_{i \in I} \widehat{G}_i$ est le sous-groupe des $x = (x_i)_{i \in I}$ éléments de G à *support fini*, c'est-à-dire tel que $\{i \in I \mid x_i \neq 0_i\}$ soit fini.

DÉFINITION V.1.4

Avec les notations ci-dessus, le sous-groupe des $x \in G$ à support fini s'appelle **somme directe (externe)** des groupes G_i , et sera noté $\coprod_{i \in I} G_i$; l'homomorphisme $\widehat{\psi}_i : G_i \rightarrow \coprod_{j \in I} G_j$, $x_i \mapsto \psi_i(x_i)$ est dit **canonique**.

Soit alors H un groupe abélien. Pour tout homomorphisme de groupes $f : \coprod_{i \in I} G_i \rightarrow H$, $(f \circ \widehat{\psi}_i)_{i \in I}$ est un élément de l'ensemble produit $\prod_{i \in I} \text{Hom}(G_i, H)$

Réciproquement, soit $(f_i)_{i \in I}$ un élément de cet ensemble ; si $x = (x_i)_{i \in I}$ est élément de $\prod_{i \in I} G_i$, la famille $(f_i(x_i))_{i \in I}$ est à support fini dans H , donc l'élément $\sum_{i \in I} f_i(x_i) = f(x) \in H$ a un sens (cf. § III.3). On vérifie immédiatement que $f : \prod_{i \in I} G_i \rightarrow H$ est un homomorphisme de groupes, et que $f \circ \widehat{\psi}_i = f_i$ pour tout $i \in I$. De plus, si $f : \prod_{i \in I} G_i \rightarrow H$ est un homomorphisme de groupes, et si $f_i = f \circ \widehat{\psi}_i$ pour $i \in I$, on a : $f((x_i)_{i \in I}) = \sum_{i \in I} f_i(x_i)$ pour tout $(x_i)_{i \in I} \in \prod_{i \in I} G_i$. On a établi :

THÉORÈME V.1.3

Avec les notations de la définition V.1.4, pour tout groupe abélien H , l'application

$$\text{Hom} \left(\prod_{i \in I} G_i, H \right) \rightarrow \prod_{i \in I} \text{Hom} (G_i, H), f \mapsto (f \circ \widehat{\psi}_i)_{i \in I}$$

est bijective.

Avec ces notations, $\text{Hom} \left(\prod_{i \in I} G_i, H \right)$ et tous les $\text{Hom} (G_i, H)$ peuvent être munis de leur structure naturelle de groupe abélien ; munissons alors $\prod_{i \in I} \text{Hom} (G_i, H)$ de la structure de groupe abélien produit de ces groupes $\text{Hom} (G_i, H)$: il est immédiat que **la bijection du théorème V.1.3 devient ainsi un isomorphisme de groupes abéliens.**

Un cas particulier fondamental est celui où tous les groupes G_i sont égaux à un même groupe abélien Γ . Alors le groupe $\prod_{i \in I} G_i$ se note simplement $\Gamma^{(I)}$: c'est le sous-groupe du groupe produit $\Gamma^I = \mathcal{F}(I, \Gamma)$ formé des familles à support fini d'éléments de Γ . Et dans ce cas, le théorème V.1.3 signifie qu'on a un *isomorphisme naturel de groupes abéliens* $\text{Hom} (\Gamma^{(I)}, H) \rightarrow (\text{Hom} (\Gamma, H))^I$, pour tout groupe abélien H , qui envoie f sur $(f \circ \widehat{\psi}_i)_{i \in I}$.

Pour achever, remarquons que lorsque les G_i sont abéliens, si I est fini, il n'y a pas lieu de distinguer les groupes $\prod_{i \in I} G_i$ et $\prod_{i \in I} G_i$, qui sont égaux.

Exemple 1 : A chaque élément $\alpha = (\alpha_n)_{n \in \mathbb{N}^*}$ de $\mathbb{Z}^{(\mathbb{N}^*)}$, associons le rationnel $f(\alpha) = \prod_{n \in \mathbb{N}^*} p_n^{\alpha_n}$, où $(p_n)_{n \geq 1}$ désigne la suite *strictement croissante* des nombres premiers dans \mathbb{N} .

Le théorème d'existence et d'unicité de la décomposition en facteurs premiers dans \mathbb{N} permet facilement de voir que $f : \mathbb{Z}^{(\mathbb{N}^*)} \rightarrow \mathbb{Q}_+^*$ est un *isomorphisme du groupe additif* $\mathbb{Z}^{(\mathbb{N}^*)}$ *sur le groupe multiplicatif* \mathbb{Q}_+^* .

Exercice 1 : Vérifier que le groupe additif \mathbb{Q} n'est pas de type fini. Quels sont ses sous-groupes de type fini ? (Réponse : ce sont les sous-groupes monogènes). Ce résultat subsiste-t-il avec les sous-groupes additifs de \mathbb{R} ?

Exercice 2 : Démontrer, par un raisonnement direct :

- Que le groupe $\mathbb{Z}^{(\mathbb{N})}$ n'est pas de type fini.
- Que le groupe $\mathbb{Z}^{\mathbb{N}}$ n'est pas de type fini.

Exercice 3 : Soit G un groupe ; un élément μ de G est dit *mou* ssi il possède la propriété suivante : pour toute partie génératrice E de G , la partie $E \setminus \{\mu\}$ est encore génératrice. Montrer que l'ensemble \mathcal{M} des éléments mous de G , augmenté de e_G , est un sous-groupe de G , et vérifier que $s(\mathcal{M}) = \mathcal{M}$ pour tout automorphisme s du groupe G . Calculer \mathcal{M} dans les cas suivants :

- $G = (\mathbb{Q}, +)$.
- $G = \mathbb{Z}^{(\mathbb{N})}$.
- $G = \mathbb{Z}$ (réponse : $\mathcal{M} = \{0\}$).
- $G = \mathbb{Z}/n\mathbb{Z}$.

Exercice 4 : Montrer que le groupe multiplicatif \mathbb{Q}^* est isomorphe au groupe produit $\mu_2 \times \mathbb{Z}^{(\mathbb{N}^*)}$, où μ_2 est le groupe multiplicatif formé des nombres -1 et 1 .

Exercice 5 : Montrer que $(\mathbb{Q}, +)$ n'est pas isomorphe à un produit $G_1 \times G_2 \times \cdots \times G_n$, où chaque G_i serait un sous-groupe de \mathbb{Q} non réduit à $\{0\}$.

Exercice 6 : Reprendre la question précédente à partir de $(\mathbb{R}, +)$.

Exercice 7 : Soit G le sous-groupe du groupe $GL(2, \mathbb{R})$ (des matrices carrées réelles inversibles d'ordre 2) engendré par les matrices $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Soit H le sous-ensemble de G formé des matrices $M = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \in G$ telles que $a = b = 1$. Montrer que H est un sous-groupe de G , et que le groupe H n'est pas de type fini. (N.B. Cela prouve qu'il existe des groupes de type fini admettant des sous-groupes qui ne le sont pas.)

Exercice 8 : Soit $SL(2, \mathbb{Z})$ le sous-groupe de $GL(2, \mathbb{R})$ formé des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telles que $(a, b, c, d) \in \mathbb{Z}^4$ et $ad - bc = 1$. On considère les deux matrices suivantes dans $SL(2, \mathbb{Z})$: $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. On note I_2 la matrice $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et on considère le sous-groupe G engendré par $\{S, T\}$ dans $SL(2, \mathbb{Z})$.

- Vérifier que $-I_2 \in G$, et que $U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in G$. Calculer T^λ et U^λ pour $\lambda \in \mathbb{Z}$.
- Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z})$. On note $\mathfrak{M}(M)$ l'ensemble des matrices $\{AMB\}_{A \in G, B \in G}$. Montrer :
 - si $(a = 0 \text{ ou } c = 0)$, alors $I_2 \in \mathfrak{M}(M)$ (si $a = 0$, calculer MS) ;
 - si $a \neq 0$, il existe $M' = \begin{pmatrix} a & b' \\ c' & d' \end{pmatrix} \in \mathfrak{M}(M)$ avec $|c'| < |a|$ (calculer $U^\lambda M$ et utiliser la division euclidienne dans \mathbb{Z}) ;
 - si $c \neq 0$, il existe $M'' = \begin{pmatrix} a'' & b'' \\ c & d'' \end{pmatrix} \in \mathfrak{M}(M)$ avec $|a''| < |c|$.

En déduire que dans tous les cas il existe $M_1 = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathfrak{M}(M)$ telle que $(\alpha = 0 \text{ ou } \gamma = 0)$, puis, que $I_2 \in \mathfrak{M}(M)$.

- Déduire de ce qui précède que $G = SL(2, \mathbb{Z})$, autrement dit que le groupe $SL(2, \mathbb{Z})$ est engendré par $\{S, T\}$.

Exercice 9 : Soit G_1, G_2, \dots, G_n des groupes et H_1, H_2, \dots, H_n des sous-groupes de G_1, G_2, \dots, G_n . Montrer que $H_1 \times H_2 \times \cdots \times H_n$ est un sous-groupe du groupe produit $\prod_{i \in \llbracket 1, n \rrbracket} G_i$.

Exercice 10 : Construire le carré direct du groupe $(\mathbb{Z}/2\mathbb{Z}, +)$ et comparer avec le groupe cité en exemple après la définition IV.1.2. Chercher tous les systèmes possibles de générateurs.

Exercice 11 : Soit G un groupe tel que $x^2 = e_G$ pour tout x . Montrer qu'il

Exercice 12 : Soit p premier et G l'ensemble des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients dans $\mathbb{Z}/p\mathbb{Z}$ qui sont inversibles ($ad - bc \neq \bar{0}$). Montrer que G est un groupe multiplicatif noté $GL(2, \mathbb{Z}/p\mathbb{Z})$. Donner le nombre de ses éléments.

§ V.2 ORDRE D'UN ÉLÉMENT

Nous avons vu au § III.6 qu'à chaque élément de torsion d'un groupe G on associe un entier ≤ 1 , appelé son **ordre** ; par convention, un élément sans torsion de G est dit *d'ordre infini*. Voici quelques propriétés liées à cette notion.

THÉORÈME V.2.1

|| Soit g un élément de torsion d'un groupe G , n l'ordre de g . Si $m \in \mathbb{Z}$, l'élément $h = g^m$ est de torsion, d'ordre n/d , où $d = \text{pgcd}(m, n)$.

Démonstration :

Soit $m' = \frac{m}{d}$ et $n' = \frac{n}{d}$, d'où $\text{pgcd}(m', n') = 1$. On a d'abord :

$$h^{n'} = g^{dm'n'} = (g^{dn'})^{m'} = (g^n)^{m'} = e_G^{m'} = e_G,$$

donc h est de torsion et son ordre ω divise n' . D'autre part, puisque $h^\omega = e_G$, il s'ensuit : $g^{m\omega} = e_G$, c'est-à-dire : $m\omega \equiv 0 \pmod{n}$; ou encore : $m'd\omega \equiv 0 \pmod{n}$, d'où $m'\omega \equiv 0 \pmod{n'}$. Mais comme m' et n' sont premiers entre eux, le théorème de Gauss montre que $\omega \equiv 0 \pmod{n'}$. Finalement $\omega = n'$. ■

THÉORÈME V.2.2

|| Dans un groupe G , soit deux éléments de torsion **permutables** u et v , d'ordres respectifs m et n ; on suppose m et n premiers entre eux. Alors uv est élément de torsion et son ordre est mn .

Démonstration :

Comme u et v sont permutables, on a d'abord

$$(uv)^{mn} = u^{mn} v^{mn} = (u^m)^n (v^n)^m = e_G e_G = e_G,$$

donc uv est élément de torsion, et son ordre ω divise mn . D'autre part, de $(uv)^\omega = e_G$, on déduit : $(uv)^{\omega n} = e_G$, d'où $u^{\omega n} = e_G$ qui entraîne $\omega n \equiv 0 \pmod{m}$. On montre de même que $\omega \equiv 0 \pmod{n}$, à partir de $v^{\omega m} = e_G$. Comme m et n sont premiers entre eux, ω , multiple commun de m et n , est multiple de leur produit mn , et en fin de compte

THÉORÈME V.2.3

Si $f: G \rightarrow G'$ est un homomorphisme de groupes **injectif**, et si $x \in G$ est sans torsion, alors $f(x)$ est sans torsion. Si x est de torsion et d'ordre n , alors $f(x)$ est de torsion et d'ordre n . En particulier l'ordre (fini ou non) d'un élément reste **invariant par isomorphisme de groupes**.

Démonstration :

Pour $m \in \mathbb{Z}$, on a : $f(x^m) = (f(x))^m$; la relation $(f(x))^m = e_{G'}$ équivaut à $f(x^m) = e_{G'}$, donc à $x^m = e_G$ puisque f est injectif.

Donc les sous-groupes de $(\mathbb{Z}, +)$: $\{m \in \mathbb{Z} \mid (f(x))^m = e_{G'}\}$ et $\{m \in \mathbb{Z} \mid x^m = e_G\}$ sont égaux, d'où le théorème, compte tenu de la définition de l'ordre (cf. § III.6). ■

Revenons à l'étude des **groupes monogènes**, qui sont des groupes abéliens particuliers. Un groupe monogène infini G est isomorphe à $(\mathbb{Z}, +)$ et il a donc exactement deux générateurs possibles : si g est l'un d'eux, l'autre est g^{-1} ; et les sous-groupes de G sont très simples : ce sont les groupes $\text{Gr}(g^\alpha)$ où $\alpha \in \mathbb{N}$, qui sont des groupes monogènes infinis si $\alpha \geq 1$.

Soit maintenant un **groupe cyclique** G , c'est-à-dire un groupe monogène fini. Si g est un générateur de G et si g est d'ordre n , alors $\text{card}(G) = n$. Dans ce cas l'application $\llbracket 0, n-1 \rrbracket \rightarrow G, k \mapsto g^k$ est bijective. Si $x = g^k \in G$ ($0 \leq k \leq n-1$), le groupe engendré par x est cyclique et son cardinal est $\frac{n}{d}$, avec $d = \text{pgcd}(k, n)$, d'après le théorème V.2.1. En particulier pour que g^k engendre G , il faut et il suffit que $\text{pgcd}(k, n) = 1$. Or, il y a exactement $\varphi(n)$ entiers $k \in \llbracket 0, n-1 \rrbracket$ tels que $\text{pgcd}(k, n) = 1$, où φ désigne l'indicateur d'Euler défini au § IV.3. Cela signifie qu'il y a, dans G , exactement $\varphi(n)$ générateurs, qui sont les g^k , $0 \leq k \leq n-1$ et $\text{pgcd}(k, n) = 1$.

THÉORÈME V.2.4

Soit G un groupe **cyclique**, de cardinal n , engendré par g .

(I) Le groupe G possède exactement $\varphi(n)$ générateurs (où φ est l'indicateur d'Euler). Ce sont les g^k , où $k \in \llbracket 0, n-1 \rrbracket$ et $\text{pgcd}(k, n) = 1$.

(II) Tout sous-groupe de G est cyclique, de cardinal un diviseur de n .

Démonstration :

L'assertion (I) résulte de l'étude précédente. Il reste à prouver l'assertion (II). Pour cela, introduisons l'homomorphisme de groupes $\varphi_g: \mathbb{Z} \rightarrow G, m \mapsto g^m$, dont on sait qu'il est surjectif et de noyau $n\mathbb{Z}$ (cf. Chap. III). Si H est un sous-groupe de G , $(\varphi_g)^{-1}(H) =$

groupe de $(\mathbb{Z}, +)$, et puisque φ_g est surjectif, $H = \varphi_g(\Gamma)$. Mais on sait (cf. § III.6) que Γ est monogène : $\Gamma = k\mathbb{Z}$ avec $k \in \mathbb{N}$; de plus, $\text{Ker}(\varphi_g) = n\mathbb{Z} \subset \Gamma$, donc k divise n et $H = \varphi_g(k\mathbb{Z}) = \text{Gr}(g^k)$ est monogène, engendré par g^k , donc cyclique de cardinal $\frac{n}{k}$ d'après l'étude précédant l'énoncé du théorème V.2.4. ■

Exercice 1 : Soit G un groupe cyclique de cardinal n , engendré par g . On donne $r \in \mathbb{Z}^*$ et on pose $\delta = r \vee n$. Montrer que l'homomorphisme de groupes $f_r : G \rightarrow G$, $x \mapsto x^r$ admet pour image $\text{Gr}(g^\delta)$ et pour noyau $\text{Gr}(g^{n/\delta})$. Si $y \in G$, calculer $\text{card}(f_r^{-1}(y))$.

Exercice 2 : Soit G un groupe cyclique de cardinal n , engendré par g . Montrer que l'application $\mathcal{D}(n) \rightarrow \mathcal{G}(G)$, $k \mapsto \text{Gr}(g^k)$ est une bijection de l'ensemble $\mathcal{D}(n)$ des diviseurs de n dans \mathbb{N} sur l'ensemble $\mathcal{G}(G)$ des sous-groupes de G .

Exercice 3 : Soit G un groupe cyclique de cardinal n , engendré par g . A chaque r on associe l'endomorphisme f_r de G tel que $f_r(x) = x^r$ pour tout x de G . Montrer que $r \mapsto f_r$ définit une bijection de $\llbracket 0, n-1 \rrbracket$ sur l'ensemble $\text{Hom}(G, G)$. Soit $Y_n : \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ l'application canonique. On munit $\text{Hom}(G, G)$ de sa structure naturelle de groupe. Montrer que $s \mapsto f_{\hat{Y}_n^{-1}(s)}$ est un isomorphisme de $(\mathbb{Z}/n\mathbb{Z}, +)$ sur $\text{Hom}(G, G)$. (On note \hat{Y}_n la restriction de Y_n à $\llbracket 0, n-1 \rrbracket$.)

Exercice 4 : Soit G_1, G_2, \dots, G_n des groupes en nombre fini, $g_i \in G_i$ ($1 \leq i \leq n$) et g l'élément (g_1, \dots, g_n) dans le groupe produit $G_1 \times G_2 \times \dots \times G_n$.

a) Si chaque g_i est de torsion dans G_i , d'ordre ω_i , montrer que g est de torsion dans G , d'ordre $\text{ppcm}(\omega_1, \omega_2, \dots, \omega_n)$.

b) On suppose chaque groupe G_i cyclique. Montrer que G est cyclique si, et seulement si, les entiers $\text{card}(G_i)$ sont deux à deux premiers entre eux.

Exercice 5 :

a) Soit p un nombre premier et $\alpha \in \mathbb{N}^*$. Pour $k \in \mathbb{N}^*$, calculer le nombre d'éléments d'ordre p^α dans le groupe $(\mathbb{Z}/p^\alpha\mathbb{Z})^k$.

b) Soit $n \in \mathbb{N}$, $n \geq 2$. On décompose n en facteurs premiers : $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ (avec les p_i premiers et distincts, les $\alpha_i \geq 1$). Pour $k \in \mathbb{N}^*$, montrer que les groupes additifs $(\mathbb{Z}/n\mathbb{Z})^k$ et $(\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})^k \times \dots \times (\mathbb{Z}/p_r^{\alpha_r}\mathbb{Z})^k$ sont isomorphes. En déduire le nombre d'éléments d'ordre n dans $(\mathbb{Z}/n\mathbb{Z})^k$. (Réponse : c'est $n^k \prod_{i=1}^r \left(1 - \frac{1}{p_i^k}\right)$.)

Exercice 6 : Soit $n \in \mathbb{N}^*$, $n \geq 2$. Sur l'ensemble $D_n = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, on considère la loi de composition suivante : $(a, b) * (a', b') = (a + (-1)^b a', b + b')$.

a) Vérifier que $(D_n, *)$ est un groupe. Combien ce groupe possède-t-il d'éléments d'ordre 2 ?

b) Soit G un groupe de cardinal $2n$ possédant exactement n éléments d'ordre 2. Démontrer que n est impair et que G est isomorphe à D_n .

c) Donner un modèle géométrique simple de D_n , à l'aide d'un polygone régulier à n côtés.

Exercice 7 : Soit G un groupe fini de cardinal N possédant la propriété suivante : (P) pour tout diviseur d de N , le nombre $\psi(d)$ d'éléments de G dont l'ordre divise d est $\leq d$.

On note φ l'indicateur d'Euler et $\mathcal{D}(N)$ l'ensemble des diviseurs de N .

a) Montrer que pour tout $d \in \mathcal{D}(N)$, on a : $\mathcal{S}(d) \leq \varphi(d)$, où $\mathcal{S}(d)$ est le nombre d'éléments d'ordre d dans G .

b) En s'appuyant sur le théorème III.4.1, montrer que $\sum_{d \in \mathcal{D}(N)} \mathcal{S}(d) = N$

c) En utilisant la relation $\sum_{d \in \mathcal{D}(N)} \varphi(d) = N$ (cf. § IV.5, exercice 15), montrer que $\mathcal{S}(d) = \varphi(d)$ pour tout $d \in \mathcal{D}(N)$, et en déduire que G est cyclique.

Exercice 8 : Soit G un groupe et $x \in G$ un élément de torsion d'ordre mn (m et n entiers premiers entre eux). Montrer qu'il existe un et un seul couple $(y, z) \in G \times G$ tel que $x = yz$, $yz = zy$ et : ordre de $y = m$, ordre de $z = n$.

Exercice 9 : Soit x et y deux éléments d'un groupe G . Montrer que xy et yx ont même ordre, fini ou infini.

Indication : pour $\sigma \in G$, l'application $G \rightarrow G$, $u \mapsto \sigma u \sigma^{-1}$ est un automorphisme (dit *intérieur*).

Exercice 10 : Soit x et y deux éléments d'un groupe G . On suppose $yx = x^m y^n$ avec m et n dans \mathbb{Z} . Montrer que les éléments $x^m y^{n-2}$, $x^{m-2} y^n$ et xy^{-1} ont le même ordre, fini ou infini.

Exercice 11 : Dans un groupe fini de cardinal pair, le nombre d'éléments d'ordre 2 est *impair*. Illustrer par des exemples. Former les groupes de cardinal 4.

Exercice 12 : Tout groupe fini de cardinal premier est nécessairement cyclique. Quels sont ses générateurs ?

Exercice 13 : Vérifier que les 16 classes d'entiers, non nulles, mod (17) forment un groupe multiplicatif et que $\overline{3}$ en est un générateur. Former ses sous-groupes.

§ V.3 CLASSES SUIVANT UN SOUS-GROUPE. INDICE

Soit G un groupe (noté multiplicativement), et H un sous-groupe de G . À l'aide de H , on définit sur G deux relations binaires \mathcal{R}_g et \mathcal{R}_d respectivement par :

- (1) $(\forall x \in G)(\forall y \in G) \quad x \mathcal{R}_g y \quad \text{ssi} \quad x^{-1}y \in H$
- (2) $(\forall x \in G)(\forall y \in G) \quad x \mathcal{R}_d y \quad \text{ssi} \quad xy^{-1} \in H.$

Il est immédiat que \mathcal{R}_g et \mathcal{R}_d sont des *relations d'équivalence*, généralement distinctes, mais qui coïncident évidemment lorsque G est abélien.

Les classes mod (\mathcal{R}_g) s'appellent **classes à gauche de G mod (H)** .

Les classes mod (\mathcal{R}_d) s'appellent **classes à droite de G mod (H)** . Les ensembles quotients G/\mathcal{R}_g (resp. G/\mathcal{R}_d) seront notés $(G/H)_g$ (resp. $(G/H)_d$).

Soit $X \in (G/H)_g$. Si $x \in X$, l'ensemble X est $\{xh\}_{h \in H}$, noté xH : c'est l'ensemble image de H par la *translation* $\tau_x : G \rightarrow G$, $u \mapsto xu$, appelée **translation à gauche par x** . De même, si $Y \in (G/H)_d$ et si $y \in Y$, on a $Y = \{hy\}_{h \in H}$, ensemble noté Hy : c'est l'ensemble image de H par la translation $\tau^y : G \rightarrow G$, $v \mapsto vy$, appelée **translation à droite par y** . Ainsi l'application canonique $G \rightarrow (G/H)_g$ (resp. $G \rightarrow (G/H)_d$) est $x \mapsto xH$ (resp. $x \mapsto Hx$). Les translations à droite et à gauche, sont des bijections de G sur lui-même (conséquence immédiate de la structure de groupe).

PROPOSITION V.3.1

|| Avec les notations ci-dessus, toute classe à gauche (resp. à droite) mod (H) est un ensemble équipotent à H .

Dans les mêmes conditions :

PROPOSITION V.3.2

|| Les ensembles $(G/H)_g$ et $(G/H)_d$ sont équipotents.

Démonstration :

Soit I la bijection $G \rightarrow G, x \mapsto x^{-1}$. Pour $x \in G, y \in G$ la relation $x^{-1} y \in H$ entraîne (puisque H est un sous-groupe)

$$y^{-1} x = y^{-1} (x^{-1})^{-1} = I(y)(I(x))^{-1} \in H.$$

On en déduit facilement que si X est une classe à gauche mod (H) , alors son image $I(X)$ est une classe à droite mod (H) . On définit ainsi une application $f : (G/H)_g \rightarrow (G/H)_d$. On définit de la même manière $g : (G/H)_d \rightarrow (G/H)_g$ par $Y \mapsto I(Y)$. Du fait que $I \circ I = \text{Id}_G$, on déduit enfin que $g \circ f = \text{Id}_{(G/H)_g}$ et que $f \circ g = \text{Id}_{(G/H)_d}$, donc f et g sont des bijections réciproques l'une de l'autre. ■

Bien entendu, si G est abélien, les classes à gauche et à droite d'un même élément x coïncident, et *a fortiori* les ensembles $(G/H)_g$ et $(G/H)_d$ sont égaux, mais pour un groupe G non abélien, les classes à gauche et à droite d'un même élément x sont généralement distinctes. S'il arrivait que, pour tout $x \in G$, il y ait coïncidence de la classe à gauche de x et de sa classe à droite, c'est que H serait un sous-groupe particulièrement remarquable. Nous étudierons cette éventualité au § V.7.

DÉFINITION V.3.1

~ Soit H un sous-groupe d'un groupe. Si les ensembles $(G/H)_g$ et $(G/H)_d$ sont infinis, on dit que **H est d'indice infini dans G** , et on note : $[G : H] = \infty$. Lorsque ces ensembles sont finis (et d'après la proposition V.3.2, ils le sont ensemble et quand ils le sont, ils ont le même cardinal), on appelle **indice de H dans G** ⁽¹⁾ l'entier $k \in \mathbb{N}^*$ tel que $k = \text{card}(G/H)_g = \text{card}(G/H)_d$, et on le note $[G : H]$.

On a $[G : H] = 1$ ssi $H = G$; d'autre part si $H = \{e_G\}$, il y a une bijection naturelle évidente entre les ensembles $(G/H)_g$ et G (resp. $(G/H)_d$ et G). Donc dire que le groupe G est fini revient à dire que l'indice de $\{e_G\}$ dans G est fini et si c'est le cas on a : $\text{card}(G) = [G : \{e_G\}]$.

⁽¹⁾ Certains auteurs appellent $[G : H]$ l'index de H dans G .

THÉORÈME V.3.1

Soit G un groupe, et H et K des sous-groupes de G tels que $K \subset H$. Alors les ensembles $(G/H)_g \times (H/K)_g$ et $(G/K)_g$ sont équipotents. En conséquence, pour que $[G : K]$ soit fini, il faut et il suffit que $[G : H]$ et $[H : K]$ le soient, et s'il en est ainsi, on a :

$$(I) \quad [G : K] = [G : H] \times [H : K]$$

(formule de transitivité des indices).

Démonstration :

Soit \mathcal{G} une partie de G et \mathcal{H} une partie de H , telles que $\alpha : \mathcal{G} \rightarrow (G/H)_g, x \mapsto xH$ et $\beta : \mathcal{H} \rightarrow (H/K)_g, y \mapsto yK$ soient des bijections. On définit une application $\varphi : \mathcal{G} \times \mathcal{H} \rightarrow (G/K)_g$ par $(g, h) \mapsto ghK$. Pour prouver le théorème il suffit de prouver que φ est une bijection.

Surjectivité de φ : Soit $l \in G$; alors $lH \in (G/H)_g$. On a donc un $g \in \mathcal{G}$ tel que $lH = gH$; alors $g^{-1}lH = H$, donc $g^{-1}l \in H$. On a donc un $h \in \mathcal{H}$ tel que $g^{-1}lK = hK$. On en déduit $lK = ghK = \varphi(g, h)$.

Injectivité de φ : Si $ghK = g'h'K$ avec $g, g' \in \mathcal{G}$ et $h, h' \in \mathcal{H}$, on en déduit $g'^{-1}ghK = h'K \subset H$. Donc $g'^{-1}gH \cap H$ n'est pas vide puisque $h'K \subset H$ et $h'K \subset g'^{-1}gH$, et puisque $g'^{-1}gH$ et H sont deux classes à gauche mod (H) , nécessairement $g'^{-1}gH = H$, d'où $gH = g'H$, d'où $g = g'$ puisque α est bijective. On en déduit : $hK = h'K$, d'où $h = h'$ puisque β est bijective, et finalement $(g, h) = (g', h')$. ■

Remarque 1 : Avec les notations de cette démonstration, si $x \in G$, l'application $\mathcal{H} \rightarrow (G/K)_g, h \mapsto xhK$ définit une bijection de \mathcal{H} sur l'ensemble des $X \in (G/K)_g$ tels que $X \subset H$.

COROLLAIRE 1

Soit G un groupe fini, et H un sous-groupe de G . Alors :

$$\text{card}(G) = \text{card}(H) \times [G : H].$$

En particulier, **card (H) divise card (G) .**

(Il suffit de prendre $K = \{e_G\}$ dans le théorème V.3.1).

COROLLAIRE 2 (théorème de Lagrange) ⁽¹⁾

Soit G un groupe fini ; pour tout $x \in G$, l'ordre de x divise card (G) . En particulier, on a :

$$x^{\text{card}(G)} = e_G.$$

⁽¹⁾ Joseph Louis, comte de *Lagrange*, mathématicien français né à Turin en 1736. mort à Paris en 1813. Son œuvre la plus remarquable est la création de la mécanique.

Démonstration :

Si $x \in G$, l'ordre de x est le cardinal du sous-groupe $H = \text{Gr}(x)$ de G engendré par x : il suffit alors d'appliquer le corollaire 1. ■

Exemple 1 : Soit n un entier ≥ 2 . Pour $k \in \mathbb{Z}$, notons \bar{k} sa classe mod (n) . Appliquons le théorème de Lagrange au groupe $G(n)$ des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$. On sait que, si $k \in \mathbb{Z}$, on a : $\bar{k} \in G(n)$ ssi k et n sont premiers entre eux. Pour un tel $k \in \mathbb{Z}$, on a donc $(\bar{k})^{\text{card } G(n)} = \bar{1}$, autrement dit $k^{\text{card } G(n)} \equiv 1 \pmod{n}$. Mais nous savons que $\text{card } G(n) = \varphi(n)$ (indicateur d'Euler : cf. § IV.3). Donc, si $k \in \mathbb{Z}$ et si $k \vee n = 1$, $k^{\varphi(n)} \equiv 1 \pmod{n}$, ce qui redémontre le théorème IV.3.5.

En particulier, si n est un nombre premier p , on a :

$$G(p) = (\mathbb{Z}/p\mathbb{Z})^* = (\mathbb{Z}/p\mathbb{Z}) \setminus \{\bar{0}\}.$$

Un entier $k \in \mathbb{Z}$ est premier avec p ssi p ne divise pas k . Pour un tel k , on a : $k^{p-1} \equiv 1 \pmod{p}$ puisque $\text{card}(G(p)) = p - 1$, ce qui redémontre le « petit » théorème de Fermat vu au § IV.4.5.

COROLLAIRE 3 :

$$\left\| \begin{array}{l} \text{Soit } G \text{ un groupe fini, } G' \text{ un groupe, et } f : G \rightarrow G' \text{ un homomorphisme de groupes. Alors :} \\ \text{card}(\text{Im}(f)) \times \text{card}(\text{Ker}(f)) = \text{card}(G). \end{array} \right.$$

Démonstration :

Soit N le groupe $\text{Ker}(f)$. Nous savons que

$$\text{card}(G) = [G : N] \times \text{card}(N).$$

Or, si $y \in \text{Im}(f)$, et si $y = f(x_0)$ avec $x_0 \in G$, on a

$$\begin{aligned} f^{-1}(y) &= \{x \in G \mid f(x) = f(x_0)\} = \{x \in G \mid f(x_0)^{-1} f(x) = e_{G'}\} = \\ &= \{x \in G \mid f(x_0)^{-1} f(x) = e_{G'}\} = \{x \in G \mid f(x_0^{-1}) f(x) = e_{G'}\} \\ &= \{x \in G \mid f(x_0^{-1} x) = e_{G'}\} = x_0 N \in (G/N)_g. \end{aligned}$$

Réciproquement, si $x_0 \in G$, pour tout $x = x_0 z \in x_0 N$ ($z \in N$) on a :

$$f(x) = f(x_0) f(z) = f(x_0),$$

donc f est constante sur la classe à gauche $x_0 N$ de $G \bmod (N)$. Il résulte de tout cela que l'application $\psi : \text{Im}(f) \rightarrow (G/N)_g$, $y \mapsto f^{-1}(y)$ est bijective. Donc

$$\text{card}(\text{Im}(f)) = \text{card}(G/N)_g = [G : N],$$

d'où

$$\begin{aligned} \text{card } (G) &= [G : N] \times \text{card } (N) = [G : N] \times \text{card } (N) = \\ &= \text{card } (\text{Im } (f)) \times \text{card } (N). \quad \blacksquare \end{aligned}$$

Exposant d'un groupe abélien fini

PROPOSITION V.3.3

|| Soit G un groupe **abélien** fini, de cardinal $N > 1$. Notons μ le ppcm des ordres des divers éléments de G . Alors il existe dans G un élément d'ordre μ .

Démonstration :

Décomposons μ en facteurs premiers : $\mu = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ (les p_i premiers et distincts, les $\alpha_i \geq 1$). La formule du § IV.4 donnant le ppcm montre que, pour chaque $i \in \llbracket 1, r \rrbracket$, on peut choisir un $x_i \in G$ d'ordre ω_i tel que $\text{val}_{p_i}(\omega_i) = \alpha_i$; notons $\nu_i = \omega_i / p_i^{\alpha_i}$; d'après le théorème V.2.1, l'élément $y_i = x_i^{\nu_i}$ de G est d'ordre $p_i^{\alpha_i}$. D'après le théorème V.2.2 (qui s'étend par récurrence au cas d'un nombre fini d'éléments deux à deux permutables et d'ordres deux à deux premiers entre eux), l'élément $y = y_1 y_2 \dots y_r$ de G est d'ordre $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} = \mu$ puisque les $(p_i^{\alpha_i})$ sont deux à deux premiers entre eux. ■

DÉFINITION V.3.2

⎧ L'entier μ défini dans la proposition V.3.3 s'appelle l'**exposant** du groupe
⎩ abélien G .

L'exposant μ de G est évidemment un diviseur de $N = \text{card } (G)$ (théorème de Lagrange ci-dessus), strict en général. A cause de la proposition V.3.3, on a : $\mu = N$ ssi G est cyclique. On en déduit :

COROLLAIRE

|| Soit G un groupe abélien fini ; si, pour tout diviseur d de $N = \text{card } (G)$, le nombre d'éléments de G dont l'ordre divise d est $\leq d$, alors G est cyclique.

Démonstration :

Soit μ l'exposant de G ; l'ordre de tout élément de G divise μ , donc d'après l'hypothèse, $\text{card } (G) = N \leq \mu$. Mais d'autre part μ divise N , donc $N = \mu$. Or il y a dans G au moins un élément d'ordre μ (proposition V.3.3), d'où le résultat. ■

Exercice 1 : Soit G un groupe abélien fini, de cardinal N . Si p est un nombre premier qui divise N , montrer qu'il existe dans G un élément d'ordre p . *Indication d'une méthode possible :* soit $G = \{g_1, g_2, \dots, g_N\}$ et ω_i l'ordre de g_i . On peut définir un homomorphisme de groupes naturel surjectif $\mathbb{Z}/\omega_1 \mathbb{Z} \times \dots \times \mathbb{Z}/\omega_N \mathbb{Z} \rightarrow G$ et appliquer le corollaire 3 du théorème V.3.1. Cet exercice montre que si μ est l'exposant de G , il existe un r dans \mathbb{N}^* tel que μ^r soit multiple de N .

Exercice 2 : Soit G un groupe abélien de cardinal $N = p_1 p_2 \dots p_r$, avec les p_i premiers et distincts. Montrer que G est cyclique (utiliser l'exercice 1).

Exercice 3 : Soit G un groupe abélien fini de cardinal $N = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ (

distincts, les $\alpha_i \geq 1$, et $r \geq 1$). Pour tout diviseur d de N , on note S_d l'ensemble des $x \in G$ tels que $x^d = e_G$. On fait l'hypothèse :

(\mathcal{H}) pour tout diviseur d de N , on a $\text{card}(S_d) \leq d$.

a) Soit d un diviseur de N . Déterminer le noyau et l'image de l'homomorphisme $G \rightarrow G, x \mapsto x^d$. En déduire $\text{card}(S_d)$.

b) Pour toute partie I de $\llbracket 1, r \rrbracket$, on note $R_I = \prod_{i \in I} p_i$, avec la convention $R_\emptyset = 1$. Montrer, si $I \subset \llbracket 1, r \rrbracket$:

$$S_{N/R_I} = \bigcap_{i \in I} S_{N/p_i}.$$

c) En appliquant la formule du crible (cf. § III.4, exercice 9), en déduire :

$$\text{card} \left(\bigcup_{i=1}^r S_{N/p_i} \right) = \sum_{k=1}^r (-1)^{k-1} \sum_{\substack{I \subset \llbracket 1, r \rrbracket \\ \text{card}(I) = k}} R_I.$$

En déduire que le nombre d'éléments d'ordre N dans G est $N \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right)$, et enfin que G est cyclique.

N.B. Cet exercice n'utilise pas le résultat de l'exercice 7 du § V.2, ni la proposition V.3.3.

Exercice 4 : Soit X et Y deux sous-groupes d'un groupe G fini de cardinal N . On note XY l'ensemble $\{xy\}_{(x,y) \in X \times Y}$.

a) Montrer que $\text{card}(XY) \times \text{card}(X \cap Y) = \text{card}(X) \text{card}(Y)$.

b) Montrer que XY est un sous-groupe de G ssi $XY = YX$.

Indication pour a). Etudier les images réciproques des $z \in XY$ par l'application $X \times Y \rightarrow XY, (x, y) \mapsto xy$: elles ont toutes le même cardinal, égal à $\text{card}(X \cap Y)$.

Exercice 5 : Soit H et K deux sous-groupes d'un groupe G .

a) Montrer que la relation binaire \mathcal{R} définie sur G par :

$$x \mathcal{R} y \text{ ssi } \exists (h, k) \in H \times K \mid y = h x k$$

est d'équivalence et que, si H et K sont finis, chaque classe mod (\mathcal{R}) est finie, de cardinal diviseur de $\text{card}(H) \text{card}(K)$ (cf. exercice 4 ci-dessus).

b) Si G est fini, en déduire une expression de $\text{card}(G)$ sous la forme

$$\text{card}(G) = \text{card}(H) \text{card}(K) \sum \frac{1}{d_i},$$

où les d_i sont des entiers à préciser.

Exercice 6 : Soit p un nombre premier et m, n deux entiers ≥ 1 . On considère le groupe abélien $G = (\mathbb{Z}/p^2\mathbb{Z})^m \times (\mathbb{Z}/p\mathbb{Z})^n$. Démontrer :

a) le nombre des sous-groupes cycliques de G de cardinal p^2 est

$$p^{m+n-1} \frac{p^m - 1}{p - 1};$$

b) le nombre des sous-groupes non cycliques de G de cardinal p^2 est

$$(p^{m+n} - 1)(p^{m+n-1} - 1)/(p^2 - 1)(p - 1).$$

Exercice 7 : Soit p un nombre premier, et $\alpha_1, \alpha_2, \dots, \alpha_r$ des entiers ≥ 1 . Trouver le nombre d'éléments d'ordre p du groupe

$$\mathbb{Z}/p^{\alpha_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p^{\alpha_r}\mathbb{Z} \quad (\text{Réponse : } p^r - 1)$$

Exercice 8 : Montrer que le sous-groupe $\mathbb{Z}^{(\mathbb{N})}$ de $\mathbb{Z}^{\mathbb{N}}$ n'est pas d'indice fini dans $\mathbb{Z}^{\mathbb{N}}$.

Exercice 9 : Soit H et K deux sous-groupes d'indice fini d'un groupe $[G:H]$ et $[G:K]$ premiers entre eux. Montrer que $G = HK$.

Exercice 10 : Soit G un groupe fini de cardinal N et k un entier premier avec N . Montrer que l'application $\varphi : G \rightarrow G, y \mapsto y^k$ est bijective.

Exercice 11 : Soit H un sous-groupe additif de \mathbb{Q} tel que $\{0\} \neq H \subsetneq \mathbb{Q}$. Montrer que $[\mathbb{Q} : H]$ est infini (en fait on a mieux : le groupe quotient \mathbb{Q}/H (cf. § V.7) n'est pas de type fini).

Exercice 12 : Soit G un groupe abélien de cardinal $N = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}, r \geq 1$ (les p_i premiers distincts, les $\alpha_i \geq 1$). On note G_i le sous-groupe des $x \in G$ tels que $x^{p_i^{\alpha_i}} = e_G$, et P le groupe produit $\prod_{i=1}^r G_i$. Soit $f : P \rightarrow G$ l'homomorphisme $(x_1, \dots, x_r) \mapsto x_1 x_2 \dots x_r$.

a) Montrer que f est surjectif (*indication* : soit $N_i = N/p_i^{\alpha_i}$; utilisant une relation de Bezout entre les N_i , prouver que tout x de G est de la forme $x = x_1^{\lambda_1} \dots x_r^{\lambda_r}$ avec $x_i = x^{N_i}$).

b) Montrer que f est injectif, donc que f est un isomorphisme de groupes.

c) Démontrer que $\text{card}(G_i) = p_i^{\alpha_i}$ (utiliser l'exercice 1).

N.B. On traduit ces propriétés en disant que G est *somme directe interne* des G_i , et l'on note $G = \bigoplus_{i=1}^r G_i$. Les G_i s'appellent les p_i -composants de G .

§ V.4 GROUPES DE PERMUTATIONS

Permutations d'un ensemble

Soit E un ensemble non vide ; nous noterons \mathfrak{S}_E l'ensemble des *bijections* de E sur lui-même : \mathfrak{S}_E est appelé ensemble des **permutations** de E ; c'est un ensemble non vide, car $\text{Id}_E \in \mathfrak{S}_E$. Si $f \in \mathfrak{S}_E$ et $g \in \mathfrak{S}_E$, on a vu au § I.3 que $g \circ f \in \mathfrak{S}_E$, et que $f \langle^{-1}\rangle$ est définie et appartient à \mathfrak{S}_E , et vérifie

$$f \circ f \langle^{-1}\rangle = f \langle^{-1}\rangle \circ f = \text{Id}_E.$$

Puisque la composition d'applications est associative, tout cela prouve que la *loi de composition interne*

$$\mathfrak{S}_E \times \mathfrak{S}_E \rightarrow \mathfrak{S}_E, \quad (f, g) \mapsto f \circ g$$

munit \mathfrak{S}_E d'une structure de groupe. On pose :

DÉFINITION V.4.1

$\left\{ \begin{array}{l} \text{Si } E \text{ est un ensemble non vide, on appelle } \mathbf{groupe des permutations} \\ \mathbf{de } E \text{ le groupe obtenu en munissant l'ensemble } \mathfrak{S}_E \text{ de la loi interne} \\ (f, g) \mapsto f \circ g. \text{ Ce groupe se note } \mathfrak{S}_E, \text{ comme l'ensemble.} \end{array} \right.$

Il importe de remarquer que le groupe \mathfrak{S}_E ne dépend, à isomorphisme près, que du cardinal de E ; on vérifie en effet immédiatement le :

THÉORÈME V.4.1

$\left\| \begin{array}{l} \text{Soit } E, F \text{ deux ensembles non vides équipotents, et } \varphi : E \rightarrow F \text{ une} \\ \text{bijection. L'application } \mathfrak{S}_E \rightarrow \mathfrak{S}_F, \sigma \mapsto \varphi \circ \sigma \circ \varphi \langle^{-1}\rangle \text{ est un iso-} \\ \text{morphisme de groupes.} \end{array} \right.$

D'autre part on a vu au § III.4 que si E est un ensemble fini, de cardinal $n \geq 1$, l'ensemble \mathfrak{S}_E est fini, de cardinal $n!$. Autrement dit :

THÉORÈME V.4.2

|| Si E est un ensemble fini de cardinal $n \geq 1$, le groupe \mathfrak{S}_E est un groupe fini, de cardinal $n!$, et ce groupe ne dépend, à isomorphisme près, que de l'entier n .

(La dernière assertion provient du théorème V.4.1).

C'est en étudiant le groupe des permutations des racines d'une équation algébrique à coefficients dans un corps K commutatif que Galois ⁽¹⁾ a forgé les fondements de la Théorie des groupes (voir [13] et [14]). Cela situe l'importance des groupes de permutations, aussi bien en Algèbre qu'en Géométrie : la plupart des groupes usuels se présentent naturellement comme des sous-groupes de certains groupes de permutations.

Exemple 1 : Soit G un groupe ; tout automorphisme du groupe G est élément de \mathfrak{S}_G . Cela dit, les propriétés vues au § III.4 montrent que l'ensemble, qu'on note $\text{Aut}(G)$, des automorphismes de G forme un sous-groupe du groupe \mathfrak{S}_G . Ce groupe est appelé **groupe des automorphismes de G** et noté encore $\text{Aut}(G)$.

Exemple 2 : Soit A un anneau ; l'ensemble des automorphismes de l'anneau A forme un sous-groupe du groupe \mathfrak{S}_A . Ce groupe revêt une importance particulière lorsque A est un corps commutatif : son étude conduit à la « théorie de Galois » que nous ne pouvons qu'évoquer.

Exemple 3 : (Nécessitant la connaissance de la structure d'espace vectoriel étudiée au Chap. VI.) Soit K un corps commutatif et E un K -espace vectoriel. L'ensemble des bijections K -linéaires de E sur E forme un sous-groupe du groupe \mathfrak{S}_E . Ce groupe noté $\text{GL}_K(E)$ est appelé **groupe linéaire de E** .

De manière générale tout groupe peut être considéré comme un sous-groupe d'un groupe de permutations. En effet, soit G un groupe ; pour $a \in G$, notons τ_a la translation à gauche de G par a , c'est-à-dire la bijection $G \rightarrow G$, $x \mapsto ax$. Pour a et b dans G , on a $\tau_b \circ \tau_a = \tau_{ba}$, donc l'application $a \mapsto \tau_a$ de G dans \mathfrak{S}_G est un homomorphisme de groupes ; le noyau de cet homomorphisme est l'ensemble $\{a \in G \mid \tau_a = \text{Id}_G\}$, soit

$$\{a \in G \mid \forall x \in G \quad ax = x\},$$

c'est-à-dire $\{e_G\}$, donc c'est un homomorphisme injectif. On a donc :

⁽¹⁾ Galois (Evariste). Mathématicien français (25-10-1811†30-5-1832), un des plus purs génies de la science. Avec Abel, a débrouillé la question de la résolution des équations algébriques, entre autres. Mort à la fleur de l'âge dans des circonstances plus que louches après avoir été forcé, par provocation politique, de se battre en un duel où il n'avait

THÉORÈME V.4.3 (Cayley ⁽¹⁾)

Si G est un groupe, l'homomorphisme de groupes $G \rightarrow \mathfrak{S}_G$, $a \mapsto \tau_a$ définit un isomorphisme de G sur un sous-groupe de \mathfrak{S}_G .

Permutations d'un ensemble fini

Si $n \in \mathbb{N}^*$, on note \mathfrak{S}_n le groupe $\mathfrak{S}_{\llbracket 1, n \rrbracket}$, et on l'appelle **groupe symétrique d'ordre n** . D'après le théorème V.4.1, il est isomorphe à tout groupe \mathfrak{S}_E , où $\text{card}(E) = n$, et d'après le théorème V.4.2, c'est un groupe fini, de cardinal $n!$.

Pour $n = 1$, \mathfrak{S}_1 est réduit à un élément. Nous allons donc étudier le groupe \mathfrak{S}_n pour $n \geq 2$. Puisque par définition $\mathfrak{S}_n = \mathfrak{S}_{\llbracket 1, n \rrbracket}$, nous profiterons du fait que l'ensemble $\llbracket 1, n \rrbracket$ peut être totalement ordonné par la relation d'ordre induite par l'ordre naturel sur \mathbb{N} pour donner à chaque permutation σ , élément de \mathfrak{S}_n , une écriture standardisée : on écrit les entiers de 1 à n dans l'ordre naturel sur une première ligne avec l'image $\sigma(i)$ de i ($1 \leq i \leq n$) sur une seconde ligne juste en face de i , d'où la notation $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$. Fixons donc $n \geq 2$. A chaque permutation σ de \mathfrak{S}_n nous allons associer sa *signature* qui en est une caractéristique essentielle. Pour faciliter l'étude de cette notion centrale, nous noterons : P l'ensemble des parties à deux éléments de $\llbracket 1, n \rrbracket$, et \mathcal{C} l'ensemble des couples $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket$ tels que $i < j$. Il est clair que l'application $\chi : \mathcal{C} \rightarrow P$, $(i, j) \mapsto \{i, j\}$ est bijective.

DÉFINITION V.4.2

Pour toute permutation $\sigma \in \mathfrak{S}_n$, on appelle **nombre d'inversions** de σ le nombre, noté $I(\sigma)$, des couples $(i, j) \in \mathcal{C}$ tels que $\sigma(i) > \sigma(j)$, c'est-à-dire :

$$I(\sigma) = \text{card} \{ (i, j) \in \llbracket 1, n \rrbracket^2 \mid i < j \text{ et } \sigma(i) > \sigma(j) \} .$$

On appelle **signature** de σ l'entier relatif élément de $\{-1, +1\}$, noté $\varepsilon\{\sigma\}$ et égal à $(-1)^{I(\sigma)}$.

Remarquons que $I(\text{Id}_{\llbracket 1, n \rrbracket}) = 0$ et par conséquent $\varepsilon(\text{Id}_{\llbracket 1, n \rrbracket}) = +1$. (Pour $n = 1$, on convient que la signature de l'unique élément de \mathfrak{S}_n est aussi $+1$.)

Considérons maintenant l'ensemble \mathcal{F} des fonctions de \mathbb{Q}^n dans \mathbb{Q} . Si $f \in \mathcal{F}$ et $\lambda \in \mathbb{Q}$, on note λf l'élément de \mathcal{F} tel que

$$(\lambda f)(x_1, x_2, \dots, x_n) = \lambda f(x_1, x_2, \dots, x_n)$$

pour tous $x_i \in \mathbb{Q}$.

⁽¹⁾ Arthur Cayley (1821-1895), avocat et mathématicien anglais, est l'un des inventeurs du calcul matriciel.

Si $f \in \mathcal{F}$ et $\sigma \in \mathfrak{S}_n$, on définit une nouvelle application g , que nous noterons $\sigma * f$, de \mathbb{Q}^n dans \mathbb{Q} par :

$$g(x_1, x_2, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

pour tous $x_i \in \mathbb{Q}$.

Il est bien clair que $\sigma * (\lambda f) = \lambda (\sigma * f)$ pour tous $\lambda \in \mathbb{Q}$ et $\sigma \in \mathfrak{S}_n$, et que $\text{Id} * f = f$ pour tout $f \in \mathcal{F}$.

Pour $n \geq 2$, la fonction fondamentale permettant une étude de la signature est $\Delta \in \mathcal{F}$ définie par

$$\Delta(x_1, x_2, \dots, x_n) = \prod_{(i,j) \in \mathcal{C}} (x_j - x_i) \quad ((x_i) \in \mathbb{Q}^n).$$

THÉORÈME V.4.4

$$\left\| \begin{array}{l} \text{Pour } \sigma \in \mathfrak{S}_n \text{ (avec } n \geq 2), \text{ on a :} \\ \sigma * \Delta = \varepsilon(\sigma) \Delta. \end{array} \right.$$

Démonstration :

Pour $(x_1, x_2, \dots, x_n) \in \mathbb{Q}^n$, on a :

$$(\sigma * \Delta)(x_1, \dots, x_n) = \prod_{(i,j) \in \mathcal{C}} (x_{\sigma(j)} - x_{\sigma(i)}).$$

Pour chaque paire $u = \{\alpha, \beta\} \in P$, soit

$$D_u : \mathbb{Q}^n \rightarrow \mathbb{Q}, \quad (x_1, \dots, x_n) \mapsto x_\gamma - x_\delta,$$

où $\gamma = \text{Max } \{\alpha, \beta\}$ et $\delta = \text{Min } \{\alpha, \beta\}$. La définition du nombre d'inversions $I(\sigma)$ montre que

$$(1) \quad (\sigma * \Delta)(x_1, \dots, x_n) = \varepsilon(\sigma) \prod_{(i,j) \in \mathcal{C}} D_{\{\sigma(i), \sigma(j)\}}(x_1, \dots, x_n).$$

Mais l'application $\mathcal{C} \rightarrow P, (i, j) \mapsto \{\sigma(i), \sigma(j)\}$ est bijective, car elle est injective et les cardinaux de \mathcal{C} et P sont égaux. Donc (cf. § III.1) :

$$\begin{aligned} \prod_{(i,j) \in \mathcal{C}} D_{\{\sigma(i), \sigma(j)\}}(x_1, \dots, x_n) &= \prod_{u \in P} D_u(x_1, \dots, x_n) = \\ &= \prod_{(i,j) \in \mathcal{C}} D_{\{i,j\}}(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n). \end{aligned}$$

La relation (1) fournit bien $\sigma * \Delta = \varepsilon(\sigma) \Delta$. ■

LEMME 1

$$\left\| \begin{array}{l} \text{Pour } \sigma \in \mathfrak{S}_n, \tau \in \mathfrak{S}_n \text{ et } f \in \mathcal{F}, \text{ on a :} \\ \sigma * (\tau * f) = (\sigma\tau) * f. \end{array} \right.$$

Démonstration :

Pour $(x_1, x_2, \dots, x_n) \in \mathbb{Q}^n$, on a :

$$\begin{aligned} [(\sigma\tau) * f](x_1, x_2, \dots, x_n) &= f(x_{\sigma\tau(1)}, x_{\sigma\tau(2)}, \dots, x_{\sigma\tau(n)}) \\ (\tau * f)(x_1, x_2, \dots, x_n) &= f(x_{\tau(1)}, x_{\tau(2)}, \dots, x_{\tau(n)}) \\ [\sigma * (\tau * f)](x_1, x_2, \dots, x_n) &= (\tau * f)(y_1, y_2, \dots, y_n), \quad \text{où } y_i = x_{\sigma(i)}. \end{aligned}$$

Donc

$$\begin{aligned} [\sigma * (\tau * f)](x_1, \dots, x_n) &= f(y_{\tau(1)}, \dots, y_{\tau(n)}) = f(x_{\sigma\tau(1)}, \dots, x_{\sigma\tau(n)}) \\ &= [(\sigma\tau) * f](x_1, \dots, x_n). \quad \blacksquare \end{aligned}$$

L'ensemble $\mathcal{U}_{\mathbb{Z}} = \{-1, +1\}$, des éléments inversibles de \mathbb{Z} est un groupe pour la multiplication. De ce qui précède, on déduit :

THÉORÈME V.4.5

|| Pour $\sigma \in \mathfrak{S}_n$ et $\tau \in \mathfrak{S}_n$, on a :

$$\boxed{\varepsilon(\sigma\tau) = \varepsilon(\sigma) \varepsilon(\tau)}.$$

|| En d'autres termes, la signature ε est un homomorphisme de groupes de \mathfrak{S}_n dans $\mathcal{U}_{\mathbb{Z}}$.

Démonstration :

Soit σ et τ dans \mathfrak{S}_n . En utilisant le théorème V.4.4 et le lemme 1, on a :

$$\begin{aligned} \sigma * (\tau * \Delta) &= \sigma * (\varepsilon(\tau) \Delta) = \varepsilon(\tau)(\sigma * \Delta) = \\ &= \varepsilon(\tau)(\varepsilon(\sigma) \Delta) = \varepsilon(\sigma) \varepsilon(\tau) \Delta \end{aligned}$$

d'une part, et $\sigma * (\tau * \Delta) = (\sigma\tau) * \Delta = \varepsilon(\sigma\tau) \Delta$

d'autre part. Or, la valeur de Δ sur la suite $(1, 2, \dots, n)$ est $\prod_{i < j} (j - i) \neq 0$,

ce qui permet de simplifier par $\Delta(1, 2, \dots, n)$ la relation

$$\varepsilon(\sigma) \varepsilon(\tau) \Delta(1, 2, \dots, n) = \varepsilon(\sigma\tau) \Delta(1, 2, \dots, n),$$

d'où le théorème. \blacksquare

Le théorème V.4.5 reste vrai pour $n = 1$ avec la convention faite précédemment.

DÉFINITION V.4.3

⌋ $(n$ étant toujours supposé $\geq 2)$, on appelle **transposition** dans \mathfrak{S}_n tout élément τ de \mathfrak{S}_n du type suivant :

$$\tau(i) = j, \quad \tau(j) = i, \quad \tau(k) = k \quad \text{pour } k \notin \{i, j\},$$

⌋ avec i, j dans $\llbracket 1, n \rrbracket$ et $i \neq j$.

Il est clair que toute transposition de \mathfrak{S}_n est involutive ($\tau \circ \tau = \text{Id}$). De plus le nombre d'inversions de la transposition τ qui vient d'être définie, admet, pour i et j fixés dans $\llbracket 1, n \rrbracket$, exactement $1 + 2(j - i - 1)$ inversions, donc $\varepsilon(\tau) = -1$. Or, pour $n \geq 2$, il existe au moins une transposition, donc l'homomorphisme de groupes ε de \mathfrak{S}_n dans $\{-1, +1\}$ est *surjectif*. Le noyau $\text{Ker}(\varepsilon)$ de cet homomorphisme est un sous-groupe de \mathfrak{S}_n . D'après l'étude faite au § V.3, ce sous-groupe est d'indice 2, de cardinal $\frac{1}{2} \text{card}(\mathfrak{S}_n) = \frac{1}{2} n!$, d'où, en résumé :

THÉORÈME V.4.6

|| Pour $n \in \mathbb{N}$, $n \geq 2$, l'homomorphisme signature $\varepsilon : \mathfrak{S}_n \rightarrow \mathcal{U}_{\mathbb{Z}}$ est surjectif; son noyau $\text{Ker}(\varepsilon)$ est un sous-groupe de cardinal $\frac{1}{2} n!$ de \mathfrak{S}_n .

DÉFINITION V.4.4

~ Avec les notations du théorème V.4.6, le noyau $\text{Ker}(\varepsilon) = \{\sigma \in \mathfrak{S}_n \mid \varepsilon(\sigma) = +1\}$ est appelé **groupe alterné** d'ordre n , et se note \mathcal{U}_n . Les éléments de \mathcal{U}_n s'appellent **permutations paires**; ceux de $\mathfrak{S}_n \setminus \mathcal{U}_n$ s'appellent **permutations impaires**.

Il résulte de cette définition que toute transposition dans \mathfrak{S}_n est une permutation impaire. D'autre part il est clair que les ensembles de classes $(\mathfrak{S}_n/\mathcal{U}_n)_g$ et $(\mathfrak{S}_n/\mathcal{U}_n)_d$ coïncident tous deux avec $\{\mathcal{U}_n, \mathfrak{S}_n \setminus \mathcal{U}_n\}$.

Revenons à présent au groupe \mathfrak{S}_E , lorsque E est un ensemble fini de cardinal $n \geq 1$.

LEMME 2

|| Soit $\varphi, \psi : \llbracket 1, n \rrbracket \rightarrow E$ deux bijections; pour tout $\sigma \in \mathfrak{S}_E$, on a :

$$\varepsilon(\varphi \langle^{-1}\rangle \circ \sigma \circ \varphi) = \varepsilon(\psi \langle^{-1}\rangle \circ \sigma \circ \psi).$$

Démonstration :

En effet, soit $u = \varphi \langle^{-1}\rangle \circ \sigma \circ \varphi$, $v = \psi \langle^{-1}\rangle \circ \sigma \circ \psi$, et $\alpha = \psi \langle^{-1}\rangle \circ \varphi$, d'où $\alpha \in \mathfrak{S}_n$. On a : $u = \alpha \langle^{-1}\rangle \circ v \circ \alpha$, d'où grâce au théorème V.4.5 :

$$\begin{aligned} \varepsilon(u) &= \varepsilon(\alpha \langle^{-1}\rangle) \varepsilon(v) \varepsilon(\alpha) = \varepsilon(\alpha \langle^{-1}\rangle) \varepsilon(\alpha) \varepsilon(v) = \\ &= \varepsilon(\alpha \langle^{-1}\rangle \circ \alpha) \varepsilon(v) = \varepsilon(\text{Id}) \varepsilon(v) = \varepsilon(v). \quad \blacksquare \end{aligned}$$

On peut donc transporter sur \mathfrak{S}_E la notion de signature :

DÉFINITION V.4.5

~ Si E est un ensemble **fini**, de cardinal $n \geq 1$, pour $\sigma \in \mathfrak{S}_E$, on appelle **signature** de σ , et on note $\varepsilon(\sigma)$, l'élément ν de $\mathcal{U}_{\mathbb{Z}}$ tel que, pour toute bijection φ de $\llbracket 1, n \rrbracket$ dans E , on ait

$$\nu = \varepsilon(\varphi \langle^{-1}\rangle \circ \sigma \circ \varphi).$$

Les propriétés de la signature ainsi définie sur \mathfrak{S}_E se déduisent de celles de la signature définie sur \mathfrak{S}_n , en choisissant une bijection particulière $\varphi : \llbracket 1, n \rrbracket \rightarrow E$ et en utilisant l'isomorphisme de groupes $J_\varphi : \mathfrak{S}_n \rightarrow \mathfrak{S}_E$, $\sigma \mapsto \varphi \circ \sigma \circ \varphi^{-1}$ (cf. théorème V.4.1). Ainsi : $\varepsilon : \mathfrak{S}_E \rightarrow \mathcal{U}_{\mathbb{Z}}$ est un homomorphisme de groupes, surjectif si $n \geq 2$; son noyau $\text{Ker}(\varepsilon)$ (appelé groupe des **permutations paires** de E) est, pour $n \geq 2$, un sous-groupe de cardinal $\frac{1}{2} n!$ de \mathfrak{S}_E , noté \mathfrak{U}_E et appelé **groupe alterné** de E . Les éléments de $\mathfrak{S}_E \setminus \mathfrak{U}_E$ (non vide pour $n \geq 2$) sont appelés **permutations impaires** de E . L'isomorphisme J_φ ci-dessus transforme \mathfrak{U}_n en \mathfrak{U}_E et donc $\mathfrak{S}_n \setminus \mathfrak{U}_n$ en $\mathfrak{S}_E \setminus \mathfrak{U}_E$.

THÉORÈME V.4.7

|| Soit n un entier ≥ 2 . L'ensemble des transpositions de \mathfrak{S}_n engendre le groupe \mathfrak{S}_n .

Démonstration :

Par récurrence sur n en commençant par le cas évident $n = 2$. Supposons-le vrai à l'ordre $n - 1 \geq 2$, et montrons-le à l'ordre n . Soit $\sigma \in \mathfrak{S}_n$; si $\sigma(n) = n$, posons $s = \sigma \llbracket \llbracket 1, n-1 \rrbracket \rrbracket$, d'où $s \in \mathfrak{S}_{n-1}$; l'hypothèse de récurrence permet d'écrire $s = t_1 t_2 \dots t_k$, composé d'un nombre fini de transpositions t_i de \mathfrak{S}_{n-1} . Soit $\tau_i \in \mathfrak{S}_n$ tel que $\tau_i(n) = n$ et $\tau_i \llbracket \llbracket 1, n-1 \rrbracket \rrbracket = t_i$, alors τ_i est une transposition de \mathfrak{S}_n , et on a bien $\sigma = \tau_1 \tau_2 \dots \tau_k$. Il reste à étudier le cas où $\sigma(n) = \alpha < n$. Soit τ la transposition de \mathfrak{S}_n telle que $\tau(\alpha) = n$, $\tau(n) = \alpha$ et $\tau(j) = j$ si $j \notin \{\alpha, n\}$. Alors $\tau\sigma(n) = n$. Donc on peut écrire $\tau\sigma$ sous la forme $\tau_1 \tau_2 \dots \tau_k$ où les τ_i sont des transpositions de \mathfrak{S}_n ; d'où finalement

$$\sigma = \tau^{-1} \tau_1 \tau_2 \dots \tau_k = \tau \tau_1 \tau_2 \dots \tau_k. \quad \blacksquare$$

Bien entendu, il n'y a pas unicité de cette décomposition de σ en transpositions ($\sigma = \sigma \text{Id} = \sigma \tau' \tau'$), mais en revanche la **parité** du nombre N de transpositions qui intervient dans une décomposition de σ est parfaitement déterminée à cause du théorème V.4.5 qui s'étend par récurrence à un nombre fini de facteurs et donne $\varepsilon(\sigma) = (-1)^N$.

Soit E un ensemble fini non vide, et $\sigma \in \mathfrak{S}_E$. Si S est une partie non vide de E telle que $\sigma(S) \subset S$ (on dit alors que S est **σ -stable**), alors $\sigma(S) = S$ puisque σ est injective et que S est fini, et l'application $\sigma \llbracket_S : S \rightarrow S, x \mapsto \sigma(x)$ est une **permutation de S** , dite **induite par σ sur S** .

THÉORÈME V.4.8

|| Soit E un ensemble fini, de cardinal $n \geq 2$, σ une permutation de E , S et T deux parties non vides de E **complémentaires** l'une de l'autre, et **σ -stables**. Désignons par α et β les permutations induites par σ sur S et T . Alors : $\varepsilon(\sigma) = \varepsilon(\alpha) \varepsilon(\beta)$.

Démonstration :

Notons $p = \text{card}(S)$ et $q = \text{card}(T)$, d'où $p \geq 1$, $q \geq 1$ et $p + q = n$. Choisissons une bijection $\varphi : \llbracket 1, n \rrbracket \rightarrow E$ telle que $\varphi(\llbracket 1, p \rrbracket) = S$, $\varphi(\llbracket p+1, n \rrbracket) = T$, et posons $s = \varphi^{\langle -1 \rangle} \circ \sigma \circ \varphi$, $a = \varphi^{\langle -1 \rangle} \circ \alpha \circ \varphi$, $b = \varphi^{\langle -1 \rangle} \circ \beta \circ \varphi$. On a, par définition, $\varepsilon(\sigma) = \varepsilon(s)$, $\varepsilon(\alpha) = \varepsilon(a)$, $\varepsilon(\beta) = \varepsilon(b)$. D'autre part, $\llbracket 1, p \rrbracket$ et $\llbracket p+1, n \rrbracket$ sont s -stables, et s induit respectivement les permutations a et b sur ces deux ensembles. Il est immédiat que les nombres d'inversions de ces permutations vérifient $I(s) = I(a) + I(b)$. D'où

$$\varepsilon(s) = (-1)^{I(s)} = (-1)^{I(a)}(-1)^{I(b)} = \varepsilon(a) \varepsilon(b). \quad \blacksquare$$

Exercice 1 : Trouver tous les automorphismes du groupe \mathfrak{S}_3 et en donner une illustration géométrique.

Exercice 2 : Trouver le centre de \mathfrak{S}_n , c'est-à-dire le sous-groupe formé dans \mathfrak{S}_n par les éléments permutable avec tous les autres. Trouver le centre de \mathcal{U}_n .

Exercice 3 : Montrer que, si $n \geq 2$, le groupe \mathfrak{S}_n est engendré par les transpositions τ_{1i} , $2 \leq i \leq n$; ($\tau_{1i}(1) = i$, $\tau_{1i}(i) = 1$, $\tau_{1i}(k) = k$ pour $k \neq 1$, $k \neq i$). Montrer aussi que \mathfrak{S}_n est engendré par les transpositions $\tau_{i,i+1}$, $1 \leq i \leq n-1$. Peut-on trouver des sous-ensembles stricts de ces ensembles qui engendrent encore \mathfrak{S}_n ?

Exercice 4 : Soit $n \geq 2$. Montrer que \mathfrak{S}_n est engendré par les deux éléments suivants : τ_{12} définie à l'exercice 3, et $c_n = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 2 & 3 & \dots & n & 1 \end{pmatrix}$.

Exercice 5 : Dans un groupe G , on appelle **groupe des commutateurs de G** le sous-groupe de G engendré par les éléments $xyx^{-1}y^{-1}$, où (x, y) parcourt $G \times G$. Ce sous-groupe se note $[G, G]$; si $x, y \in G$, l'élément $xyx^{-1}y^{-1}$ s'appelle *commutateur* de x et y et se note $[x, y]$. Si $n \geq 3$, déterminer les groupes $[\mathfrak{S}_n, \mathfrak{S}_n]$ et $[\mathcal{U}_n, \mathcal{U}_n]$. Pour $n = 4$ on pourra se reporter au § V.7.

Exercice 6 : A isomorphisme près, montrer qu'il n'y a que deux groupes de cardinal 6, à savoir $(\mathbb{Z}/6\mathbb{Z}, +)$ et \mathfrak{S}_3 .

Exercice 7 : Soit $n \geq 3$. Peut-on trouver un ensemble de $n-2$ transpositions qui engendrent le groupe \mathfrak{S}_n ?

Exercice 8 : Montrer que \mathcal{U}_5 est engendré par les deux éléments

$$c = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix} \quad \text{et} \quad \gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}.$$

Exercice 9 : Soit $n \geq 2$ et c_n la permutation $\begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 2 & 3 & \dots & n & 1 \end{pmatrix}$. Montrer que si c_n est produit de k transpositions, on a $k \geq n-1$ (variante plus fine de l'exercice 7).

Exercice 10 : Soit $\sigma \in \mathfrak{S}_n$ et $\tau \in \mathfrak{S}_n$, $n \geq 2$. Compter directement le nombre exact d'inversions de $\sigma \circ \tau$ en fonction de ceux de σ et de τ .

Exercice 11 : Soit $E = \mathbb{C} \cup \{\infty\}$ et G le groupe des bijections de E (pour la composition des applications) engendré par $\varphi : x \mapsto \frac{1}{x}$ et $\psi : x \mapsto \frac{1}{1-x}$. Montrer que G

\mathfrak{S}_3 . Comparer avec les valeurs prises par le birapport $\rho = \frac{a_3 - a_1}{a_3 - a_2} : \frac{a_4 - a_1}{a_4 - a_2}$ quand on permute les variables a_i .

Pour les exercices suivants, on adoptera la notation suivante : soit $k \mapsto a_k$ une bijection de $\llbracket 1, p \rrbracket$ sur un sous-ensemble F ($\text{card}(F) = p \geq k$) de l'ensemble fini E . On notera $\langle a_1, a_2, \dots, a_k \rangle$ la permutation c de E définie par $c(a_1) = a_2, c(a_2) = a_3, \dots, c(a_{k-1}) = a_k, c(a_k) = a_1$ et $c(x) = x$ pour $x \in E \setminus F$.

Exercice 12 : Si $n \geq 2$, trouver les permutations permutables avec

$$c_n = \langle 1, 2, \dots, n \rangle \quad (c_n \in \mathfrak{S}_n).$$

Exercice 13 : Soit $s \in \mathfrak{S}_{10}$ la permutation $u \circ v$, où

$$u = \langle 1, 2, 3, 4, 5 \rangle, \quad v = \langle 6, 7, 8, 9, 10 \rangle.$$

Trouver les permutations $\sigma \in \mathfrak{S}_{10}$ telles que $\sigma \circ s = s \circ \sigma$. Montrer qu'il y en a 50 et qu'elles forment un sous-groupe de \mathfrak{S}_{10} .

Exercice 14 : Soit $n \geq 3$ et k donnés, $1 < k < n$. Trouver $uvu^{-1}v^{-1}$ pour $u = \langle 1, 2, \dots, k \rangle$ et $v = \langle k+1, \dots, n \rangle$.

Exercice 15 :

a) Démontrer, pour $n \geq 1$:

$$\begin{aligned} \langle a_1, a_2, \dots, a_{2n} \rangle &= \langle a_1, a_{2n-1} \rangle \circ \langle a_2, a_{2n-2} \rangle \circ \dots \circ \\ &\quad \circ \langle a_{n-1}, a_{n+1} \rangle \circ \langle a_1, a_{2n} \rangle \circ \dots \circ \langle a_n, a_{n+1} \rangle \\ \langle a_1, a_2, \dots, a_{2n+1} \rangle &= \langle a_1, a_{2n} \rangle \circ \langle a_2, a_{2n-1} \rangle \circ \dots \circ \\ &\quad \circ \langle a_{n-1}, a_{n+1} \rangle \circ \langle a_1, a_{2n+1} \rangle \circ \langle a_2, a_{2n} \rangle \circ \dots \circ \langle a_n, a_{n+2} \rangle. \end{aligned}$$

b) En déduire, pour un ensemble fini E , que tout $\sigma \in \mathfrak{S}_E$ peut s'écrire $\sigma = uv$ avec $u^2 = v^2 = \text{Id}_E$.

Exercice 16 : Soit $N = pq + 1$ avec $p \geq 2, p \in \mathbb{N}^*, q \in \mathbb{N}^*$. On considère les permutations

$$\begin{aligned} s_1 &= \langle 1, 2, \dots, q+1 \rangle, \quad s_2 = \langle 1, q+2, \dots, 2q+1 \rangle, \dots, \\ s_p &= \langle 1, (p-1)q+2, \dots, pq+1 \rangle. \end{aligned}$$

Démontrer :

- a) Si N est impair, $\{s_1, \dots, s_p\}$ engendrent le groupe \mathfrak{S}_N .
- b) Si N est pair, $\{s_1, \dots, s_p\}$ engendrent le groupe \mathfrak{U}_N .

Exercice 17 : Dans \mathfrak{S}_{12} , soit $u = \langle 1, 2, 3, 4, 5, 6 \rangle$ et $v = \langle 7, 8, 9, 10, 11, 12 \rangle$, et $s = uv$. Montrer que l'ensemble Γ des $\sigma \in \mathfrak{S}_{12}$ telles que $\sigma s = s \sigma$ forme un sous-groupe de cardinal 72 de \mathfrak{S}_{12} .

§ V.5 CYCLES DANS LES GROUPE \mathfrak{S}_E (E fini)

Eléments conjugués dans un groupe

Soit G un groupe, noté multiplicativement. Pour $\sigma \in G$, notons

$$f_\sigma : G \rightarrow G, x \mapsto \sigma x \sigma^{-1}.$$

On voit immédiatement que f_σ est un *endomorphisme* du groupe G , et que, si $\sigma \in G$ et $\tau \in G$, alors

$$(1) \quad f_{\sigma\tau} = f_\sigma \circ f_\tau.$$

En particulier

$$f_\sigma \circ f_{\sigma^{-1}} = f_{\sigma^{-1}} \circ f_\sigma = f_{e_G} = \text{Id}_G$$

et par suite f_σ est un **automorphisme** du groupe G , et

$$(2) \quad f_\sigma^{\langle -1 \rangle} = f_{\sigma^{-1}}.$$

DÉFINITION V.5.1

*On appelle **automorphisme intérieur** d'un groupe G tout automorphisme g de G du type $g = f_\sigma$ pour au moins un $\sigma \in G$, où $f_\sigma : G \rightarrow G$ est l'application $x \mapsto \sigma x \sigma^{-1}$.*

Notons que, d'après (1), l'application $\mathcal{F} : G \rightarrow \text{Aut}(G)$, $\sigma \mapsto f_\sigma$ (où $\text{Aut}(G)$ désigne le groupe des automorphismes de G défini dans l'exemple 1 du § V.4) est un homomorphisme de groupes. Nous l'étudierons plus loin (exemple 5 du § V.7).

DÉFINITION V.5.2

*Soit G un groupe ; deux éléments x et y de G (resp. deux sous-groupes H et K de G) sont dits **conjugués dans G** ssi il existe $\sigma \in G$ tel que $y = \sigma x \sigma^{-1}$ (resp. tel que $K = \sigma H \sigma^{-1}$, où $\sigma H \sigma^{-1}$ désigne $\{\sigma h \sigma^{-1}\}_{h \in H}$).*

Il revient au même de dire : x et y (resp. H et K) sont *conjugués dans G* ssi il existe un automorphisme intérieur g de G tel que $y = g(x)$ (resp. $K = g(H)$). Les relations (1) et (2) permettent facilement de démontrer que sur l'ensemble G , la relation binaire « x et y sont conjugués » est d'équivalence. On appelle cette relation d'équivalence **conjugaison dans G** ; ses classes sont appelées **classes de conjugaison** de G .

De même, sur l'ensemble $\mathcal{G}(G)$ des sous-groupes de G , la relation « les groupes H et K sont conjugués » est d'équivalence : on l'appelle aussi **conjugaison dans G** (des sous-groupes).

Cycles

Fixons maintenant un ensemble fini non vide E , de cardinal $n \geq 2$, et un élément σ du groupe \mathfrak{S}_E . Cet élément engendre un groupe cyclique que nous noterons $\text{Gr}(\sigma)$; si d est l'ordre de σ dans \mathfrak{S}_E , rappelons que l'application

$$[0, d-1] \rightarrow \text{Gr}(\sigma), k \mapsto \sigma^k$$

est bijective, et que $\mathbb{Z} \rightarrow \mathfrak{S}_E, k \mapsto \sigma^k$ est un homomorphisme de groupes, d'image $\text{Gr}(\sigma)$ et de noyau $d\mathbb{Z}$. Considérons alors la relation binaire \mathcal{R}_σ ainsi définie sur E : pour $x \in E$ et $y \in E$, $x \mathcal{R}_\sigma y$ ssi $(\exists k \in \mathbb{Z} \mid y = \sigma^k(x))$. A cause des propriétés ci-dessus rappelées, on voit que \mathcal{R}_σ est une relation d'équivalence sur E . Par définition, les classes d'équivalence de \mathcal{R}_σ sont appelées les **σ -orbites** de E . C'est en étudiant ces σ -orbites qu'on aboutit au concept de cycle. Remarquons tout de suite qu'une σ -orbite est un singleton $\{a\}$ ssi a est un **point fixe** de σ , i.e. tel que $\sigma(a) = a$. Si ω est une σ -orbite, il résulte de sa définition que ω est *stable* par σ (i.e. $\sigma(\omega) \subset \omega$), et puisque σ est injective, $\sigma|_\omega$ est une permutation de ω .

DÉFINITION V.5.3

Soit E un ensemble **fini** de cardinal $n \geq 2$. On appelle **cycle** sur E toute permutation $\sigma \in \mathfrak{S}_E$ pour laquelle il y a une et une seule σ -orbite de cardinal ≥ 2 . Pour un tel cycle σ , la σ -orbite de cardinal ≥ 2 s'appelle le **support** de σ , et ce cardinal s'appelle la **longueur** de σ ; si l est cette longueur, σ est aussi appelé un **l -cycle**.

A remarquer que dans ces conditions, la permutation Id_E n'est pas un cycle.

Exemple 1 : Prenons $E = \llbracket 1, n \rrbracket$; la permutation

$$c_n = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ 2 & 3 & \dots & n & 1 \end{pmatrix}$$

($c_n(k) = k + 1$ pour $k < n$ et $c_n(n) = 1$) est un cycle sur $\llbracket 1, n \rrbracket$, de support $\llbracket 1, n \rrbracket$ et de longueur n . On l'appelle **cycle canonique** de \mathfrak{S}_n .

Si $2 \leq p \leq n$, la permutation $c_{p,n} \in \mathfrak{S}_n$ telle que $c_{p,n}(x) = x$ pour $x > p$ et $c_{p,n}|_{\llbracket 1, p \rrbracket} = c_p$ est un cycle, de support $\llbracket 1, p \rrbracket$ et de longueur p :

$$c_{p,n} = \begin{pmatrix} 1 & 2 & \dots & p-1 & p & p+1 & \dots & n \\ 2 & 3 & \dots & p & 1 & p+1 & \dots & n \end{pmatrix}.$$

On l'appelle **cycle canonique de longueur p** de \mathfrak{S}_n . Avec la notation abrégée déjà utilisée dans les exercices du § V.4 on peut écrire $c_n = \langle 1, 2, \dots, n \rangle$ et $c_{p,n} = \langle 1, 2, \dots, p \rangle$.

D'après la remarque précédant la définition V.5.3, si c est un cycle de support S sur E , les points $x \in E \setminus S$ sont laissés fixes par c , et ce sont les seuls points fixes de c .

THÉORÈME V.5.1

Soit c un cycle de longueur l sur l'ensemble fini E . Alors l est égal à l'ordre de c dans le groupe \mathfrak{S}_E .

Démonstration :

Soit S le support de c . Prenons un élément particulier $a \in S$. Par définition des c -orbites de E , et du fait que le groupe $\text{Gr}(c)$ est l'ensemble $\{c^k\}_{k \in \llbracket 0, d-1 \rrbracket}$, où d est l'ordre de c , l'application

$$f: \llbracket 0, d-1 \rrbracket \rightarrow S, k \mapsto c^k(a)$$

est surjective. Montrons qu'elle est injective : si

$$0 \leq k_1 \leq k_2 \leq d-1 \quad (k_1 \in \mathbb{N}, k_2 \in \mathbb{N}),$$

et si $c^{k_1}(a) = c^{k_2}(a)$, alors $c^{k_2-k_1}(a) = a$. Soit alors $k_2 - k_1 = k$ et $b \in E$; si $b \notin S$, $c(b) = b$ donc $c^k(b) = b$; si $b \in S$ on a $b = c^r(a)$ pour un $r \in \mathbb{Z}$, d'où

$$c^k(b) = c^k(c^r(a)) = c^{k+r}(a) = c^r(c^k(a)) = c^r(a) = b.$$

Finalement $c^k = \text{Id}_E$. Donc $k \equiv 0 \pmod{d}$; et puisque $0 \leq k \leq d-1$ il en découle $k = 0$, d'où $k_1 = k_2$. Ainsi f est bien injective, donc bijective, d'où

$$l = \text{card}(S) = \text{card}(\llbracket 0, d-1 \rrbracket) = d. \quad \blacksquare$$

La démonstration précédente prouve ceci : pour tout $a \in S$, la bijection :

$$\llbracket 1, d \rrbracket \rightarrow \llbracket 1, d \rrbracket, k \mapsto 1 + f^{-1} \circ c \circ f(k-1),$$

est le *cycle canonique* de $\llbracket 1, d \rrbracket$. Ainsi, la permutation $c|_S$ se « transmue » de l manières différentes (selon le choix de a) en ce cycle canonique. Le cycle c peut se noter

$$\langle a, c(a), c^2(a), \dots, c^{d-1}(a) \rangle.$$

En particulier un cycle $\langle a, c(a) \rangle$ de longueur 2 peut être appelé *transposition* puisque, avec les notations précédentes, il se transmue en une transposition de \mathfrak{S}_n .

Décomposition en cycles d'une permutation

Soit $\sigma \in \mathfrak{S}_E$ une permutation de E , ensemble fini de cardinal $n \geq 2$. On suppose $\sigma \neq \text{Id}_E$ pour qu'il y ait des σ -orbites non réduites à un élément. Notons Ω l'ensemble de ces σ -orbites non-singleton.

On a vu plus haut que, pour chaque $\omega \in \Omega$, on a $\sigma(\omega) = \omega$, et que $\sigma|_\omega = \sigma_\omega$ est une permutation de ω . Désignons par c_ω (pour ω donné dans Ω) la permutation de E définie ainsi : $c_\omega(x) = x$ si $x \notin \omega$,

$$c_\omega(x) = \sigma(x) (= \sigma_\omega(x)) \quad \text{si } x \in \omega.$$

THÉORÈME V.5.2

Avec les notations ci-dessus :

a) Pour chaque σ -orbite ω non-singleton, c_ω est un cycle de support ω .

b) Les $(c_\omega)_{\omega \in \Omega}$ sont deux à deux **permutables**.

c) On a :

$$(I) \quad \sigma = \prod_{\omega \in \Omega} c_\omega \quad (\text{produit dans le groupe } \mathfrak{S}_E).$$

d) L'ordre de σ dans \mathfrak{S}_E est le ppcm des nombres $(\text{card}(\omega))_{\omega \in \Omega}$.

Démonstration :

a) Tout point de $E \setminus \omega$ est fixe pour c_ω ; si $x \in \omega$ et $y \in \omega$, et si $k \in \mathbb{Z}$ est tel que $\sigma^k(x) = y$, alors $c_\omega^k(x) = y$. Finalement ω est une c_ω -orbite et c'est la seule non réduite à un élément, donc c_ω est un cycle de support ω .

b) Soit $\omega \in \Omega$, $\omega' \in \Omega$, $\omega \neq \omega'$; alors $\omega \cap \omega' = \emptyset$. De ce fait on voit que $c_\omega \circ c_{\omega'}(x) = c_{\omega'} \circ c_\omega(x) = x$ si $x \in E \setminus (\omega \cup \omega')$

et $c_\omega \circ c_{\omega'}(x) = c_{\omega'} \circ c_\omega(x) = \sigma(x)$ si $x \in \omega \cup \omega'$,

d'où $c_\omega \circ c_{\omega'} = c_{\omega'} \circ c_\omega$.

c) (I) est évidente (appliquer chaque membre à un élément quelconque x de E). On note au passage que $E \setminus \left(\bigcup_{\omega \in \Omega} \omega \right)$ est l'ensemble des points fixes de σ .

d) Soit $\mu = \text{ppcm}(\text{card}(\omega))_{\omega \in \Omega}$; pour $\omega \in \Omega$, on note $d_\omega = \text{card}(\omega)$: c'est l'ordre du cycle c_ω . Enfin soit d l'ordre de σ . Les c_ω étant deux à deux permutables,

$$(I) \Rightarrow \sigma^\mu = \text{Id}_E \quad (\text{car } c_\omega^\mu = \text{Id}_E \text{ pour tout } \omega).$$

Donc d divise μ . Inversement, de la relation $\sigma^d = \text{Id}_E$, on déduit que, pour chaque ω , on a $c_\omega^d = \text{Id}_E$ (car si $x \in E \setminus \omega$, $c_\omega(x) = x$ et si $x \in \omega$, $c_\omega^d(x) = \sigma^d(x) = x$), d'où : d_ω divise d . Finalement, chaque d_ω divisant d , μ divise d , d'où $\mu = d$. ■

Deux cycles dans \mathfrak{S}_E sont dits **disjoints** ssi leurs supports sont disjoints. En reprenant le raisonnement du b) ci-dessus il est clair que deux cycles disjoints sont permutables.

La relation (I) du théorème V.5.2 est appelée la **décomposition canonique de la permutation σ en cycles disjoints**. Cette décomposition est **unique**, en le sens précis suivant :

THÉORÈME V.5.3

Soit E un ensemble fini de cardinal ≥ 2 , $\sigma \in \mathfrak{S}_E$ ($\sigma \neq \text{Id}_E$) et Ω l'ensemble des σ -orbites de E non réduites à un élément. Supposons trouvés des cycles deux à deux disjoints $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r$ tels que $\sigma = \mathcal{C}_1 \mathcal{C}_2 \dots \mathcal{C}_r$.
 Alors $r = \text{card}(\Omega)$, et il existe une bijection $\varphi : \llbracket 1, r \rrbracket \rightarrow \Omega$ telle que $\mathcal{C}_i = c_{\varphi(i)}$ pour tout $i \in \llbracket 1, r \rrbracket$ (où, pour $\omega \in \Omega$, c_ω est le cycle tel que $c_\omega|_\omega = \sigma|_\omega$ et $c_\omega(x) = x$ si $x \in E \setminus \omega$).

Démonstration (abrégée) :

On vérifie d'abord que les supports S_i des cycles \mathcal{C}_i sont les σ -orbites non-singleton de E , d'où $r = \text{card}(\Omega)$. Puis on contrôle que $\mathcal{C}_i = c_{S_i}$ pour tout i . ■

Exemple 2 : Pour tout ensemble fini E , nous avons appelé *transpositions* les 2-cycles sur E . Le théorème V.5.2 montre alors : pour qu'une permutation $\sigma \in \mathfrak{S}_E$ ($\sigma \neq \text{Id}_E$) soit d'ordre 2, il faut et il suffit que sa décomposition en cycles disjoints ne contienne que des transpositions.

Exemple 3 : Soit $n \in \mathbb{N}$, $p \in \mathbb{N}$, $2 \leq p \leq n$. Le nombre d'inversions du cycle canonique $c_{p,n}$ de \mathfrak{S}_n est évidemment $p-1$, d'où $\varepsilon(c_{p,n}) = (-1)^{p-1}$. La remarque qui suit le théorème V.5.1 permet d'en déduire : **la signature d'un cycle de longueur l est $(-1)^{l-1}$.**

En reprenant la décomposition (I) du théorème V.5.2, on en déduit une expression de la signature $\varepsilon(\sigma)$ d'une permutation $\sigma \in \mathfrak{S}_E$ quelconque ($\sigma \neq \text{Id}_E$) :

$$\varepsilon(\sigma) = (-1)^{\sum_{\omega \in \Omega} (\text{card}(\omega) - 1)}.$$

Mais si $\mathcal{S} = \bigcup_{\omega \in \Omega} \omega$, on a : $\sum_{\omega \in \Omega} \text{card}(\omega) = \text{card}(\mathcal{S})$. Posons $F_\sigma = E \setminus \mathcal{S}$: c'est l'ensemble des *points fixes* de σ .

$$\varepsilon(\sigma) = (-1)^{\text{card}(E) - \text{card}(F_\sigma) - \text{card}(\Omega)},$$

formule qu'on peut ramener à

$$(1) \quad \boxed{\varepsilon(\sigma) = (-1)^{\text{card}(E) - N_\sigma}},$$

où N_σ désigne le nombre total des σ -orbites (réduites ou non à un élément). Cette expression purement « ensembliste » de $\varepsilon(\sigma)$ fait bien voir l'indépendance de ce nombre relativement à tout choix d'un ordre total sur E de préférence à un autre.

Conjugaison dans \mathfrak{S}_E

THÉORÈME V.5.4

Soit E un ensemble fini de cardinal $n \geq 2$. Deux cycles de E sont **conjugués** dans \mathfrak{S}_E ssi ils ont même longueur.

Démonstration :

a) Si deux cycles sont conjugués, ils ont **même ordre** dans \mathfrak{S}_E , puisqu'ils se déduisent l'un de l'autre par un automorphisme particulier de \mathfrak{S}_E ; donc ils ont bien même longueur (théorème V.5.1).

b) Réciproquement, soit c et c' deux cycles de même longueur l , et ω, ω' leurs supports respectifs. Choisissons $a \in \omega, a' \in \omega'$; on sait que les applications

$$[0, l-1] \rightarrow \omega, k \mapsto c^k(a) \quad \text{et} \quad [0, l-1] \rightarrow \omega', k' \mapsto c'^k(a')$$

sont bijectives. On peut donc définir une bijection $\varphi: E \rightarrow E$ ainsi : $\varphi(c^k(a)) = c'^k(a')$ pour $k \in [0, l-1]$ et $\varphi(x) = \psi(x)$ pour $x \in E \setminus \omega$, où $\psi: E \setminus \omega \rightarrow E \setminus \omega'$ est une bijection choisie arbitrairement. On constate alors que : $\varphi \in \mathfrak{S}_E$, et $c' = \varphi \circ c \circ \varphi^{-1}$, donc c et c' sont conjugués dans \mathfrak{S}_E . ■

COROLLAIRE

Pour toute permutation $s \in \mathfrak{S}_E$, notons : $\nu_1(s)$ = nombre de points fixes de s , $\nu_k(s)$ = nombre de k -cycles dans la décomposition en cycles de s ($2 \leq k \leq n$) (pour $s = \text{Id}_E$, $\nu_1(s) = n$ et $\nu_k(s) = 0$ si $k \geq 2$). Alors, pour que deux permutations σ, σ' de \mathfrak{S}_E soient conjuguées dans \mathfrak{S}_E , il faut et il suffit que :

$$\forall k \in [1, n], \quad \nu_k(\sigma) = \nu_k(\sigma').$$

Démonstration :

C'est évident si $\sigma = \text{Id}_E$ ou $\sigma' = \text{Id}_E$.

a) Supposons σ et σ' conjugués ; soit Ω (resp. Ω') l'ensemble des σ -orbites de σ (resp. σ') non-singleton. Soit $\varphi \in \mathfrak{S}_E$ tel que $\sigma' = \varphi \sigma \varphi^{-1}$. Alors $\omega \mapsto \varphi(\omega)$ est une bijection de Ω sur Ω' d'où immédiatement la relation voulue $\nu_k(\sigma) = \nu_k(\sigma')$ pour tout $k \in [1, n]$.

b) Réciproquement supposons $\nu_k(\sigma) = \nu_k(\sigma')$ pour tout $k \in [1, n]$. Les symboles Ω et Ω' ayant la même signification qu'en a), il est clair que l'hypothèse entraîne l'existence d'une bijection $f: \Omega \rightarrow \Omega'$ telle que $\text{card}(f(\omega)) = \text{card}(\omega)$ pour tout $\omega \in \Omega$. Ayant ainsi choisi f , pour chaque $\omega \in \Omega$, les cycles c_ω et $c'_{f(\omega)}$ (où $c_\omega(x) = \sigma(x)$ si $x \in \omega$, $c_\omega(x) = x$ si $x \notin \omega$, et où $c'_{f(\omega)}$ se définit de même) sont conjugués d'après le théorème V.5.2. Soit $g_\omega \in \mathfrak{S}_E$ tel que

$$c'_{f(\omega)} = g_\omega \circ c_\omega \circ g_\omega^{-1};$$

on voit que g_ω définit une bijection de ω sur $f(\omega)$. Construisons maintenant ainsi l'élément $g \in \mathfrak{S}_E$: pour tout $\omega \in \Omega$, $g|_\omega = g_\omega|_\omega$; et si x est point fixe de σ , $g(x) = \psi(x)$, où ψ est une bijection arbitrairement choisie de l'ensemble des points fixes de σ sur celui des points fixes de σ' , ce qui est possible puisque ces ensembles ont même cardinal. Alors $\sigma' = g \circ \sigma \circ g^{-1}$. ■

Exercice 1 : Montrer que pour $N \geq 3$, \mathfrak{U}_N est engendré par les cycles de longueur 3. Montrer même que chacun des ensembles suivants suffit à engendrer \mathfrak{U}_N :

a) $\{ \langle 1, 2, 3 \rangle, \langle 1, 4, 5 \rangle, \dots, \langle 1, 2n, 2n+1 \rangle \}$ si $N = 2n+1$.

b) $\{ \langle 1, 2, 3 \rangle, \langle 1, 4, 5 \rangle, \dots, \langle 1, 2n-2, 2n-1 \rangle, \langle 1, 2, 2n \rangle \}$ si $N = 2n$.

Peut-on diminuer ces ensembles générateurs ?

Montrer également que $\{ \langle 1, 2, 3 \rangle, \langle 1, 2, 4 \rangle, \dots, \langle 1, 2, N \rangle \}$ engendre

Exercice 2 : Montrer que pour $n \geq 2$, \mathfrak{U}_{2n} est engendré par s et t :

$$s = \langle 1, 2, 3 \rangle \quad \text{et} \quad t = \langle 2, 3, \dots, 2n \rangle.$$

Exercice 3 : Soit $p \in \llbracket 2, n \rrbracket$ ($n \geq 2$). Combien y a-t-il de cycles de longueur p dans \mathfrak{S}_n ?

Exercice 4 : Dans \mathfrak{S}_{11} , soit $u = \langle 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 \rangle$ et

$$v = \langle 5, 6, 4, 10 \rangle \circ \langle 11, 8, 3, 7 \rangle.$$

Montrer que le groupe engendré par u et v dans \mathfrak{S}_{11} est de cardinal 7 920.

Exercice 5 : a) Soit G un groupe, G' un groupe abélien et $f : G \rightarrow G'$ un homomorphisme de groupes. Alors f est constant sur chaque classe de conjugaison de G .

b) Soit E un ensemble fini de cardinal ≥ 2 . Dédurre de a) que les seuls homomorphismes de \mathfrak{S}_E dans le groupe multiplicatif \mathbb{C}^* sont 1°) l'application $\chi_0 : x \mapsto 1$ ($x \in E$) 2°) la signature ε sur \mathfrak{S}_E . En déduire que le seul sous-groupe d'indice 2 de \mathfrak{S}_E est \mathfrak{U}_E .

Exercice 6 : Dans \mathfrak{S}_{12} , on donne les éléments u et v de l'exercice 4, et

$$w = \langle 1, 12 \rangle \circ \langle 2, 11 \rangle \circ \langle 3, 6 \rangle \circ \langle 4, 8 \rangle \circ \langle 5, 9 \rangle \circ \langle 7, 10 \rangle.$$

Montrer que $\{u, v, w\}$ engendre dans \mathfrak{S}_{12} un groupe de cardinal 95 040.

Exercice 7 : Soit $n \in \mathbb{N}$, $n \geq 2$ et E un ensemble fini de cardinal n . On note $\nu_k(\sigma)$ le nombre de cycles de longueur k de $\sigma \in \mathfrak{S}_E$ et par \mathcal{P}_n l'ensemble des $(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^*$ tels que

$$\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = n,$$

de sorte que

$$(\nu_1(\sigma), \nu_2(\sigma), \dots, \nu_n(\sigma)) \in \mathcal{P}_n.$$

Soit

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathcal{P}_n.$$

Montrer que le nombre d'éléments de \mathfrak{S}_n tels que $(\nu_1(\sigma), \dots, \nu_n(\sigma)) = \alpha$ est $\frac{n!}{\alpha_1! \dots \alpha_n! 1^{\alpha_1} 2^{\alpha_2} \dots n^{\alpha_n}}$ (commencer par compter les cycles de support donné).

Exercice 8 : Si p est un nombre premier, combien y a-t-il d'éléments d'ordre p dans \mathfrak{S}_p ?

Exercice 9 : a) Si $s = \langle a_1, a_2, a_3 \rangle$ et $t = \langle a_1, a_3, a_2 \rangle$ calculer st^{-1} .

b) Soit $n \geq 3$, et $\sigma \in \mathfrak{S}_n$ une permutation d'ordre ≥ 3 . Dédurre du a) qu'il existe $\tau \in \mathfrak{S}_n$ telle que $\sigma\tau$ soit un cycle de longueur 3 et que σ et τ soient conjuguées dans \mathfrak{S}_n .

Exercice 10 : (Cauchy). Soit $n \in \mathbb{N}$, $n \geq 5$; on note p le plus grand facteur premier de n , et l'on suppose $p \geq 5$. Soit enfin $e \in \llbracket 3, p-1 \rrbracket$. Montrer que \mathfrak{S}_n n'a aucun sous-groupe d'indice e .

Indication : a) les éléments d'ordre p de \mathfrak{S}_n engendrent \mathfrak{U}_n (cf. exercice 9).

b) Si H est un sous-groupe d'indice e de \mathfrak{S}_n , il ne contient pas tous les éléments d'ordre p .

c) Soit s un élément d'ordre p et $\text{Gr}(s) = L$. Montrer que $L \subset H$ et que $L \cap H \neq \{\text{Id}\}$. Conclure. Voir aussi exercice 7 du § V.6.

Exercice 11 : Soit $n \in \mathbb{N}$, $n \geq 2$. Le nombre de sous-groupes cycliques de \mathfrak{S}_n engendrés par un cycle de longueur n est $\frac{(n-1)!}{\varphi(n)}$, où φ est l'indicateur d'Euler.

Exercice 12 : a) Soit c un cycle de longueur n dans \mathfrak{S}_n ($n \geq 2$). Si $k \in \mathbb{N}^*$, quelle est la décomposition en cycles de c^k ?

b) Réciproquement, soit $\sigma \in \mathfrak{S}_n$ dont la décomposition en cycles disjoints ne contient que des cycles de même longueur l (une telle permutation est alors dite *régulière*) ($l \geq 2$), d'où l divise n . Existe-t-il des cycles c de longueur n tels que $c^{n/l} = \sigma$? Combien y en a-t-il ?

Exercice 13 : Décomposer en cycles disjoints les permutations suivantes :

a) $\langle 1, 2, \dots, n-1 \rangle \circ \langle 1, n \rangle$ dans \mathfrak{S}_n .

$$\begin{aligned} \text{b) } & \langle 1, 2, \dots, r, p+1, p+2, p+3, r+1, r+2, \dots, r+s \rangle \circ \\ & \circ \langle r, r-1, \dots, 1, p+1, p+2, p+3, q+1, q+2, \dots, q+t \rangle \end{aligned}$$

où $q = r + s$ et $p = r + s + t = n$, dans \mathfrak{S}_{p+3} .

Exercice 14 : Dans \mathfrak{S}_{mn} , avec $m \geq 2, n \geq 2$, soit deux permutations régulières (cf. exercice 12) d'ordres respectifs m et n et permutables σ et σ' ; si m et n sont premiers entre eux, montrer que $\sigma\sigma'$ est un cycle de longueur mn de \mathfrak{S}_{mn} .

§ V.6 OPÉRATION D'UN GROUPE SUR UN ENSEMBLE

DÉFINITION V.6.1

On appelle **opération à gauche** d'un groupe G sur un ensemble non vide E toute application $G \times E \rightarrow E$, $(\sigma, x) \mapsto \sigma \cdot x$ possédant les propriétés suivantes :

(O₁) pour tous $\sigma \in G, \tau \in G$ et $x \in E$ $\sigma \cdot (\tau \cdot x) = (\sigma\tau) \cdot x$

(O₂) pour tout $x \in E, e_G \cdot x = x$.

On dit parfois **action à gauche** au lieu de « opération à gauche ».

Remarque 1 : Une **action à droite** de G sur E est une application $E \times G \rightarrow E$, $(x, \sigma) \mapsto x \cdot \sigma$ qui vérifie

$$\begin{aligned} (\text{O}'_1) \quad & \forall (\sigma, \tau) \in G^2, \quad \forall x \in E \quad (x \cdot \sigma) \cdot \tau = x \cdot (\sigma\tau) \\ (\text{O}'_2) \quad & \forall x \in E \quad x \cdot e_G = x. \end{aligned}$$

Les propriétés des actions à gauche s'étendent aux actions à droite avec des modifications généralement mineures (l'équivalent du théorème V.6.1 nécessite cependant un peu d'attention et fera l'objet de l'exercice 2).

Dans tout ce qui va suivre (excepté les exemples), nous fixerons un ensemble E non vide, un groupe G (noté multiplicativement), et des actions à gauche de G sur E (avec la notation $(\sigma, x) \mapsto \sigma \cdot x$).

Si $\sigma \in G$, désignons par f_σ l'application $E \rightarrow E$, $x \mapsto \sigma \cdot x$. D'après (O₁),

$$(1) \quad f_\sigma \circ f_\tau = f_{\sigma\tau} \quad \text{pour } \sigma \in G \text{ et } \tau \in G.$$

D'après (O₂), $f_{e_G} = \text{Id}_E$. En conséquence, pour $\sigma \in G$,

$$f_\sigma \circ f_{\sigma^{-1}} = f_{\sigma^{-1}} \circ f_\sigma = f_{e_G} = \text{Id}_E.$$

Donc f_σ est une permutation de E , et sa permutation réciproque est $f_{\sigma^{-1}}$. Ainsi on a défini une application $f: G \rightarrow \mathfrak{S}_E$, $\sigma \mapsto f_\sigma$ qui est un homomorphisme de groupes, dit **associé** à l'action donnée de E sur G .

Réciproquement, soit $g: G \rightarrow \mathfrak{S}_E$ un homomorphisme de \mathfrak{S}_E

$\sigma \in G$ et $x \in E$, posons $\sigma \cdot x = [g(\sigma)](x)$. On vérifie immédiatement que $(\sigma, x) \mapsto \sigma \cdot x$ satisfait les axiomes (O_1) et (O_2) , donc est une action à gauche de G sur E , et que l'homomorphisme de G dans \mathfrak{S}_E associé à cette action est g ; en résumé :

THÉORÈME V.6.1

|| En faisant correspondre, à chaque action à gauche de G sur E , l'homomorphisme de groupes associé de G dans \mathfrak{S}_E , on obtient une bijection, de l'ensemble des actions à gauche de G sur E , sur l'ensemble $\text{Hom}(G, \mathfrak{S}_E)$.

Soit maintenant une action à gauche de G sur E et $\Phi : G \rightarrow \mathfrak{S}_E$ l'homomorphisme associé. Le noyau $\text{Ker}(\Phi)$ est l'ensemble des $\sigma \in G$ tels que $f_\sigma = \text{Id}_E$, c'est-à-dire tels que $\sigma \cdot x = x$ pour tout $x \in E$. L'opération est dite **fidèle** ssi ce noyau est $\{e_G\}$, c'est-à-dire ssi Φ est **injective**, ce qui signifie : *le seul $\sigma \in G$ tel que $\sigma \cdot x = x$ pour tout $x \in E$ est $\sigma = e_G$.*

Une opération fidèle d'un groupe G sur un ensemble E permet de « réaliser » le groupe G comme un sous-groupe de \mathfrak{S}_E , et cela constitue une puissante méthode d'étude de groupes.

Exemple 1 : Soit \mathcal{E} un espace affine de dimension finie $n \geq 1$ sur un corps commutatif K , d'espace directeur E . On a le théorème suivant : « si A est une partie de \mathcal{E} qui engendre affinement \mathcal{E} , et si f est une bijection affine de \mathcal{E} telle que $f(A) = A$ et $f|_A = \text{Id}_A$, alors $f = \text{Id}_{\mathcal{E}}$ ». Prenons pour A une partie *finie* de \mathcal{E} qui engendre affinement \mathcal{E} , et soit G le sous-groupe du groupe affine de \mathcal{E} formé des *bijections affines* f de \mathcal{E} telles que $f(A) = A$. Le théorème rappelé ci-dessus signifie que l'action naturelle de G sur A est **fidèle**, ce qui permet de lui associer un homomorphisme injectif $\Phi : G \rightarrow \mathfrak{S}_A$; au moyen de Φ , on voit que G s'identifie à un sous-groupe du groupe \mathfrak{S}_A ; on ramène ainsi l'étude de tels groupes G à celle de certains sous-groupes de \mathfrak{S}_q , où $q = \text{card}(A)$.

On associe à l'action donnée de G sur E une relation binaire \mathcal{R} ainsi définie :

$$(2) \quad x \mathcal{R} y \quad \text{ssi} \quad \exists \sigma \in G \mid y = \sigma \cdot x$$

dont on vérifie facilement qu'elle est *d'équivalence*.

DÉFINITION V.6.2

~ On appelle **G-orbites de E**, pour l'action à gauche donnée de G sur E , les classes d'équivalence de la relation d'équivalence \mathcal{R} définie par (2). L'action de G sur E est dite **transitive** ssi il y a une seule G -orbite (E tout entier), i.e.

$$\forall x \in E, \quad \forall y \in E, \quad \exists \sigma \in G \mid y = \sigma x.$$

Pour qu'une G -orbite ω soit un singleton $\{a\}$, il faut et il suffit que a soit un point fixe de G , c'est-à-dire vérifie $\sigma \cdot a = a$ pour tout $\sigma \in G$. Une G -orbite ω est G -stable, c'est-à-dire vérifie $f_\sigma(\omega) \subset \omega$ pour tout

particulier, on aura $f_\sigma(\omega) \subset \omega$ et $f_{\sigma^{-1}} \subset \omega$ pour $\sigma \in G$, donc $f_\sigma(\omega) = \omega$, puisque $f_{\sigma^{-1}} = (f_\sigma)^{\langle -1 \rangle}$. Donc, pour tout $\sigma \in G$, $f_\sigma|_\omega$ est une permutation de ω .

Si $a \in E$, l'ensemble $G_a = \{\sigma \in G \mid \sigma \cdot a = a\}$ est, comme on le vérifie aisément, un sous-groupe de G .

DÉFINITION V.6.3

$\}$ Pour tout $a \in E$, le sous-groupe G_a des $\sigma \in G$ tels que $\sigma \cdot a = a$ (on dit : « qui laissent fixe a ») est appelé **groupe d'isotropie** de a , ou encore **stabilisateur** de a .

Si a et b appartiennent à une même orbite ω , choisissons un $\sigma_0 \in G$ tel que $\sigma_0(a) = b$. On constate que $G_b = \sigma_0 \cdot G_a \cdot \sigma_0^{-1}$ (car

$$G_b = \{\sigma \in G \mid \sigma \cdot (\sigma_0 a) = \sigma_0 \cdot a\} = \{\sigma \in G \mid (\sigma_0^{-1} \sigma \sigma_0) \cdot a = a\},$$

autrement dit les sous-groupes G_a et G_b sont conjugués dans G et a fortiori isomorphes.

Exemple 2 : Soit Γ un sous-groupe de \mathfrak{S}_E ; à $\sigma \in \Gamma$ et à $x \in E$, associons $\sigma \cdot x = \sigma(x)$. On définit ainsi une action à gauche de Γ sur E , dite **action naturelle** de Γ sur E . Si E est fini non vide et si $\sigma \in \mathfrak{S}_E$, $\sigma \neq \text{Id}_E$, en prenant pour Γ le groupe engendré par σ , on voit que les Γ -orbites ne sont autres que les σ -orbites définies au § V.5, celles-là même qui ont conduit à la décomposition en cycles de σ .

Exemple 3 : Soit H un sous-groupe de G . À tout $\sigma \in H$ et tout $x \in G$, associons $\sigma * x = \sigma x \sigma^{-1}$. On obtient ainsi une action à gauche de H sur G , dite **opération de H sur G par automorphismes intérieurs**.

Considérons alors l'opération de G sur lui-même par automorphismes intérieurs. Pour que $\sigma \in G$ définisse $f_\sigma = \text{Id}_G$, il faut et suffit que : $\sigma x \sigma^{-1} = x$ pour tout $x \in G$, ou encore $\sigma x = x \sigma$ pour tout $x \in G$. Les éléments ainsi définis sont dits **centraux** dans G . Ils forment un sous-groupe, qui n'est autre que le noyau de l'homomorphisme $G \rightarrow \mathfrak{S}_G$ associé à l'action considérée, que l'on note $\mathcal{Z}(G)$ et qu'on appelle le **centre** de G . Le centre de tout groupe G est évidemment abélien. Pour que $\mathcal{Z}(G) = G$, il faut et il suffit que G soit abélien.

Un élément $x \in G$ est central ssi $\sigma x = x \sigma$ pour tout $\sigma \in G$, c'est-à-dire $\sigma x \sigma^{-1} = x$ pour tout $\sigma \in G$. Donc l'ensemble des éléments centraux $\mathcal{Z}(G)$ est aussi l'ensemble des points fixes du groupe G pour l'action par conjugaison. Les G -orbites pour l'action de G sur lui-même par conjugaison s'appellent **classes de conjugaison** de G (comparer à la définition V.5.2).

Exemple 4 : Soit H un sous-groupe de G . On peut faire opérer H sur G à gauche, par *translation à gauche*, avec $H \times G \rightarrow G$, $(\sigma, x) \mapsto \sigma x$, et aussi à droite, par *translation à droite*, avec $G \times H \rightarrow G$, $(x, \sigma) \mapsto x \sigma$. Les H -orbites pour l'action par translation à gauche (resp. à droite) sont les **classes à gauche** de $G \bmod (H)$ (resp. les **classes à droite** de $G \bmod (H)$).

Exemple 5 : Soit $\mathcal{G}(G)$ l'ensemble des sous-groupes de G ; alors G opère à gauche sur $\mathcal{G}(G)$ par la loi : $G \times \mathcal{G}(G) \rightarrow \mathcal{G}(G)$, $(\sigma, H) \mapsto \sigma H \sigma^{-1}$.

Cette opération est dite *action par conjugaison* de G sur $\mathcal{G}(G)$. Pour un sous-groupe H de G , le stabilisateur, noté \mathcal{N}_H de H est le groupe $\{\sigma \in G \mid \sigma H \sigma^{-1} = H\}$. On voit que $H \subset \mathcal{N}_H$. Ce sous-groupe \mathcal{N}_H s'appelle le **normalisateur de H dans G** .

Exemple 6 : Soit p un entier ≥ 2 , et notons $\mathcal{J}_p(E)$ l'ensemble des p -séquences injectives de E , c'est-à-dire l'ensemble des applications injectives de $\llbracket 1, p \rrbracket$ dans E . Pour $\sigma \in G$ et $a = (a_1, \dots, a_p) \in \mathcal{J}_p(E)$ posons $\sigma \cdot a = (\sigma \cdot a_1, \sigma \cdot a_2, \dots, \sigma \cdot a_p)$. On définit ainsi une action à gauche de G sur $\mathcal{J}_p(E)$, dite *extension à $\mathcal{J}_p(E)$ de l'action donnée de G sur E* . On dit que l'action de G sur E est **p -fois transitive** ssi cette extension à $\mathcal{J}_p(E)$ est transitive, c'est-à-dire ssi on a :

$$\forall a \in \mathcal{J}_p(E), \quad \forall b \in \mathcal{J}_p(E),$$

$$a = (a_i), \quad b = (b_i) \quad \exists \sigma \in G \mid \sigma \cdot a_1 = b_1, \dots, \sigma \cdot a_p = b_p.$$

Exemple 7 : Soit n un entier ≥ 1 et \mathcal{F} l'ensemble des fonctions de \mathbb{Q}^n dans \mathbb{Q} . L'application $\mathfrak{S}_n \times \mathcal{F} \rightarrow \mathcal{F}$, $(\sigma, f) \mapsto \sigma * f$ définie au § V.4, page 171, est une action à gauche de \mathfrak{S}_n sur \mathcal{F} , comme il résulte du lemme 1, § V.4 et du fait que $\text{Id}_{\llbracket 1, n \rrbracket} * f = f$ pour toute $f \in \mathcal{F}$.

Nous rencontrerons dans la suite de cet ouvrage de nombreux autres exemples d'opérations de groupe sur des ensembles sans parler des exemples de géométrie élémentaire que chacun peut facilement imaginer (groupes d'isométries laissant invariante une figure donnée telle qu'un triangle équilatéral, un losange, un carré, un tétraèdre régulier, un cube, ... voire le Rubik cube).

Le résultat essentiel sur cette notion est donné par le :

THÉORÈME V.6.2

Soit ω une G -orbite pour une action à gauche donnée de G sur E .
Donnons-nous $a \in \omega$.
Pour chaque $x \in \omega$, posons $\mathcal{C}_x = \{\sigma \in G \mid \sigma \cdot a = x\}$; alors :

- (I) Chaque \mathcal{C}_x est une **classe à gauche** de G modulo le stabilisateur G_a de a .
- (II) L'application $\omega \rightarrow (G/G_a)_g, x \mapsto \mathcal{C}_x$ est bijective.

Démonstration :

(I) : Si $x \in \omega$, soit $\sigma_0 \in G$ tel que $\sigma_0 \cdot a = x$. Alors

$$\begin{aligned} \mathcal{C}_x &= \{\sigma \in G \mid \sigma \cdot a = \sigma_0 \cdot a\} = \{\sigma \in G \mid \sigma_0^{-1} \sigma \cdot a = a\} = \\ &= \{\sigma \in G \mid \sigma_0^{-1} \sigma \in G_a\}. \end{aligned}$$

Donc $\mathcal{C}_x = \sigma_0 G_a$ est bien une classe à gauche de G mod (

(II) : Si Γ est une classe à gauche $G \bmod (G_a)$, pour $\sigma \in \Gamma$ et $\tau \in \Gamma$, on a : $\tau^{-1} \sigma \in G_a$, donc $(\tau^{-1} \sigma) \cdot a = a$ et par suite $\sigma \cdot a = \tau \cdot a$. On a donc un unique élément $x = \Psi(\Gamma)$ tel que $\sigma \cdot a = x$ pour $\sigma \in \Gamma$. Si $\Phi : \omega \rightarrow (G/G_a)_g$ désigne l'application $x \mapsto \mathcal{C}_x$, on voit à présent que $\Psi \circ \Phi = \text{Id}_\omega$ et que $\Phi \circ \Psi = \text{Id}_{(G/G_a)_g}$, donc Φ et Ψ sont des bijections réciproques l'une de l'autre. ■

COROLLAIRE 1

Avec les notations et hypothèses du théorème V.6.2, pour que l'orbite ω soit finie, il faut et il suffit que l'indice $[G : G_a]$ soit fini, et si c'est le cas, on a :

$$\text{card}(\omega) = [G : G_a] .$$

En conséquence, si G est fini (ce qui entraîne que toute G -orbite est finie), on a :

$$\text{card}(\omega) = \frac{\text{card}(G)}{\text{card}(G_a)} .$$

COROLLAIRE 2

Supposons l'ensemble E fini ; soit \mathcal{C} une partie de E rencontrant chaque G -orbite suivant un singleton. Alors :

$$\text{card}(E) = \sum_{a \in \mathcal{C}} [G : G_a] \quad (\text{« équation aux classes »}) .$$

Démonstration :

Soit \mathcal{O} l'ensemble des G -orbites. L'application $a \mapsto \text{orbite de } a$ est, par hypothèse une bijection de \mathcal{C} sur \mathcal{O} . Mais \mathcal{O} définit une partition de E , donc, par le principe des bergers :

$$\text{card}(E) = \sum_{\omega \in \mathcal{O}} \text{card}(\omega) = \sum_{a \in \mathcal{C}} \text{card}(\text{orb}(a)) .$$

Par le corollaire 1, on a : $\text{card}(\text{orb}(a)) = [G : G_a]$ pour tout a , d'où le résultat. ■

COROLLAIRE 3 (Burnside ⁽¹⁾)

*Supposons E et G finis. Pour tout $\sigma \in G$, soit $\mathcal{F}(\sigma)$ l'ensemble des **points fixes** de σ et $N(\sigma) = \text{card}(\mathcal{F}(\sigma))$. Soit \mathcal{O} l'ensemble des G -orbites ; alors :*

$$\text{card}(\mathcal{O}) = \frac{1}{\text{card}(G)} \left(\sum_{\sigma \in G} N(\sigma) \right) .$$

⁽¹⁾ Burnside (William Snow), mathématicien irlandais (1852-1921).

Démonstration :

On introduit l'ensemble $\Gamma \subset G \times E$ formé des $(\sigma, x) \in G \times E$ pour lesquels $\sigma \cdot x = x$ et on calcule $\text{card}(\Gamma)$ de deux manières grâce au principe de Fubini (cf. § III.3).

1^{re} manière :

$$\text{card}(\Gamma) = \sum_{\sigma \in G} \text{card}(\{x \in E \mid \sigma \cdot x = x\}) = \sum_{\sigma \in G} N(\sigma).$$

2^e manière :

$$\text{card}(\Gamma) = \sum_{x \in E} \text{card}(\{\sigma \in G \mid \sigma \cdot x = x\}) = \sum_{x \in E} \text{card}(G_x)$$

(où G_x désigne le *stabilisateur* de x), soit, puisque $\text{card}(G_x) \times [G : G_x] = \text{card}(G)$:

$$\text{card}(\Gamma) = \sum_{x \in E} \frac{\text{card}(G)}{[G : G_x]} = \text{card}(G) \sum_{x \in E} \frac{1}{\text{card}(\text{orb}(x))}$$

(où $\text{orb}(x)$ désigne l'orbite de x).

Mais par associativité

$$\sum_{x \in E} \frac{1}{\text{card}(\text{orb}(x))} = \sum_{\omega \in \mathcal{O}} \sum_{x \in \omega} \frac{1}{\text{card}(\omega)} = \sum_{\omega \in \mathcal{O}} 1 = \text{card}(\mathcal{O})$$

d'où $\text{card}(\Gamma) = \text{card}(G) \text{card}(\mathcal{O})$. En égalant les deux expressions obtenues pour $\text{card}(\Gamma)$, on obtient le résultat désiré. ■

Exemple 8 (centre d'un p -groupe) :

Soit p un nombre *premier* et G un groupe fini de cardinal p^α , $\alpha \in \mathbb{N}^*$ (un tel groupe de cardinal primaire est appelé un **p -groupe**). Faisons opérer G sur lui-même par conjugaison et écrivons l'équation aux classes après avoir fait choix d'un ensemble \mathcal{C} contenant un et un seul élément de chaque classe de conjugaison : $p^\alpha = \text{card}(G) = \sum_{a \in \mathcal{C}} (G : G_a)$. L'ensemble \mathcal{C} contient nécessairement les *points*

fixes de G dont on a vu dans l'exemple 3 qu'ils constituaient le *centre* $\mathcal{Z}(G)$. Les orbites des $a \in \mathcal{C} \setminus \mathcal{Z}(G) = \mathcal{C}'$ ne sont pas des singletons, c'est-à-dire : si $a \in \mathcal{C}'$, $[G : G_a] > 1$. Mais si $a \in \mathcal{C}'$, $[G : G_a]$ divise $\text{card}(G) = p^\alpha$; c'est donc que $[G : G_a] \equiv 0 \pmod{p}$. Donc

$$p^\alpha = \left(\sum_{a \in \mathcal{Z}(G)} 1 \right) + \sum_{a \in \mathcal{C}'} [G : G_a] = \text{card}(\mathcal{Z}(G)) + S$$

avec

$$S = \sum_{a \in \mathcal{C}'} [G : G_a] \equiv 0 \pmod{p}, \quad \text{d'où} \quad \text{card}(\mathcal{Z}(G)) = p^\alpha - S \equiv 0 \pmod{p}.$$

En particulier, $\mathcal{Z}(G)$ possède au moins p éléments, ce qui prouve que le **centre d'un p -groupe n'est jamais réduit à l'élément neutre**.

Exercice 1 : Soit Γ le sous-groupe de $\text{GL}(2, \mathbb{R})$ formé des matrices $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$

$b > 0$. On fait opérer Γ à gauche sur \mathbb{R}^2 ainsi :

$$\left(\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, (x, y) \right) \mapsto (ax, by).$$

- Trouver les stabilisateurs de $(0, 0)$, de $(1, 1)$ et de $(0, 1)$.
- Trouver les Γ -orbites dans \mathbb{R}^2 . Combien y en a-t-il ?

Exercice 2 :

a) Soit G un groupe ; montrer que la loi $(x, y) \mapsto yx$ noté $x * y$ munit G d'une structure de groupe (dite *opposée* à celle de G et notée G^0), et que l'application $x \mapsto x^{-1}$ est un isomorphisme de G sur G^0 .

b) Etendre le théorème V.6.1 en considérant sur \mathfrak{S}_E la structure de groupe opposée \mathfrak{S}_E^0 à sa structure habituelle.

Exercice 3 : Soit p un nombre premier, et F_p le corps $\mathbb{Z}/p\mathbb{Z}$. Dans le groupe \mathfrak{S}_{F_p} on considère les sous-ensembles suivants :

- l'ensemble \mathcal{T} des bijections du type $x \mapsto x + h$, où $h \in F_p$ (translations de F_p) ;
- l'ensemble Γ des bijections du type $x \mapsto \alpha x + \beta$, où $\alpha \in F_p^*$ et $\beta \in F_p$ (similitudes directes de F_p).

Vérifier que \mathcal{T} et Γ sont des sous-groupes de \mathfrak{S}_{F_p} et préciser leur cardinal. Montrer que Γ est le normalisateur de \mathcal{T} dans \mathfrak{S}_{F_p} .

Exercice 4 : Dans le groupe $\mathfrak{S}_{\mathbb{R}}$, on considère les sous-groupes \mathcal{T} et Γ (des translations et des similitudes directes) définies comme dans l'exercice 3. Soit $\mathcal{N}_{\mathcal{T}}$ le normalisateur de \mathcal{T} dans $\mathfrak{S}_{\mathbb{R}}$. Vérifier que $\Gamma \subsetneq \mathcal{N}_{\mathcal{T}}$. Chercher les $f \in \mathcal{N}_{\mathcal{T}}$ qui sont en même temps des fonctions continues de \mathbb{R} dans \mathbb{R} .

Exercice 5 : Soit n un entier ≥ 2 ; un sous-groupe G de \mathfrak{S}_n est dit **régulier** ssi

1°) $\text{card}(G) = n$

et 2°) G opère *transitivement* sur $\llbracket 1, n \rrbracket$ par l'opération naturelle. Montrer que tout $\sigma \in G$, $\sigma \neq \text{Id}$ est un *dérangement*, c'est-à-dire vérifie $\sigma(x) \neq x$ pour tout $x \in \llbracket 1, n \rrbracket$.

Exercice 6 : Soit p un nombre premier ≥ 3 et G un groupe de cardinal $p + 1$. On suppose trouvé un automorphisme α de G d'ordre p . Démontrer que G est abélien.

Indication : $E = G \setminus \{e\}$ est α -stable. Si $\beta = \alpha \parallel_E$, montrer que β est un cycle de longueur p sur E . Montrer ensuite que G contient au moins un élément d'ordre 2, puis, que tous ses éléments (sauf Id) sont d'ordre 2, et conclure.

Exercice 7 : (théorème de Cauchy, preuve directe).

Soit G un groupe de cardinal n et soit p un nombre premier divisant n . On se propose de montrer que G admet au moins un élément d'ordre p (le cas où G est abélien est supposé connu : voir l'exercice 1 du § V.3).

Vérifier la propriété pour $p = n$. Si $n = kp$ (k entier ≥ 2), supposons la propriété vraie à tout ordre $k' < k$ et montrons-la pour k .

- Etudier le cas où G admet un sous-groupe *strict* H de cardinal divisible par p .
- Etudier le cas contraire en faisant opérer G sur lui-même par conjugaison ; écrire l'équation aux classes, en déduire que p divise $\text{card}(\mathcal{Z}(G))$. Conclure.

Exercice 8 :

a) Vérifier que le centre du groupe \mathfrak{S}_4 est $\{\text{Id}\}$. En déduire que le groupe $\text{Int}(\mathfrak{S}_4)$ des automorphismes intérieurs de \mathfrak{S}_4 est isomorphe à \mathfrak{S}_4 .

b) Vérifier que \mathfrak{S}_4 contient exactement 4 sous-groupes de cardinal 3, et que ces groupes sont $(G_i)_{1 \leq i \leq 4}$ où G_i est le stabilisateur de i pour l'action naturelle de \mathfrak{S}_4 sur $\llbracket 1, 4 \rrbracket$. On désignera par \mathcal{G}_3 l'ensemble $\{G_i\}_{1 \leq i \leq 4}$.

c) Soit Γ le groupe $\text{Aut}(\mathfrak{S}_4)$. Montrer que l'opération à gauche de Γ sur \mathcal{G}_3 par conjugaison est fidèle. En remarquant que $\text{Int}(\mathfrak{S}_4) \subset \Gamma$, en déduire $\text{Int}(\mathfrak{S}_4) = \Gamma$, et le groupe Γ est isomorphe à \mathfrak{S}_4 .

d) Comparer \mathfrak{S}_4 au groupe des isométries conservant un tétraèdre régulier.

Exercice 9 : Pour chaque entier $n \geq 1$, on considère l'opération naturelle du groupe \mathfrak{S}_n sur l'ensemble $\mathcal{F}_n = \mathcal{F}(\mathbb{Q}^n, \mathbb{Q})$ des applications de \mathbb{Q}^n dans \mathbb{Q} , donnée par $(\sigma, f) \mapsto \sigma * f = g$ où $g(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ (cf. § V.4). On prend $n = 4$. Démontrer que les stabilisateurs des 3 fonctions f_1, f_2, f_3 suivantes sont des groupes à 8 éléments isomorphes entre eux :

$$f_1(x_1, x_2, x_3, x_4) = (x_1 + x_2)(x_3 + x_4); \quad f_2(x_1, x_2, x_3, x_4) = x_1 x_2 + x_3 x_4;$$

$$f_3(x_1, x_2, x_3, x_4) = (x_1 + x_2 - x_3 - x_4)^2.$$

Exercice 10 : (plus difficile) : Avec les notations de l'exercice 9 on prend $n = 7$. Soit f la fonction

$$x = (x_i)_{1 \leq i \leq 7} \mapsto x_1 x_2 x_4 + x_2 x_3 x_5 + x_3 x_4 x_6 + x_4 x_5 x_7 + x_5 x_6 x_1 + x_6 x_7 x_2 + x_7 x_1 x_3.$$

Démontrer que le stabilisateur de f dans \mathfrak{S}_7 est un groupe à 168 éléments.

Exercice 11 : Soit n un entier ≥ 3 tel que $n \nmid \varphi(n) = 1$, où φ est l'indicateur d'Euler. Montrer que tout groupe de cardinal un tel n est abélien, donc cyclique, compte tenu de l'exercice 2 du § V.3.

N.B. Cet exercice « sec » ayant été résolu en quelques heures par un élève de spéciale M' (promotion 1983) qui en a donné une solution originale et élégante (est-ce parce qu'il y avait une récompense à la clé ?), nous ne fournissons pas d'indication.

Exercice 12 : Soit n et d deux entiers, $n \geq 2, d \geq 1$. On note $S_{n,d}$ l'ensemble

$$\{(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n \mid \alpha_1 + \alpha_2 + \dots + \alpha_n = d\}.$$

On fait opérer \mathfrak{S}_n sur $S_{n,d}$ à gauche ainsi : si $\sigma \in \mathfrak{S}_n$ et

$$\alpha = (\alpha_i) \in S_{n,d}, \quad \sigma \cdot \alpha = (\alpha_{\sigma^{-1}(1)}, \dots, \alpha_{\sigma^{-1}(n)}).$$

a) Soit $\mathcal{T}_{n,d}$ l'ensemble des suites $(k_0, k_1, \dots, k_d) \in \mathbb{N}^{d+1}$ telles que

$$k_0 + k_1 + \dots + k_d = n \quad \text{et} \quad k_1 + 2k_2 + \dots + dk_d = d.$$

Si $\alpha = (\alpha_i) \in S_{n,d}$, on définit $\varphi(\alpha) = (k_0, k_1, \dots, k_d) \in \mathcal{T}_{n,d}$ de la manière suivante :

$$k_i = \text{card}(\{q \in \llbracket 1, n \rrbracket \mid \alpha_q = i\}) \quad \text{pour} \quad 0 \leq i \leq d.$$

Montrer : $\varphi(\alpha) = \varphi(\beta) \Leftrightarrow \alpha$ et β sont dans la même \mathfrak{S}_n -orbite de $S_{n,d}$. Si Ω est l'ensemble de ces orbites, en déduire une bijection naturelle $\Phi : \Omega \rightarrow \mathcal{T}_{n,d}$.

b) Soit $k = (k_0, \dots, k_d) \in \mathcal{T}_{n,d}$ et $\omega = \Phi^{-1}(k)$. Montrer :

$$\text{card}(\omega) = \frac{n!}{k_0! k_1! \dots k_d!}.$$

En déduire

$$\sum_{k \in \mathcal{T}_{n,d}} \frac{n!}{k_0! k_1! \dots k_d!} = \binom{n+d-1}{n-1}.$$

c) On développe le multinôme $(x_1 + x_2 + \dots + x_n)^d$ où les x_i sont éléments d'un anneau commutatif A . Montrer :

$$(x_1 + x_2 + \dots + x_n)^d = \sum_{\omega \in \Omega} \frac{d!}{(1!)^{k_1(\omega)} (2!)^{k_2(\omega)} \dots (d!)^{k_d(\omega)}} F_{\omega},$$

où $(k_0(\omega), k_1(\omega), \dots, k_d(\omega)) = \Phi(\omega)$ pour $\omega \in \Omega$

et où $F_{\omega} = \sum_{(\alpha_1, \alpha_2, \dots, \alpha_n) \in \omega} x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ pour $\omega \in \Omega$.

d) Application numérique à $d = 7$.

§ V.7 SOUS-GROUPES DISTINGUÉS. GROUPE QUOTIENT

DÉFINITION V.7.1

Un sous-groupe H d'un groupe G est dit **distingué** dans G ⁽¹⁾, et l'on écrit $H \triangleleft G$, ssi l'on a : $\sigma H \sigma^{-1} = H$ pour tout $\sigma \in G$.

Cette condition signifie que $\sigma H = H \sigma$ pour tout $\sigma \in G$, c'est-à-dire : Pour tout élément σ de G , la classe à droite et la classe à gauche de $\sigma \bmod (H)$ coïncident. En particulier, si $H \triangleleft G$, les ensembles des classes à gauche et à droite $(G/H)_g$ et $(G/H)_d$ sont égaux à un même ensemble qu'il est naturel de noter G/H et d'appeler ensemble des classes $\bmod (H)$.

Remarquons qu'on a toujours $\{e_G\} \triangleleft G$ et $G \triangleleft G$. Dans les cas où G n'admet pas d'autre sous-groupe distingué que $\{e_G\}$ et G , on dit par définition que c'est un groupe **simple** ⁽²⁾.

Exemple 1 : Si G est abélien, il est évident que tout sous-groupe de G est distingué dans G .

Il existe des groupes dans lesquels tout sous-groupe est distingué et qui cependant ne sont pas abéliens.

Les assertions du théorème suivant se vérifient de façon élémentaire :

THÉORÈME V.7.1

- (I) Toute intersection de sous-groupes distingués dans G est un sous-groupe distingué dans G .
- (II) Si $f : G \rightarrow G'$ est un homomorphisme de groupes, pour tout sous-groupe $H' \triangleleft G'$, on a : $f^{-1}(H') \triangleleft G$; pour tout sous-groupe $H \triangleleft G$, on a : $f(H) \triangleleft \text{Im}(f)$.
- (III) Si $(G_i)_{i \in I}$ est une famille non vide de groupes, et si pour tout $i \in I$, H_i est un sous-groupe distingué dans G_i , alors $\prod_{i \in I} H_i \triangleleft \prod_{i \in I} G_i$.

COROLLAIRE

- Si $f : G \rightarrow G'$ est un homomorphisme de groupes, le noyau $\text{Ker}(f)$ est un sous-groupe distingué dans G .

Ce corollaire est un outil puissant pour établir qu'un sous-groupe est distingué.

⁽¹⁾ On dit encore : sous-groupe « invariant », ou : sous-groupe « normal ».

⁽²⁾ La classification des groupes finis simples, commencée dans les années 1860 et achevée depuis peu, n'a pas été une mince affaire. Cf. « Pour la Science », février 1986. « Le Théorème géant ».

Exemple 2 : Soit A un anneau commutatif ; on note \mathcal{S}_A le sous-groupe de \mathfrak{S}_A formé des bijections de la forme $g_{\alpha, \beta} : x \mapsto \alpha x + \beta$, où $\beta \in A$ et où $\alpha \in \mathcal{U}_A$ (groupe des éléments inversibles de A). Pour chaque $f \in \mathcal{S}_A$, il y a unicité du couple $(\alpha, \beta) \in \mathcal{U}_A \times A$ tel que $f = g_{\alpha, \beta}$. On vérifie que l'application $\varphi : \mathcal{S}_A \rightarrow \mathcal{U}_A, f \mapsto \alpha$, est un homomorphisme de groupes. Or $\text{Ker}(\varphi)$ est le groupe \mathcal{T}_A des translations de A (bijections du type $x \mapsto x + h$). Donc $\mathcal{T}_A \triangleleft \mathcal{S}_A$.

Exemple 3 : Soit K un corps commutatif et n un entier ≥ 2 ; l'application déterminant notée $\det : \text{GL}(n, K) \rightarrow K^*$ est un homomorphisme de groupes (voir le chapitre « Déterminants »). Son noyau est donc un sous-groupe distingué de $\text{GL}(n, K)$, qu'on note $\text{SL}(n, K)$.

L'intérêt majeur de la notion de sous-groupe distingué est de donner lieu à une structure naturelle de groupe sur l'ensemble G/H des classes mod (H) :

THÉORÈME V.7.2

Soit G un groupe, H un sous-groupe distingué de G , et $\varphi : G \rightarrow G/H$ l'application canonique. Il existe une, et une seule, structure de groupe sur G/H telle que φ soit un homomorphisme de groupes. Avec cette structure, on a :

$$\text{Ker}(\varphi) = H, \quad \text{Im}(\varphi) = G/H.$$

Démonstration :

Tout d'abord, φ est surjectif, en tant qu'application de G sur un de ses ensembles quotients. Si une structure de groupe $(X, Y) \mapsto X * Y$ existe sur G/H de manière que $\varphi(x) * \varphi(y) = \varphi(xy)$ pour tout $(x, y) \in G^2$ on voit que nécessairement, pour $X \in G/H, Y \in G/H$, l'élément $\varphi(xy)$ ne dépend pas du choix de (x, y) dans $X * Y$, et que cet élément est XY .

Réciproquement, soit $X, Y \in G/H$. Donnons-nous $(x, y) \in X \times Y, (x', y') \in X \times Y$. On a :

$$(x' y')^{-1} xy = y'^{-1} x'^{-1} xy.$$

Or $x'^{-1}x \in H$, et comme $y^{-1}y' \in H, y' = ys$ avec $s \in H$. D'où

$$y'^{-1} x'^{-1} xy = s^{-1}(y^{-1} x'^{-1} xy).$$

Comme $x'^{-1}x \in H$ et $H \triangleleft G$, on a $y^{-1} x'^{-1} xy \in H$; d'où

$$s^{-1}(y^{-1} x'^{-1} xy) = (x' y')^{-1} xy \in H,$$

ce qui prouve que $\varphi(x' y') = \varphi(xy)$.

Ainsi l'élément $\varphi(x, y)$, où $(x, y) \in X \times Y$ ne dépend pas du choix du couple (x, y) . On peut donc définir une loi sur G/H

$X * Y =$ valeur commune des $\varphi(xy)$ pour $(x, y) \in X * Y$. Il est alors élémentaire de vérifier que cette loi est une loi de groupe grâce à la relation $\varphi(xy) = \varphi(x) * \varphi(y)$ qui est vraie par définition de $*$ et du fait que G est un groupe. Le neutre est H . L'inverse de $X \in G'/H$ est $X^{-1} = \{x^{-1}\}_{x \in H}$. Enfin, puisque

$$\varphi(xy) = \varphi(x) * \varphi(y) \quad (x \in G, y \in G'),$$

φ est un homomorphisme de groupes. Son noyau est $\{x \in G \mid \varphi(x) = H\}$, c'est-à-dire l'ensemble H . ■

Si G est *abélien*, la démonstration ci-dessus se simplifie considérablement, car pour prouver $(x' y')^{-1} xy \in H$, on écrit simplement :

$$(x' y')^{-1} xy = (x'^{-1} x)(y'^{-1} y).$$

De plus, dans ce cas, la surjectivité de φ et la relation

$$\varphi(xy) = \varphi(yx) = \varphi(x) * \varphi(y) = \varphi(y) * \varphi(x) \quad (x \in G, y \in G)$$

montrent que la loi de groupe $*$ sur G/H est *commutative*.

DÉFINITION V.7.2

Si H est un sous-groupe **distingué** du groupe G , la structure de groupe G/H définie dans le théorème V.7.2 s'appelle **structure de groupe quotient** de G par H . Le groupe obtenu se note G/H (**groupe quotient** de G par H) et φ s'appelle l'**homomorphisme canonique** de G dans G/H .

De ce qui précède, on déduit aussitôt que : **tout groupe quotient d'un groupe abélien est abélien.**

Les théorèmes qui suivent permettent d'utiliser de façon mécanique la structure de groupe quotient :

THÉORÈME V.7.3

Soit $f : G \rightarrow G'$ un homomorphisme de groupes, N le sous-groupe $\text{Ker}(f)$ de G , I le sous-groupe $\text{Im}(f)$ de G' ,

$$\varphi : G \rightarrow G/N$$

l'homomorphisme canonique et $j : I \rightarrow G'$ l'injection canonique (qui est un homomorphisme de groupes).

(I) Il existe une et une seule application $\bar{f} : G/N \rightarrow I$ telle que

$$j \circ \bar{f} \circ \varphi = f.$$

(II) \bar{f} est un isomorphisme de groupes.

Cela se traduit par le diagramme commutatif ci-contre. On dit que l'isomorphisme \bar{f} est **déduit de f par passage au quotient**. La décomposition $f = j \circ \bar{f} \circ \varphi$ s'appelle **décomposition canonique** de f .

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \varphi \downarrow & & \uparrow j \\ G/N & \dashrightarrow & I \end{array}$$

Démonstration :

Remarquons d'abord que f est constante sur chaque classe mod (N) , car si $x \in G$, $y \in G$ et $x^{-1}y \in N$, cela entraîne $f(x^{-1}y) = e_{G'}$, d'où $f(x) = f(y)$.

On en déduit facilement qu'il existe une et une seule application $\bar{f}: G/N \rightarrow I$ telle que $f = j \circ \bar{f} \circ \varphi$; \bar{f} est l'application qui, à chaque $X \in G/N$, associe la valeur constante de f sur X . D'où l'assertion (I). De plus, \bar{f} est surjective par définition de I .

Montrons que \bar{f} est un homomorphisme de groupes : soit $X \in G/N, Y \in G/N$. Par définition de \bar{f} , $\bar{f}(XY)$ est la valeur constante de f sur XY . Soit $x \in X$ et $y \in Y$, alors $xy \in XY$, donc

$$\bar{f}(XY) = f(xy) = f(x) f(y) = \bar{f}(X) \bar{f}(Y),$$

d'où l'assertion. Enfin montrons que \bar{f} est injectif, ce qui achèvera la preuve : il s'agit de voir que

$$\text{Ker}(\bar{f}) = \{\text{neutre de } G/N\} = \{N\}.$$

Or

$$\text{Ker}(\bar{f}) = \{X \in G/N \mid \forall x \in X, f(x) = e_{G'}\} = \{\text{Ker}(f)\} = \{N\},$$

d'où le résultat. ■

Dans la plupart des cas, le théorème V.7.3 autorise une description du groupe image I en fonction de la structure de G seul.

Exemple 4 : Soit x un élément d'un groupe G , et $\psi_x: \mathbb{Z} \rightarrow G$ l'homomorphisme $k \mapsto x^k$. L'image de ψ_x est le groupe $\text{Gr}(x)$ engendré par x . Si x est sans torsion, on a : $\text{Ker}(\psi_x) = \{0\}$ et ψ_x définit un isomorphisme de \mathbb{Z} sur $\text{Gr}(x)$. Si x est de torsion et d'ordre $d \geq 1$, on a : $\text{Ker}(\psi_x) = d\mathbb{Z}$ et ψ_x définit par passage au quotient un isomorphisme du groupe additif $\mathbb{Z}/d\mathbb{Z}$ sur $\text{Gr}(x)$, ce qui décrit parfaitement $\text{Gr}(x)$.

Exemple 5 : Soit G un groupe ; à chaque $\sigma \in G$ associons l'automorphisme intérieur f_σ de G ($f_\sigma: x \mapsto \sigma x \sigma^{-1}$). On a vu au § V.5 que l'application $\Phi: G \rightarrow \text{Aut}(G)$, $\sigma \mapsto f_\sigma$ est un homomorphisme

Son image est $\text{Int}(G)$, groupe formé par tous les f_σ , appelé *groupe des automorphismes intérieurs* de G . Son noyau est

$$\{\sigma \in G \mid \forall x \in G \quad x\sigma = \sigma x\} = \mathcal{Z}(G):$$

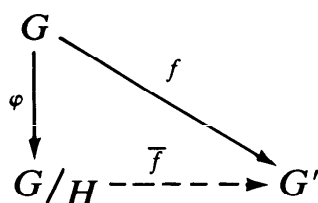
c'est le **centre de G** qui est donc un sous-groupe distingué de G (ce qui est immédiat à vérifier directement). Par passage au quotient, l'homomorphisme Φ définit un isomorphisme de groupes $G/\mathcal{Z}(G) \xrightarrow{\cong} \text{Int}(G)$, (le symbole \cong est là pour signifier que la flèche est un isomorphisme), et cela donne une description du groupe $\text{Int}(G)$.

THÉOREME V.7.4 (Propriété universelle du quotient)

|| Soit G un groupe et H un sous-groupe **distingué** de G , $\varphi: G \rightarrow G/H$ l'homomorphisme canonique. Soit G' un autre groupe, et $f: G \rightarrow G'$ un homomorphisme de groupes tel que $H \subset \text{Ker}(f)$. Alors il existe une application unique $\bar{f}: G/H \rightarrow G'$ telle que $f = \bar{f} \circ \varphi$, et cette application est un homomorphisme de groupes.

Démonstration :

Du fait que $H \subset \text{Ker}(f)$, on déduit que f est *constante* sur les classes de $G \bmod (H)$, d'où l'existence et l'unicité de \bar{f} . Le fait que \bar{f} soit un homomorphisme de groupes s'en déduit exactement comme dans la démonstration du théorème V.7.3. ■



Le diagramme commutatif ci-dessus illustre le théorème V.7.4 qui s'énonce brièvement en disant que f « se factorise à travers G/H en \bar{f} ».

Etude détaillée du groupe \mathfrak{S}_4

A titre d'illustration concrète de l'ensemble des notions exposées dans ce chapitre sur la théorie des groupes, nous allons étudier en détail le groupe symétrique \mathfrak{S}_4 , qui joue d'ailleurs un rôle très particulier parmi les groupes \mathfrak{S}_n .

1^{re} étape : classes de conjugaison dans \mathfrak{S}_4

Rappelons d'abord que toute permutation σ de \mathfrak{S}_n se décompose en produit de cycles disjoints et que deux cycles de $[[1, n]]$ sont conjugués ssi ils ont même longueur. Le corollaire du théorème V.5.4 donne immédiatement les classes de conjugaison dans \mathfrak{S}_4 . Si c_1 est le nombre des points fixes de σ , et c_i le nombre de cycles de longueur i dans la décomposition en cycles disjoints de σ pour $i = 2, 3, 4$, une classe de conjugaison est caractérisée par la suite (c_1, c_2, c_3, c_4) où $c_i \in \mathbb{N}$ et $c_1 + 2c_2 + 3c_3 + 4c_4 = 4$ qui détermine l'ensemble des $\sigma \in$

à la classe correspondante. Or les seules possibilités pour les suites $(c_i)_{1 \leq i \leq 4}$ sont : $(4, 0, 0, 0)$; $(2, 1, 0, 0)$; $(0, 2, 0, 0)$; $(1, 0, 3, 0)$ et $(0, 0, 0, 1)$. Un décompte des nombres de parties à 2 ou 3 éléments de $\llbracket 1, 4 \rrbracket$ donne donc les cinq classes de conjugaison suivantes dans \mathfrak{S}_4 :

- Classe du neutre $((c_i) = (4, 0, 0, 0))$, réduite à $e = \text{Id}_{\llbracket 1, 4 \rrbracket}$, notée \mathcal{C}_1 .
- Classe \mathcal{C}_2 formée des *six transpositions* : $((c_i) = (2, 1, 0, 0))$.
- Classe \mathcal{C}_3 formée des *trois doubles transpositions* : $((c_i) = (0, 2, 0, 0))$. Ces doubles transpositions sont précisément :

$$\langle 1, 2 \rangle \cdot \langle 3, 4 \rangle = u, \quad v = \langle 1, 3 \rangle \cdot \langle 2, 4 \rangle, \quad w = \langle 1, 4 \rangle \cdot \langle 2, 3 \rangle.$$

- Classe \mathcal{C}_4 formée des 8 cycles d'ordre 3 $((c_i) = (1, 0, 3, 0))$.
- Classe \mathcal{C}_5 formée des 6 cycles d'ordre 4 $((c_i) = (0, 0, 0, 1))$.

Si l'on pose $\alpha = \langle 1, 4, 2, 3 \rangle$, $\beta = \langle 1, 2, 3, 4 \rangle$, $\gamma = \langle 1, 3, 4, 2 \rangle$ on constate que $\mathcal{C}_5 = \{\alpha, \alpha^3, \beta, \beta^3, \gamma, \gamma^3\}$.

Si l'on s'intéresse aux permutations **paires**, il faut les rechercher dans \mathcal{C}_1 , dans \mathcal{C}_3 ou dans \mathcal{C}_4 , ce qui donne le *groupe alterné* $\mathcal{U}_4 = \mathcal{C}_1 \cup \mathcal{C}_3 \cup \mathcal{C}_4$. Un sous-groupe presque évident de \mathcal{U}_4 est $K = \mathcal{C}_1 \cup \mathcal{C}_3 = \{e, u, v, w\}$: c'est le groupe de Klein, isomorphe à $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +)$.

Quant aux permutations **impaires**, on les trouve dans $\mathfrak{S}_4 \setminus \mathcal{U}_4 = \mathcal{C}_2 \cup \mathcal{C}_5$.

2^e étape : opération naturelle de \mathfrak{S}_4 sur \mathcal{C}_3 .

A l'aide du corollaire du théorème V.5.4, on voit qu'on fait opérer à gauche \mathfrak{S}_4 sur \mathcal{C}_3 par conjugaison, i.e. par la loi $\mathfrak{S}_4 \times \mathcal{C}_3 \rightarrow \mathcal{C}_3$, $(\sigma, \tau) \mapsto \sigma\tau\sigma^{-1}$. Notons $\Phi : \mathfrak{S}_4 \rightarrow \mathfrak{S}_{\mathcal{C}_3}$ l'homomorphisme de groupes associé à cette opération. Montrons d'abord que Φ est *surjectif* : il suffit pour cela de vérifier que toute transposition de \mathcal{C}_3 est dans $\text{Im}(\Phi)$, puisque les transpositions engendrent $\mathfrak{S}_{\mathcal{C}_3}$. Soit par exemple à vérifier que $\langle u, v \rangle \in \text{Im}(\Phi)$: en choisissant $\tau = \langle 2, 3 \rangle$ il est facile de vérifier que $\tau w \tau^{-1} = w$, $\tau u \tau^{-1} = v$ et $\tau v \tau^{-1} = u$, donc $\langle u, v \rangle \in \text{Im}(\Phi)$ et de même pour $\langle u, w \rangle$ et pour $\langle v, w \rangle$. Donc Φ est bien surjectif. *A fortiori* cette opération de \mathfrak{S}_4 sur \mathcal{C}_3 est *transitive*. Comme $\text{card}(\mathcal{C}_3) = 3$, on en déduit (cf. corollaire 3 du théorème V.3.1) que

$$\text{card}(\text{Ker}(\Phi)) = \text{card} \mathfrak{S}_4 / \text{card}(\mathfrak{S}_{\mathcal{C}_3}) = \frac{24}{6} = 4.$$

De plus, les stabilisateurs G_u , G_v et G_w sont des sous-groupes deux à deux conjugués dans \mathfrak{S}_4 et de cardinal $\text{card}(\mathfrak{S}_4) / \text{card}(\mathcal{C}_3) = 8$. Ces stabilisateurs contiennent tous $\text{Ker}(\Phi)$. *Montrons alors que* $\text{Ker}(\Phi) = K$. Comme K est abélien, il est clair que $K \subset \text{Ker}(\Phi)$, et comme $\text{card}(K) = 4 = \text{card}(\text{Ker}(\Phi))$, il en résulte que $K = \text{Ker}(\Phi)$. En particulier le groupe K est **distingué** dans \mathfrak{S}_4 , et le groupe quotient \mathfrak{S}_4/K est isomorphe à $\mathfrak{S}_{\mathcal{C}_3}$ par application du théorème V.7.3, c'est-à-dire finalement à \mathfrak{S}_3 . Par ailleurs $K \subset \mathcal{U}_4$, d'où $K \triangleleft \mathcal{U}_4$ et le groupe quotient \mathcal{U}_4/K est isomorphe à \mathcal{U}_3 (c'est-à-dire à $(\mathbb{Z}/3\mathbb{Z}, +)$).

3^e étape : sous-groupes distingués de \mathfrak{S}_4 .

On déduit de ce qui précède que les sous-groupes distingués de \mathfrak{S}_4 autres que $\{e\}$ et \mathfrak{S}_4 sont K et \mathcal{U}_4 . En effet, si Γ est un tel sous-groupe, il est union de certaines classes de conjugaison, dont obligatoirement \mathcal{C}_1 .

Si $\Gamma \subset \mathcal{U}_4$ et $\mathcal{C}_4 \subset \Gamma$, alors $\Gamma = \mathcal{U}_4$ (engendré par les cycles d'ordre 3).

Si $\Gamma \subset \mathcal{U}_4$ et $\Gamma \cap \mathcal{C}_4 = \emptyset$, alors nécessairement $\mathcal{C}_3 \subset \Gamma$ donc $\Gamma = K$.

Si $\Gamma \not\subset \mathcal{U}_4$ et si $\mathcal{C}_2 \subset \Gamma$, nécessairement $\Gamma = \mathfrak{S}_4$ (engendré par les transpositions).

Si $\Gamma \not\subset \mathcal{U}_4$ et si $\mathcal{C}_5 \subset \Gamma$, alors $\mathcal{C}_3 \subset \Gamma$ car, par exemple, $\alpha^2 \in \mathcal{C}_3$, d'où encore $\Gamma = \mathfrak{S}_4$ car $\mathcal{C}_3 \cup \mathcal{C}_5$ engendre \mathfrak{S}_4 .

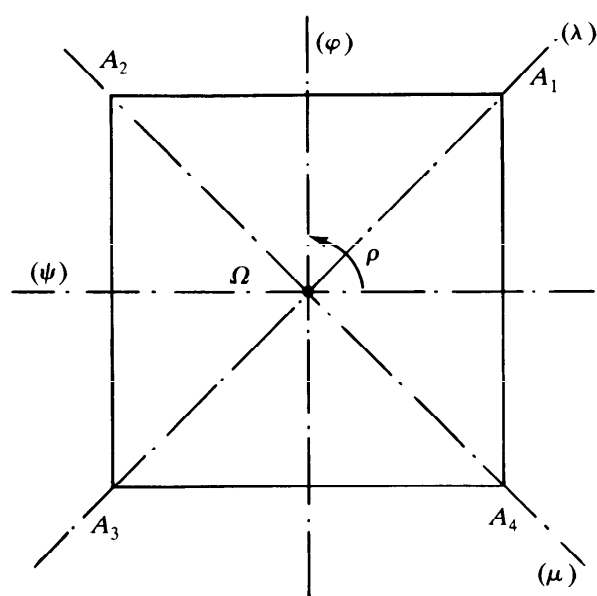
4^e étape : étude des groupes G_u, G_v, G_w .

Chacun de ces groupes est respectivement engendré par $K \cup \{\alpha\}$, $K \cup \{\beta\}$, $K \cup \{\gamma\}$ et on a :

$$G_u \cap G_v \cap G_w = K = G_v \cap G_w = G_w \cap G_u = G_u \cap G_v.$$

Rappelons que ces groupes G_u, G_v, G_w sont deux à deux conjugués dans \mathfrak{S}_4 , donc isomorphes entre eux. L'intersection de chacun d'eux avec \mathcal{U}_4 est K . Chacun de ces groupes est, nous allons le voir, isomorphe au **groupe du carré** Γ , c'est-à-dire au groupe Γ des isométries d'un plan affine euclidien \mathcal{E} laissant un carré donné globalement invariant. Pour le voir, notons $\{A_1, A_2, A_3, A_4\}$ le carré, Ω son centre, ρ et $\bar{\rho}$ les deux rotations de centre Ω , d'angle droit, λ et μ les symétries orthogonales autour des diagonales $(A_1 A_3)$ et $(A_2 A_4)$, φ et ψ les symétries orthogonales autour des médianes du carré, s la symétrie centrale de centre ω . La numérotation des (A_i) est choisie pour que ρ induise $\langle A_1, A_2, A_3, A_4 \rangle$ sur les sommets du carré, et pour que φ induise $\langle A_1, A_2 \rangle \cdot \langle A_3, A_4 \rangle$.

On constate que l'opération naturelle de Γ sur l'ensemble $\mathcal{S} = \{A_1, A_2, A_3, A_4\}$ est *fidèle*, ce qui permet d'identifier Γ à un sous-groupe de $\mathfrak{S}_{\mathcal{S}}$, c'est-à-dire de \mathfrak{S}_4 (la bijection $A_i \mapsto i$ permet d'identifier \mathcal{S} et $\llbracket 1, 4 \rrbracket$, donc $\mathfrak{S}_{\mathcal{S}}$ et \mathfrak{S}_4 , ce que nous ferons). Cela dit, on voit bien que, dans cette identification, le groupe Γ correspond au groupe G_v engendré par $K \cup \{\beta\}$, d'où l'isomorphisme recherché $\Gamma \cong G_v$. D'ailleurs Γ apparaît, ainsi que G_v , comme le sous-groupe de \mathfrak{S}_4 formé des $\sigma \in \mathfrak{S}_4$ qui laissent globalement invariant l'ensemble $\{\{1, 3\}, \{2, 4\}\}$ (on pourra comparer Γ au groupe étudié dans l'exercice 9 du § V.6 : ils sont isomorphes). Le lecteur n'aura aucun mal à dresser la table du groupe du carré en s'aidant au besoin de la figure ci-dessous.



$d \backslash g$	e	s	ρ	$\bar{\rho}$	φ	ψ	λ	μ
e	e	s	ρ	$\bar{\rho}$	φ	ψ	λ	μ
s	s	e	$\bar{\rho}$	ρ	ψ	φ	μ	λ
ρ	ρ	$\bar{\rho}$	s	e	μ	λ	φ	ψ
$\bar{\rho}$	$\bar{\rho}$	ρ	e	s	λ	μ	ψ	φ
φ	φ	ψ	λ	μ	e	s	ρ	$\bar{\rho}$
ψ	ψ	φ	μ	λ	s	e	$\bar{\rho}$	ρ
λ	λ	μ	ψ	φ	$\bar{\rho}$	ρ	e	s
μ	μ	λ	φ	ψ	ρ	$\bar{\rho}$	s	e

(Se souvenir que dans gd le facteur de gauche g correspond à l'isométrie effectuée en second, conformément à la notation utilisée pour composer des ap

5^e étape : recensement de tous les sous-groupes de \mathfrak{S}_4 .

En possession des résultats précédents, le lecteur pourra dresser une liste complète des sous-groupes de \mathfrak{S}_4 . Nous avons déjà rencontré $\{e\}$, K , G_u , G_v , G_w , \mathfrak{U}_4 , \mathfrak{S}_4 . Nous avons remarqué que K est un sous-groupe *distingué* de \mathfrak{U}_4 , autrement dit que \mathfrak{U}_4 n'est pas simple (c'est cette propriété qui permet la résolution « par radicaux » des équations du 4^e degré ; l'exercice 15 propose de prouver que pour $n \geq 5$ le groupe \mathfrak{U}_n est simple). Il faut ajouter à cette liste les 4 sous-groupes (isomorphes à \mathfrak{S}_3) qui stabilisent les points de $\llbracket 1, 4 \rrbracket$ pour l'action naturelle de \mathfrak{S}_4 (\cong groupe du triangle équilatéral) et les 4 sous-groupes des précédents isomorphes à \mathfrak{U}_3 (rotations de $\frac{2k\pi}{3}$). En comptant les 6 sous-groupes engendrés par une transposition, les trois sous-groupes d'ordre 2 de $K : \{e, u\}$, $\{e, v\}$, $\{e, w\}$, les 3 sous-groupes engendrés par 2 transpositions disjointes, et les 3 sous-groupes engendrés par un 4 cycle, cela fait en tout 30 sous-groupes de \mathfrak{S}_4 (y compris $\{e\}$ et \mathfrak{S}_4) et nous laissons au lecteur le soin de préciser ceux de ces sous-groupes qui sont distingués dans \mathfrak{S}_4 , ceux qui sont commutatifs. Il pourra aussi écrire des chaînes (pour l'inclinaison) de sous-groupes distingués, en cherchant celles de longueur maximum.

Exercice 1 : Soit p un nombre premier. Montrer que tous les groupes de cardinal p sont isomorphes à $(\mathbb{Z}/p\mathbb{Z}, +)$.

Exercice 2 : Soit $\mathcal{Z}(G)$ le centre d'un groupe G . Montrer que si $G/\mathcal{Z}(G)$ est un groupe cyclique, alors G est abélien. En déduire que si G est fini de cardinal p^2 , avec p premier, alors G est abélien.

Indication : utiliser l'exemple 8 du § V.6.

Exercice 3 : Soit G un groupe et $[G, G]$ son groupe des commutateurs (voir la définition au § V.4, exercice 5). Démontrer que $[G, G] \triangleleft G$, et que le groupe quotient $G/[G, G]$ est abélien. Montrer que si H est un sous-groupe de G , distingué dans G , tel que G/H soit abélien, on a : $[G, G] \subset H$. En déduire que pour tout homomorphisme $f : G \rightarrow G'$, où G' est un groupe abélien, on a un, et un seul, homomorphisme $\bar{f} : G/[G, G] \rightarrow G'$ tel que $\bar{f} \circ \varphi = f$, où φ est l'homomorphisme canonique de G dans $G/[G, G]$.

Exercice 4 : Soit G un groupe et H et K deux sous-groupes tels que $K \subset H, K \triangleleft G, H \triangleleft G$; vérifier que H/K est un sous-groupe de G/K et que $H/K \triangleleft G/K$. Démontrer que les groupes $(G/K)/(H/K)$ et G/H sont isomorphes, en étudiant l'homomorphisme composé de $\text{can} : G \rightarrow G/K$ et $\text{can} : G/K \rightarrow (G/K)/(H/K)$ (« second théorème d'isomorphisme de Nøther » ⁽¹⁾).

Exercice 5 : Soit H et K deux sous-groupes d'un groupe G , avec $K \triangleleft G$.

- Vérifier que $HK = \{hk\}_{(h,k) \in H \times K}$ est un sous-groupe de G , et que $K \triangleleft HK$.
- Vérifier $H \cap K \triangleleft H$.
- Montrer que les groupes $H/H \cap K$ et HK/K sont isomorphes, en étudiant l'homomorphisme $\varphi \circ j$ où $j : H \rightarrow HK, x \mapsto x$ et où $\varphi : HK \rightarrow HK/K$ est l'homomorphisme canonique (« premier théorème d'isomorphisme de Nøther »).

Exercice 6 :

- Déterminer les sous-groupes finis du groupe quotient R/\mathbb{Z} ; de \mathbb{C}/\mathbb{Z} .
- Soit n entier ≥ 2 . Déterminer les sous-groupes finis du groupe quotient $\mathbb{R}^n/\mathbb{Z}^n$.

⁽¹⁾ Emmy Nøther (1882-1935), mathématicienne allemande, qui a contribué à la création de l'Algèbre dite moderne.

Exercice 7 : Soit H un sous-groupe d'un groupe G , d'indice n dans G . Démontrer qu'il existe un sous-groupe N de G , distingué dans G , tel que $N \subset H$ et $[G : N] \mid n!$

Indication : Faire opérer G de manière convenable sur l'ensemble $(G/H)_g$.

Application : Si n est le plus petit facteur premier de $\text{card}(G)$, alors $H \triangleleft G$.

Exercice 8 : Soit G un groupe produit de deux autres groupes G_1 et G_2 : $G = G_1 \times G_2$. On note $\overline{G}_1 = G_1 \times \{e_{G_2}\}$, $\overline{G}_2 = \{e_{G_1}\} \times G_2$. Soit H un sous-groupe distingué dans G tel que $H \cap \overline{G}_i = \{e_G\}$ pour $i = 1, 2$. Montrer que H est abélien.

Indication : tous ses éléments sont involutifs.

Exercice 9 : Montrer que tout sous-groupe fini du groupe quotient $G = \mathbb{Q}/\mathbb{Z}$ est cyclique (cf. exercice 2), et que \mathbb{Q}/\mathbb{Z} est union de ses sous-groupes finis, mais n'est pas de type fini. Pour tout sous-groupe fini H de G , montrer que $G/H \cong G$.

Exercice 10 : On considère le groupe Γ des éléments inversibles de l'anneau $A = \mathbb{Z}/4\mathbb{Z}$. Montrer que le groupe G des bijections $f: A \rightarrow A$ de la forme $x \mapsto \alpha x + \beta$, $\alpha \in \Gamma$, $\beta \in A$, est isomorphe au groupe du carré. Quel est le centre \mathcal{Z} de G ? le groupe quotient G/\mathcal{Z} ?

Exercice 11 : Soit p un nombre premier. Dans le groupe quotient \mathbb{Q}/\mathbb{Z} , on considère le sous-groupe G formé des éléments dont l'ordre est une puissance de p (on vérifiera que G est bien un sous-groupe). Montrer que G n'est pas de type fini et que tout sous-groupe strict H de G est fini et vérifie : G/H est isomorphe à G .

Exercice 12 : Soit G le groupe engendré dans $\text{GL}(2, \mathbb{C})$ par les matrices

$$A = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Montrer que G est de cardinal 8, non abélien, que tous ses sous-groupes sont distingués et que G contient, outre e_G , un élément d'ordre 2 et 6 éléments d'ordre 4 (G s'appelle le **groupe quaternionique**). Quel est le centre de G ? le groupe quotient $G/\mathcal{Z}(G)$?

Exercice 13 : Faire le recensement de tous les groupes de cardinal 8. On montrera que tout groupe non abélien de cardinal 8 est isomorphe soit au groupe du carré, soit au groupe quaternionique. Quant aux groupes abéliens on en trouvera trois différents isomorphes à $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ additifs.

Exercice 14 : Soit G un groupe tel que le groupe quotient $G/\mathcal{Z}(G)$ soit fini ($\mathcal{Z}(G)$ désigne le centre de G). Démontrer que le groupe des commutateurs $[G, G]$ est fini (cf. définition du § V.4, exercice 5).

Indication d'une méthode possible : soit $n = [G : \mathcal{Z}(G)]$.

a) Montrer que le nombre L des commutateurs $xyx^{-1}y^{-1}$ distincts est fini, $L \leq n^2$.

b) Montrer que tout élément de $[G, G]$ est un produit de commutateurs dans lequel chacun des L commutateurs intervient au plus n fois (si $c = [x, y]$ intervient $n+1$ fois dans $X \in [G, G]$ on pourra ramener X à la forme $c^{n+1} \gamma'_{n+2} \dots \gamma'_N$ et remplacer $[x, y]^{n+1}$ par $[x, y^2][xyx^{-1}, y]^n$). Noter que $[x, y]^n \in \mathcal{Z}(G)$.

c) En déduire : $\text{card}([G, G]) \leq n^{2n^3}$.

Exercice 15 : Soit $n \in \mathbb{N}$, $n \geq 5$. On se propose de montrer que le groupe \mathcal{U}_n est simple.

a) Soit G , un sous-groupe distingué de \mathcal{U}_5 distinct de $\{e\}$. Si G contient un élément d'ordre 3, montrer que $G = \mathcal{U}_5$. Si G contient un élément d'ordre 2, soit c , montrer que l'on peut trouver un 3-cycle s tel que $scs^{-1}c$ soit un 3-cycle et en déduire que l'on a encore $G = \mathcal{U}_5$. Si G contient un élément d'ordre 5, soit c , montrer qu'il existe $t \in \mathcal{U}_5$ d'ordre 2 tel que $tct^{-1}c$ soit un 3-cycle et en déduire que $G = \mathcal{U}_5$.

b) On suppose maintenant $n \geq 6$. Soit G un sous-groupe distingué de \mathcal{U}_n distinct de $\{e\}$:

1° Si G contient un élément $\sigma \neq e$ laissant fixes tous les éléments de \mathcal{U}_n sauf 5 au plus, alors $G = \mathcal{U}_n$.

2° On suppose que tout $\sigma \in G \setminus \{e\}$ dérange au moins 6 éléments. Soit F_σ l'ensemble des points fixes d'un tel σ (par hypothèse $\text{card}(F_\sigma) \leq n - 6$). Montrer qu'il existe $a, b \in \llbracket 1, n \rrbracket$ tels que $a, b, \sigma(a), \sigma(b)$ soient distincts et non dans F_σ . Soit E l'ensemble $F_\sigma \cup \{a, b, \sigma(a), \sigma(b)\}$, a et b étant ainsi choisis. Montrer qu'il existe $s \in \mathcal{U}_n$ tel que $s(b) = a, s(\sigma(b)) = c$, avec $c \notin E$, et $s(x) = x$ si $x \in F_\sigma \cup \{a, b\}$.

3° Soit alors $\tau = \sigma^{-1} s^{-1} \sigma s$. Montrer que $\tau \neq e$, $\tau \in G$, et $\text{card}(F_\tau) \geq 1 + \text{card}(F_\sigma)$. Par un argument de récurrence descendante, en déduire à l'aide de b) 1° que $G = \mathcal{U}_n$.

Application : Montrer que si $n \geq 5$, le seul sous-groupe de \mathfrak{S}_n dont l'indice dans \mathfrak{S}_n soit $\leq n$ est \mathcal{U}_n , qui est d'indice 2.

Exercice 16 : Montrer que si un groupe fini de cardinal $2n$ a un sous-groupe de cardinal n , ce ne peut être qu'un sous-groupe distingué. En donner des illustrations (groupe symétrique, groupe alterné, mais aussi **groupe diédral** D_n des isométries laissant globalement invariant un polygone régulier de n côtés).

Exercice 17 : Étudier le groupe des 24 rotations qui font coïncider un cube donné avec lui-même. Former tous les sous-groupes et préciser ceux qui sont distingués.

Exercice 18 : Soit G le groupe défini dans l'exercice 10 du § V.6.

On considère le tableau :

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 1 \\ 4 & 5 & 6 & 7 & 1 & 2 & 3 \end{bmatrix}.$$

Vérifier que G s'identifie au groupe des $\sigma \in \mathfrak{S}_7$ qui laissent globalement invariant l'ensemble des 7 colonnes du tableau quand on considère chaque colonne comme **ensemble** à 3 éléments.

Faire opérer G de la manière naturelle sur ces 7 éléments, puis sur les *parties à deux éléments* de l'ensemble de ces 7 colonnes. Pour cette dernière opération, montrer que les stabilisateurs ont 4 éléments. Montrer que G est engendré par $\{u, v\}$, où $u = \langle 1, 2, 3, 4, 5, 6, 7 \rangle$ et $v = \langle 2, 4 \rangle \circ \langle 5, 6 \rangle$. Enfin, chercher s'il existe des sous-groupes de G isomorphes à \mathfrak{S}_4 ou \mathcal{U}_4 , et prouver que G est un groupe *simple*.

Chapitre VI

STRUCTURES D'ESPACE VECTORIEL ET D'ALGÈBRE ; NOMBRES COMPLEXES

§ VI.1 STRUCTURE D'ESPACE VECTORIEL

DÉFINITION VI.1.1

Soit K un corps commutatif. On appelle **K-espace vectoriel** (en abrégé : **K-ev**) tout ensemble E muni d'une **loi interne** (que l'on note toujours additivement) et d'une application $K \times E \rightarrow E$, notée en général $(\lambda, x) \mapsto \lambda \cdot x$ ou λx , et appelée **loi externe de domaine K**, de manière que les propriétés suivantes soient satisfaites :

(EV₁) E est un **groupe abélien** pour sa loi interne.

(EV₂) Pour tous $\lambda_1 \in K$, $\lambda_2 \in K$, $x_1 \in E$ et $x_2 \in E$, on a :

$$a) (\lambda_1 + \lambda_2) \cdot x_1 = \lambda_1 \cdot x_1 + \lambda_2 \cdot x_1$$

$$b) \lambda_1 \cdot (x_1 + x_2) = \lambda_1 \cdot x_1 + \lambda_1 \cdot x_2.$$

(EV₃) Pour tout $x \in E$, $1_K \cdot x = x$.

(EV₄) Pour tous $\lambda \in K$, $\mu \in K$ et $x \in E$, on a :

$$\lambda \cdot (\mu \cdot x) = (\lambda \mu) \cdot x.$$

Les propriétés (EV₁) à (EV₄) s'appellent *axiomes de la structure de K-ev*. La loi externe $K \times E \rightarrow E$ s'appelle *multiplication (des éléments de E) par les scalaires*, le terme « scalaires » désignant les éléments de K . Les éléments de E sont alors appelés *vecteurs*. La propriété (EV₂a) (resp. EV₂b) exprime la *distributivité* de la multiplication par rapport aux scalaires (resp. par rapport à l'addition des vecteurs).

Un espace vectoriel n'est *jamais vide* (il contient l'élément neutre du groupe $(E, +)$ noté 0_E et appelé *vecteur nul*). Si E se réduit à $\{0_E\}$ on dit que E est un *espace vectoriel nul*.

Soit un K -ev quelconque, noté E . A un vecteur x de E on peut associer l'application $\psi_x : K \rightarrow E, \lambda \mapsto \lambda \cdot x$ qui est un *homomorphisme de groupes additifs* grâce à (EV_{2a}) ; en particulier $0_K \cdot x = 0_E$. De même à un scalaire λ de K on peut associer l'application $h_\lambda : E \rightarrow E, x \mapsto \lambda \cdot x$ qui est un *endomorphisme du groupe additif* $(E, +)$ grâce à (EV_{2b}), appelé **homothétie de rapport λ** ; en particulier $\lambda \cdot 0_E = 0_E$.

On remarque que ψ_{0_E} est l'homomorphisme nul de $(K, +)$ dans $(E, +)$ et que h_{0_K} est l'endomorphisme nul du groupe additif E .

THÉORÈME VI.1.1

Avec les notations qui précèdent :

- (I) Si $(\lambda, x) \in K \times E$, on a $\lambda \cdot x = 0_E$ ssi $\lambda = 0_K$ ou $x = 0_E$.
- (II) Si $x \in E \setminus \{0_E\}$, l'homomorphisme ψ_x est **injectif**.
- (III) Si $\lambda \in K^* = K \setminus \{0_K\}$, l'homothétie h_λ est une **bijection** de E sur E (donc un **automorphisme du groupe** $(E, +)$), et la bijection réciproque est l'homothétie $h_{\lambda^{-1}}$.

Démonstration :

Pour prouver (I) supposons que $\lambda \cdot x = 0_E$, avec $\lambda \neq 0_K$. Alors, en utilisant (EV₄) :

$$\lambda^{-1} \cdot (\lambda x) = (\lambda^{-1} \lambda) \cdot x = 1_K \cdot x$$

et grâce à (EV₃) : $\lambda^{-1} \cdot (\lambda x) = x$; mais par hypothèse

$$\lambda^{-1} \cdot (\lambda x) = \lambda^{-1} \cdot 0_E = 0_E$$

(déjà vu), d'où $x = 0_E$.

On en déduit que si $x \in E \setminus \{0_E\}$, le noyau de l'homomorphisme ψ_x est $\{0_K\}$, donc ψ_x est injectif.

Enfin si $\lambda \in K^*$, on a vu que $\lambda^{-1} \cdot (\lambda x) = x (\forall x \in E)$, ce qui se traduit par : $h_{\lambda^{-1}} \circ h_\lambda = \text{Id}_E$. De même $h_\lambda \circ h_{\lambda^{-1}} = \text{Id}_E$, ce qui montre bien que h_λ et $h_{\lambda^{-1}}$ sont des bijections réciproques l'une de l'autre. ■

Notons également que si $x \in E$, l'opposé de x dans le groupe additif E n'est autre que $(-1_K) \cdot x$; en effet :

$$x + (-1_K) \cdot x = 1_K \cdot x + (-1_K) \cdot x = (1_K + (-1_K)) \cdot x = 0_K \cdot x = 0_E.$$

Exemple 1 : Soit K un corps commutatif ; l'addition de K et la multiplication de K munissent de manière évidente K d'une structure de K -ev, dite *naturelle* (c'est la distinction entre vecteurs et scalaires qui paraît ici artificielle !).

Exemple 2 : Soit I un ensemble non vide et K un corps commutatif. Considérons le groupe additif $K^I = \mathcal{F}(I, K)$ (cf. § II.4, exerr

munit d'une structure de K -ev, dite *naturelle*, en définissant la multiplication par les scalaires « point par point », c'est-à-dire en posant $\lambda \cdot x = (\lambda x_i)_{i \in I}$. Le K -ev obtenu se notera K^I , comme le groupe. En particulier si $I = \llbracket 1, n \rrbracket$, on note ce K -ev K^n : il est construit sur l'ensemble des suites (x_1, x_2, \dots, x_n) de n éléments de K ; si $n = 1$ on reconnaît le K -ev de l'exemple 1.

Exemple 3 : Soit L un corps commutatif et K un sous-corps de L ; considérons un L -ev E ; conservons la structure de groupe additif de E et munissons-le de la loi externe de domaine K obtenue par restriction à $K \times E$ de l'application $L \times E \rightarrow E$ qui définit la structure de L -ev de E . On obtient un K -ev dont le groupe additif est E , *déduit du L -ev E par restriction des scalaires à K* , et souvent noté $E_{(K)}$. En particulier $L_{(K)}$ est un K -espace vectoriel.

Sous-espaces vectoriels

DÉFINITION VI.1.2

Soit K un corps commutatif et E un K -ev ; on appelle **sous- K -espace vectoriel de E** (en abrégé : **sous- K -ev**) toute partie F de E qui est un sous-groupe additif de E , et telle que $(\forall \lambda \in K)(\forall x \in F) \lambda \cdot x \in F$.

Si F est un sous- K -ev de E , le sous-groupe additif F de E peut être muni de la multiplication par les scalaires $K \times F \rightarrow F$, $(\lambda, x) \mapsto \lambda \cdot x$, ce dernier produit étant *calculé dans le K -ev E* . On vérifie alors immédiatement qu'on a ainsi muni F d'une structure de K -ev, dite **induite par E sur F** .

Réciproquement, si F est une partie de E **stable** pour l'addition et le produit par un scalaire de K , et **qui est non vide**, alors F est un sous- K -ev de E . En effet si $x \in F$ on a : $(-1_K) \cdot x = -x \in F$, et comme F possède au moins un élément x_0 , $x_0 + (-x_0) = 0_E \in F$, ce qui prouve bien que F est un sous-groupe additif de E , d'où finalement :

THÉORÈME VI.1.2

Soit F une partie non vide d'un K -ev E ; pour que F soit un sous- K -ev de E , il faut et il suffit que F soit **stable pour l'addition et la loi externe de E** .

Exemple 4 : Pour tout K -ev E , $\{0_E\}$ et E sont des sous- K -ev de E .

Exemple 5 : Les seuls sous- K -ev de K sont $\{0_K\}$ et K . En effet soit F un tel sous- K -ev autre que $\{0_K\}$ et soit $x \neq 0_K$ un de ses éléments ; alors $x^{-1} \cdot x = 1_K \in F$, d'où, pour tout $\lambda \in K$: $\lambda \cdot 1_K = \lambda \in F$ et $F = K$.

Exemple 6 : Soit I un ensemble non vide et K un corps commutatif ; alors le sous-groupe additif $K^{(I)}$ du K -ev K^I est un sous- K -ev de K^I , qu'on note encore $K^{(I)}$. Si I est fini, on a : $K^{(I)} = K^I$. (Cf. la définition de $K^{(I)}$ page 157.)

Exemple 7 : Soit L un corps commutatif de caractéristique $p > 0$, et soit P le sous-corps premier de L . Alors $L_{(P)}$ est un P -ev (cf. exemple 3) ; et les sous- P -ev de $L_{(P)}$ ne sont autres que les sous-groupes additifs de L , mais ce résultat devient faux si L est de caractéristique 0.

Une conséquence évidente du théorème VI.1.2 est la suivante :

THÉORÈME VI.1.3

|| Si $(F_i)_{i \in I}$ est une famille de sous- K -ev d'un K -ev E , alors l'intersection $\bigcap_{i \in I} F_i$ est un sous- K -ev de E .

Exercice 1 : Soit (F_i) une famille non vide de sous- K -ev d'un K -ev E pour laquelle $(\forall i \in I)(\forall j \in I) \exists k \in I \mid F_i \cup F_j \subset F_k$. Démontrer qu'alors $\bigcup_{i \in I} F_i$ est un sous- K -ev de E .

Exercice 2 : Soit $E = \mathbb{R}_+^* \times \mathbb{R}$. On définit l'addition dans E par $(a, b) + (c, d) = (ac, b + d)$ et la loi externe de domaine \mathbb{R} par $\lambda \cdot (a, b) = (a^\lambda, \lambda b)$. Montrer que $(E, +, \cdot)$ est un \mathbb{R} -espace vectoriel.

Exercice 3 : On munit K^2 de l'addition habituelle et de la loi externe de domaine K définie par $\lambda \cdot (a, b) = (\lambda a, 0)$. Pourquoi n'obtient-on pas un K -ev ?

Exercice 4 : Soit le K -ev $E = K^n$. Les sous-ensembles suivants sont-ils des sous- K -cv ?

- a) les n -uplets (x_1, x_2, \dots, x_n) tels que $x_1 = 0$ et $x_2 = 0$
- b) tels que $x_1 + x_2 = 0$
- c) tels que $x_1 \neq 0$
- d) tels que $x_1 = 0$ ou $x_2 = 0$
- e) tels que $x_1 = x_2$.

Exercice 5 : Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$ quels sont parmi les sous-ensembles suivants ceux qui sont des sous- \mathbb{R} -cv ?

- a) les trinômes du second degré
- b) les fonctions telles que $f(1) = 2f(0)$
- c) les fonctions telles que $f(1) - f(0) = 1$
- d) les fonctions telles que $f(x) = f(a - x) \forall x \in \mathbb{R}$.

Exercice 6 : L'ensemble des fonctions continues de \mathbb{R} dans \mathbb{R} est-il un \mathbb{R} -ev ? L'ensemble des fonctions dérivables de \mathbb{R} dans \mathbb{R} en est-il un sous- \mathbb{R} -ev ? Citer d'autres exemples.

Exercice 7 : Soit G un groupe abélien, noté additivement.

- a) Montrer que G peut être muni d'au plus une structure de \mathbb{Q} -ev.
- b) Montrer que, pour que G puisse être muni d'une structure de \mathbb{Q} -cv, il faut et il suffit qu'il vérifie les conditions :
 - 1° G est *sans torsion* (i.e. G n'a pas d'élément de torsion) et
 - 2° G est *divisible* (i.e. $\forall x \in G \forall n \in \mathbb{N}^*, \exists y \in G \mid ny = x$).

§ VI.2 APPLICATIONS LINÉAIRES

DÉFINITION VI.2.1

Soit K un corps commutatif et E_1, E_2 deux K -ev. On appelle **application K -linéaire** de E_1 dans E_2 (ou **homomorphisme de K -ev** de E_1 dans E_2) toute application $f : E_1 \rightarrow E_2$ qui est un **homomorphisme de groupes additifs** et qui vérifie

$$(AL) \quad (\forall \lambda \in K) \quad (\forall x \in E_1) \quad f(\lambda \cdot x) = \lambda \cdot f(x).$$

Lorsqu'il n'y a pas possibilité d'ambiguïté sur K , on dit « application linéaire ». La propriété (AL) s'appelle *homogénéité par rapport aux scalaires*.

DÉFINITION VI.2.2

Soit K un corps commutatif et E_1, E_2 deux K -ev. On appelle **isomorphisme de K -ev** de E_1 sur E_2 toute bijection $f : E_1 \rightarrow E_2$ telle que f et $f^{\langle -1 \rangle}$ soient K -linéaires. On dit que E_1 et E_2 sont **isomorphes** ssi il existe au moins un isomorphisme de K -ev de E_1 sur E_2 .

Si E est un K -ev, une application K -linéaire de E dans E est appelée un **endomorphisme de K -ev** de E . Un isomorphisme de K -ev $E \rightarrow E$ est appelé un **automorphisme** du K -ev E .

Les propriétés suivantes sont élémentaires :

($\mathcal{L}I$) Si E est un K -ev et si F est un sous K -ev de E , l'injection canonique $j : F \rightarrow E$ est K -linéaire. En particulier Id_E est un automorphisme du K -ev E .

($\mathcal{L}II$) Si E_1, E_2, E_3 sont des K -ev, si $f : E_1 \rightarrow E_2$ et $g : E_2 \rightarrow E_3$ sont K -linéaires, alors $g \circ f : E_1 \rightarrow E_3$ est K -linéaire. En particulier toute restriction d'une application K -linéaire à un sous- K -ev est encore K -linéaire.

($\mathcal{L}III$) Soit E_1, E_2 deux K -ev ; si $f : E_1 \rightarrow E_2$ est une **bijection K -linéaire**, alors f est un **isomorphisme** de K -ev car $f^{\langle -1 \rangle}$ est aussi K -linéaire, et $f^{\langle -1 \rangle} : E_2 \rightarrow E_1$ est un autre isomorphisme de K -ev.

($\mathcal{L}IV$) Le composé de deux isomorphismes de K -ev est un isomorphisme de K -ev.

La relation entre espaces vectoriels définis sur le même corps commutatif K « le K -ev E est isomorphe au K -ev F » est donc réflexive, symétrique et transitive. Nous nous garderons cependant de prétendre que c'est une relation d'équivalence car la relation « E est un K -ev » n'est pas collective : il n'existe aucun ensemble dont tout K -ev soit éléme

THÉORÈME VI.2.1

|| Soit E_1 et E_2 deux K -ev et $f : E_1 \rightarrow E_2$ une application K -linéaire. Alors pour tout sous- K -ev F_1 de E_1 , $f(F_1)$ est un sous- K -ev de E_2 ; et pour tout sous- K -ev F_2 de E_2 , $f^{-1}(F_2)$ est un sous- K -ev de E_1 .

C'est une conséquence immédiate du théorème VI.1.2.

En particulier, $f(E_1) = \text{Im}(f)$ est un sous- K -ev de E_2 .

$$f^{-1}(\{0_{E_2}\}) = f^{-1}(0_{E_2})$$

est un sous- K -ev de E_1 appelé **noyau** de f , et noté $\text{Ker}(f)$. Du théorème II.4.3, il découle aussitôt :

THÉORÈME VI.2.2

|| Soit E_1, E_2 deux K -ev et $f : E_1 \rightarrow E_2$ une application K -linéaire. Pour que f soit **injective**, il faut et il suffit que $\text{Ker}(f) = \{0_{E_1}\}$.

Exemple 1 : Soit I un ensemble non vide et K un corps commutatif. Pour tout $i \in I$, la projection canonique $p_i : K^I \rightarrow K$ est K -linéaire ; on a :

$$\text{Im}(p_i) = K \quad \text{et} \quad \text{Ker}(p_i) = \{x \in K^I \mid x_i = 0_K\}.$$

Exemple 2 : Avec les notations K et I de l'exemple 1, l'application $S :$

$$K^{(I)} \rightarrow K, x = (x_i)_{i \in I} \mapsto S(x) = \sum_{i \in I} x_i,$$

est K -linéaire et surjective.

Exemple 3 : Soit E un K -ev ; pour tout $\lambda \in K$, l'homothétie de rapport λ (définie au § VI.1) de E est K -linéaire ; si $\lambda \neq 0_K$, c'est un automorphisme du K -ev E .

Donnons-nous maintenant deux K -ev E_1 et E_2 et considérons l'ensemble des applications K -linéaires de E_1 dans E_2 que l'on note $\text{Hom}_K(E_1, E_2)$, ou plus simplement $\text{Hom}_K(E)$ si $E_1 = E_2 = E$. Il est facile de munir ces ensembles de façon très naturelle de structures algébriques remarquables.

Soit $f \in \text{Hom}_K(E_1, E_2)$ et $g \in \text{Hom}_K(E_1, E_2)$. L'application

$$f + g : E_1 \rightarrow E_2$$

définie par $x \mapsto f(x) + g(x)$ où $x \in E_1$ est un homomorphisme de groupes additifs. Soit $\lambda \in K$. L'application $\lambda f : E_1 \rightarrow E_2$ définie par $x \mapsto \lambda \cdot f(x)$ est aussi un homomorphisme de groupes additifs. De plus

$$f + g \in \text{Hom}_K(E_1, E_2)$$

car

$$\begin{aligned}(f + g)(\lambda x) &= f(\lambda x) + g(\lambda x) = \lambda \cdot f(x) + \lambda \cdot g(x) = \\ &= \lambda \cdot (f(x) + g(x)) = \lambda \cdot (f + g)(x).\end{aligned}$$

De même $\lambda f \in \text{Hom}_K(E_1, E_2)$ car

$$\begin{aligned}(\lambda f)(\mu x) &= \lambda \cdot f(\mu x) = \lambda (\mu f(x)) = (\lambda \mu) \cdot f(x) = \\ &= \mu \cdot (\lambda f(x)) = \mu \cdot (\lambda f)(x).\end{aligned}$$

Il suffit alors de quelques vérifications immédiates pour obtenir :

THÉORÈME VI.2.3

Sur $\text{Hom}_K(E_1, E_2)$, la loi interne $(f, g) \mapsto f + g$ telle que

$$(\forall x \in E_1) \quad (f + g)(x) = f(x) + g(x),$$

et la loi externe de domaine $K : (\lambda, f) \mapsto \lambda f$, telle que

$$(\forall x \in E_1) \quad (\lambda f)(x) = \lambda \cdot (f(x)).$$

définissent une structure de K -espace vectoriel.

L'ensemble $\text{Hom}_K(E_1, E_2)$ sera systématiquement muni de cette structure dite **naturelle** de K -ev, et le K -ev ainsi obtenu se note encore $\text{Hom}_K(E_1, E_2)$.

Remarque 1 : On aura noté au passage le rôle essentiel joué par la *commutativité de la multiplication de K* pour démontrer que, si $f \in \text{Hom}_K(E_1, E_2)$ et si $\lambda \in K$, alors $\lambda f \in \text{Hom}_K(E_1, E_2)$.

Donnons-nous trois K -ev E_1, E_2, E_3 ; grâce à ($\mathcal{L}II$), on définit une application de $\text{Hom}_K(E_2, E_3) \times \text{Hom}_K(E_1, E_2)$ dans $\text{Hom}_K(E_1, E_3)$ en associant, à chaque couple (v, u) , l'application linéaire $v \circ u$. Cette application vérifie les propriétés suivantes :

(1) pour $u \in \text{Hom}_K(E_1, E_2)$ fixé, l'application

$$v \mapsto v \circ u, \quad \text{Hom}_K(E_2, E_3) \rightarrow \text{Hom}_K(E_1, E_3)$$

est K -linéaire

(2) pour $v \in \text{Hom}_K(E_2, E_3)$ fixé, l'application

$$u \mapsto v \circ u, \quad \text{Hom}_K(E_1, E_2) \rightarrow \text{Hom}_K(E_1, E_3)$$

est K -linéaire.

En effet, vérifions (1) par exemple. Si $x \in E_1$ et $\lambda \in K$, on a :

$$[(\lambda v) \circ u](x) = (\lambda v)(u(x)) = \lambda (v \circ u(x)),$$

d'où $(\lambda v) \circ u = \lambda (v \circ u)$; de même, avec

$$v_i \in \text{Hom}_K(E_2, E_3) \quad (i \in \{1, 2\}),$$

et $x \in E_1$:

$$\begin{aligned} [(v_1 + v_2) \circ u](x) &= (v_1 + v_2)(u(x)) = \\ &= v_1(u(x)) + v_2(u(x)) = v_1 \circ u(x) + v_2 \circ u(x), \end{aligned}$$

$$\text{d'où} \quad (v_1 + v_2) \circ u = v_1 \circ u + v_2 \circ u.$$

La propriété (2) se prouve de même.

Considérons maintenant l'ensemble $\text{Hom}_K(E)$, (où E est un K -ev), dont on sait déjà qu'il est muni d'une structure naturelle de K -ev. Mais, grâce à (\mathcal{L} II), on peut doter $\text{Hom}_K(E)$ d'une nouvelle loi interne, à savoir l'application : $\text{Hom}_K(E) \times \text{Hom}_K(E) \rightarrow \text{Hom}_K(E)$, $(u, v) \mapsto u \circ v$ que nous appellerons simplement **produit** (ou **multiplication**). Alors :

THÉORÈME VI.2.4

|| *L'addition du K -ev $\text{Hom}_K(E)$ et le produit : $(u, v) \mapsto u \circ v$ munissent l'ensemble $\text{Hom}_K(E)$ d'une **structure d'anneau**. L'élément unité de cet anneau est Id_E .*

Démonstration (abrégée) :

Le fait que Id_E est élément neutre pour le produit est évident. Comme $\text{Hom}_K(E)$ est déjà un groupe abélien pour son addition il reste à vérifier que le produit est associatif et qu'il est distributif à droite et à gauche par rapport à l'addition.

L'associativité est une propriété générale de la composition d'applications quelconques (cf. § I.3). La distributivité à droite et à gauche du produit par rapport à la somme est conséquence immédiate des propriétés (1) et (2) ci-dessus. ■

L'ensemble $\text{Hom}_K(E)$ sera systématiquement muni de cette structure d'anneau. Observons que l'anneau $\text{Hom}_K(E)$ est nul ssi $E = \{0_E\}$. En effet Id_E est distinct de l'application nulle ssi $E \neq \{0_E\}$.

Exemple 4 : Pour un corps commutatif K , considéré comme K -ev, l'anneau $\text{Hom}_K(K)$ s'identifie au corps K .

Supposons $E \neq \{0_E\}$. Alors l'anneau $\text{Hom}_K(E)$ est non nul. On se souvient que dans tout anneau non nul les éléments inversibles de l'anneau forment un groupe pour la multiplication (cf. théorème II.5.3). Cela donne ici :

DÉFINITION VI.2.3

{ Soit E un K -ev non nul sur un corps commutatif K . Le groupe des
 { éléments inversibles de l'anneau $\text{Hom}_K(E)$ s'appelle **groupe linéaire**
 { **de E** , et se note $\text{GL}_K(E)$.

Il est clair que $\text{GL}_K(E)$ n'est autre que l'ensemble des automorphismes du K -ev E muni de la loi interne $(u, v) \mapsto u \circ v$ qui définit la structure de groupe de $\text{GL}_K(E)$ (cf. propriétés (\mathcal{L} III) et (\mathcal{L} IV)).

Exemple 5 : Le corps commutatif K étant considéré comme K -ev, le groupe linéaire $\text{GL}_K(K)$ s'identifie au groupe multiplicatif K^* : il suffit d'associer à tout $\lambda \in K^*$ l'homothétie h_λ de K pour obtenir un isomorphisme de groupes de K^* sur $\text{GL}_K(K)$.

Exemple 6 : Soit E un K -ev non nul. Pour $\lambda \in K$, soit h_λ l'homothétie de rapport λ dans E . L'application $H : K \rightarrow \text{Hom}_K(E)$, $\lambda \mapsto h_\lambda$ est un homomorphisme d'anneaux. Il est injectif puisque K est un corps et $\text{Hom}_K(E) \neq \{0\}$. Par restriction à K^* , H définit un *homomorphisme injectif de groupes* $H^* : K^* \rightarrow \text{GL}_K(E)$, dont l'image est le sous-groupe de $\text{GL}_K(E)$ formé des homothéties de rapport non nul. Le sous-anneau $H(K)$ de $\text{Hom}_K(E)$, isomorphe au corps K , s'appelle **corps des homothéties de E** . Le sous-groupe $H(K^*)$ de $\text{GL}_K(E)$, isomorphe au groupe multiplicatif K^* , s'appelle **groupe des homothéties de E** (ici, on sous-entend : de rapport $\neq 0$). En général, on a bien entendu $H(K^*) \subsetneq \text{GL}_K(E)$.

Exercice 1 : Soit E un K -ev et $f \in \text{Hom}_K(E)$. On pose $f \circ f = f^2$. Montrer que

$$\text{Ker}(f) = \text{Ker}(f^2) \quad \text{ssi} \quad \text{Im}(f) \cap \text{Ker}(f) = \{0_E\}.$$

Exercice 2 : Soit f et g deux endomorphismes de l'espace vectoriel E . Montrer que

$$f[\text{Ker}(g \circ f)] = \text{Ker}(g) \cap \text{Im}(f).$$

Exercice 3 : Soit E le \mathbb{R} -ev des fonctions de classe $\mathcal{C}^\infty(\mathbb{R})$ et 2π -périodiques. A $f \in E$ on associe $Tf = f''$; T est un endomorphisme de E . Chercher son noyau et son image.

Exercice 4 : Soit u et v deux endomorphismes, avec u inversible, du K -ev E tels que $u \circ v = v \circ u$ et $v^4 = 0$. Montrer que u est inversible et que $(u^{-1} \circ v)^4 = 0$. Montrer que $u + v$ est inversible.

Exercice 5 : Soit E_1, E_2, E_3 trois sous-espaces d'un espace vectoriel E sur le corps commutatif K . On considère $f \in \text{Hom}_K(E_1, E_2)$, $g \in \text{Hom}_K(E_1, E_3)$ et on suppose que g est surjective. Prouver que $\text{Ker}(f) \supset \text{Ker}(g) \Leftrightarrow \exists h \in \text{Hom}_K(E_3, E_2)$ tel que $f = h \circ g$.

Exercice 6 : On définit dans $\text{Hom}_K(E)$ une relation binaire \mathcal{R} par $f \mathcal{R} g$ ssi $\text{Ker}(f) = \text{Ker}(g)$. Montrer que \mathcal{R} est une relation d'équivalence. On pose maintenant $f \mathcal{S} g$ ssi $\text{Ker}(f) \subset \text{Ker}(g)$. Obtient-on une relation d'ordre ?

Exercice 7 : Soit K un corps commutatif de cardinal ≥ 4 , E un K -ev et F un sous- K -ev de E distinct de E . On considère une application $f \in \text{Hom}_K(E)$ telle que

$$(\forall x \in E \setminus F) \quad \exists \lambda_x \in K \mid f(x) = \lambda_x \cdot x.$$

Prouver que f est une homothétie.

Exercice 8 : Vérifier que le groupe des homothéties d'un K -ev E est distingué dans $\text{GL}_K(E)$.

§ VI.3 COMBINAISONS LINÉAIRES ; INDÉPENDANCE LINÉAIRE ; BASES

Dans ce paragraphe, le corps commutatif K est fixé.

DÉFINITION VI.3.1

Soit $(x_i)_{i \in I}$ une famille non vide d'éléments d'un K -ev E . On appelle **combinaison K -linéaire** des x_i tout vecteur x de E de la forme $\sum_{i \in I} \lambda_i x_i$, où $\lambda = (\lambda_i)_{i \in I}$ est une famille à support fini d'éléments de K , c'est-à-dire un élément de $K^{(I)}$. Les λ_i sont appelés les **coefficients** de la combinaison linéaire.

Cette définition est cohérente car, pour $\lambda \in K^{(I)}$, la famille $(\lambda_i x_i)_{i \in I}$ est à support fini dans E .

Si aucune ambiguïté n'est à craindre sur K , on parlera de **combinaisons linéaires**.

La famille $x = (x_i)_{i \in I}$ étant fixée dans E , notons $\mathcal{C}_K(x)$ l'ensemble de toutes les combinaisons K -linéaires des x_i . On obtient :

THÉORÈME VI.3.1

L'ensemble $\mathcal{C}_K(x)$ est un sous- K -ev de E . Pour l'inclusion, c'est le plus petit sous- K -ev de E contenant chaque x_i . Autrement dit, c'est l'intersection des sous- K -ev de E contenant tous les x_i (cf. théorème VI.1.3).

Démonstration :

Il est évident que si un sous- K -ev F de E contient chaque x_i , il contient nécessairement les combinaisons linéaires des x_i , donc $\mathcal{C}_K(x) \subset F$. Or, l'application $K^{(I)} \rightarrow E$, $(\lambda_i) \mapsto \sum_{i \in I} \lambda_i x_i$ est K -linéaire, comme on voit aisément. Son image est donc un sous- K -ev de E . Or c'est justement $\mathcal{C}_K(x)$. ■

DÉFINITION VI.3.2

Si $(x_i)_{i \in I}$ est une famille non vide dans le K -ev E , on appelle **sous-espace vectoriel de E engendré par $(x_i)_{i \in I}$** le sous- K -ev $\mathcal{C}_K(x)$ des combinaisons linéaires des x_i . Si A est une partie non vide de E , on appelle **sous-espace engendré par A** le sous-espace engendré par la famille canoniquement associée à A (cf. § I.4).

Cela nous conduit à préférer la notation $\text{Vect}_K((x_i)_{i \in I})$ ou $\text{Vect}((x_i)_{i \in I})$ à la place de $\mathcal{C}_K((x_i)_{i \in I})$. Pour $A \subset E$ on notera de même $\text{Vect}_K(A)$ ou $\text{Vect}(A)$ le sous-espace engendré par A , c'est-à

ble des vecteurs $\sum_{a \in A} \lambda_a \cdot a$, où $(\lambda_a)_{a \in A}$ décrit $K^{(A)}$. On convient que le sous- K -ev de E engendré par la famille vide de E , ou par la partie vide de E , est $\{0_E\}$. Moyennant quoi on vérifie les propriétés suivantes :

(CL1) Soit A une partie du K -ev E ; alors $A = \text{Vect}_K(A)$ ssi A est un sous- K -ev de E .

(CL2) Donc, si $A \subset E$, on a : $\text{Vect}_K(\text{Vect}_K(A)) = \text{Vect}_K(A)$.

DÉFINITION VI.3.3

Soit $(x_i)_{i \in I}$ une famille dans un K -ev E , avec $I \neq \emptyset$.

(I) On dit que les $(x_i)_{i \in I}$ sont **linéairement indépendants** sur K (ou que la famille (x_i) est **K -libre**) ssi la seule famille $(\lambda_i)_{i \in I} \in K^{(I)}$ telle que $\sum_{i \in I} \lambda_i x_i = 0_E$ est la famille nulle (i.e. telle que $\forall i \in I$, $\lambda_i = 0_K$). Dans le cas contraire on dit que les (x_i) sont **linéairement dépendants** sur K (ou **K -liés**), et on appelle alors **relation de dépendance linéaire** entre les (x_i) toute relation du type $\sum_{i \in I} \lambda_i x_i = 0_E$, avec $(\lambda_i) \in K^{(I)}$ et les λ_i non tous nuls.

(II) On dit que les (x_i) **engendrent** E ssi $\text{Vect}_K((x_i)_{i \in I}) = E$.

(III) On dit que la famille $(x_i)_{i \in I}$ est une **base** du K -ev E ssi elle est **K -libre** et les (x_i) **engendrent** le K -ev E .

On complète cette définition en convenant que la famille vide de E est libre. En revanche une famille contenant le vecteur 0_E est nécessairement liée (même si c'est un singleton !).

Fixons une famille non vide $x = (x_i)_{i \in I}$ dans le K -ev E . A cette famille associons l'application K -linéaire $f_x : K^{(I)} \rightarrow E$, $(\lambda_i) \mapsto \sum_{i \in I} \lambda_i x_i$. La défini-

tion VI.3.3 peut alors se traduire ainsi : **pour que les (x_i) soient K -linéairement indépendants, il faut et il suffit que $\text{Ker}(f_x) = \{0_{K^{(I)}}\}$, c'est-à-dire que f_x soit injective.**

Pour que les (x_i) engendrent le K -ev E , il faut et il suffit que f_x soit surjective. En combinant ces deux assertions, on obtient :

Pour que les (x_i) forment une base du K -ev E , il faut et il suffit que f_x soit un isomorphisme de K -ev.

L'importance de la notion de **base** d'un espace vectoriel E provient de la très simple remarque suivante : lorsque les $(x_i)_{i \in I}$ forment une base, tout élément $v \in E$ s'écrit **d'une manière et d'une seule** sous la forme $v = \sum_{i \in I} \lambda_i x_i$ avec $(\lambda_i) \in K^{(I)}$, c'est-à-dire comme combinaison linéaire **finie**

de vecteurs de la base. En particulier, pour chaque $i \in I$, on a une application $\psi_i : E \rightarrow K$ qui associe, à chaque $v \in E$, le terme d'indice i de la famille $(\lambda_j) \in K^{(I)}$ telle que $\sum_{j \in I} \lambda_j x_j = v$. Cette application ψ_i est appelée la

i -ième fonction coordonnée dans la base $(x_i)_{i \in I}$. Si l'on note $p_i : K^{(I)} \rightarrow K$ la restriction à $K^{(I)}$ de la i -ième projection $K^I \rightarrow K$, qui est K -linéaire, on voit que $\psi_i = p_i \circ f_x^{\langle -1 \rangle}$, ce qui montre que ψ_i est K -linéaire. Il est important d'observer que, pour chaque $i \in I$, ψ_i dépend de toute la famille $(x_j)_{j \in I}$.

Lorsqu'on applique tous les concepts ci-dessus à la famille canoniquement associée à une partie A de E , on obtient les notions de **partie libre de E** (ou partie K -linéairement indépendante de E), de **partie liée de E** , de **partie génératrice de E** , et de **partie-base du K -ev E** . On convient que \emptyset est une partie libre de E .

Exemple 1 : Considérons le K -ev K . Les parties-base en sont les singletons $\{x\}$, où $x \in K$ et $x \neq 0_K$.

Les propriétés suivantes se vérifient immédiatement :

(CL3) Soit $(x_i)_{i \in I}$ une famille K -linéairement indépendante du K -ev E . Alors l'application $i \mapsto x_i$ est injective (en effet, si $i \in I$, $j \in I$, $i \neq j$, la famille $\lambda = (\lambda_K)_{K \in I} \in K^{(I)}$ telle que $\lambda_K = 0_K$ si $k \notin \{i, j\}$, $\lambda_i = 1_K$, $\lambda_j = -1_K$, n'est pas nulle, donc $x_i - x_j = f_x(\lambda) \neq 0_E$). On remarque que chaque x_i est $\neq 0_E$ (sinon $1_K \cdot x_i = 0_E$ serait une relation de dépendance linéaire entre les (x_j)).

(CL4) Soit E_1 et E_2 deux K -ev, $f \in \text{Hom}_K(E_1, E_2)$ et $(x_i)_{i \in I}$ une famille dans E_1 . Si les $(x_i)_{i \in I}$ sont liés, les $(f(x_i))_{i \in I}$ le sont aussi. Donc si les $(f(x_i))_{i \in I}$ sont K -libres, les $(x_i)_{i \in I}$ sont aussi K -libres.

(CL5) Soit $(x_i)_{i \in I}$ une famille libre dans le K -ev E , et soit J une partie de I . Alors la famille $(x_i)_{i \in J}$ est libre elle aussi.

Exemple 2 : Soit I un ensemble non vide. Dans le K -ev $K^{(I)}$ considérons la famille $(e_i)_{i \in I}$ suivante ; pour chaque $i \in I$, e_i est l'élément $(\delta_{ij})_{j \in I}$ de $K^{(I)}$ ainsi défini : $\delta_{ii} = 1_K$ et $\delta_{ij} = 0_K$ si $j \neq i$ (δ_{ij} est appelé **symbole de Kronecker** ⁽¹⁾). En revenant aux définitions, on voit immédiatement que $(e_i)_{i \in I}$ est une base du K -ev $K^{(I)}$. Cette base est appelée la **base canonique** de $K^{(I)}$.

THÉORÈME VI.3.2

Soit $(e_i)_{i \in I}$ une base du K -ev E_1 , avec $I \neq \emptyset$. Si E_2 est un K -ev, et si $(b_i)_{i \in I}$ est une famille arbitraire, indexée par I , de vecteurs de E_2 , il existe une et une seule application K -linéaire $f : E_1 \rightarrow E_2$ telle que $(\forall i \in I) f(e_i) = b_i$.

De plus f est injective ssi $(b_i)_{i \in I}$ est K -libre.

f est surjective ssi $(b_i)_{i \in I}$ est génératrice du K -ev E_2 .

f est bijective ssi $(b_i)_{i \in I}$ est une base de E_2 .

⁽¹⁾ Leopold Kronecker (1823-1891), mathématicien allemand, a effectué des travaux sur les fonctions elliptiques, les nombres algébriques, l'algèbre linéaire et multilinéaire.

Démonstration :

Soit $\Phi : K^{(I)} \rightarrow E_1, (\lambda_i) \mapsto \sum_{i \in I} \lambda_i e_i$. On sait que Φ est un isomorphisme de K -ev. Soit de même $\Psi : K^{(I)} \rightarrow E_2, (\lambda_i) \mapsto \sum_{i \in I} \lambda_i b_i$; Ψ est K -linéaire. Si f existe, pour tout $x = \sum_{i \in I} \lambda_i e_i = \Phi((\lambda_i))$ de E_1 , on a :

$$f(x) = \sum_{i \in I} \lambda_i f(e_i) = \sum_{i \in I} \lambda_i b_i = \Psi((\lambda_i)), \text{ donc } f = \Psi \circ \Phi^{-1}.$$

Réciproquement, $\Psi \circ \Phi^{-1}$ est K -linéaire et envoie e_i sur b_i pour tout $i \in I$. Donc f existe et est unique et vaut $\Psi \circ \Phi^{-1}$. On sait que (b_i) est K -libre ssi Ψ est injective, K -génératrice ssi Ψ est surjective, et que (b_i) est une base ssi Ψ est bijective. Puisque Φ est bijective, les dernières assertions du théorème en résultent. ■

Avec les notations du théorème VI.3.2, on dit que f s'obtient en **prolongeant par K -linéarité** les relations $(\forall i \in I) f(e_i) = b_i$.

Exercice 1 : Soit K un corps commutatif et I un ensemble *infini*. Montrer que les K -ev $K^{(I)}$ et K^I n'admettent aucune partie génératrice finie.

Exercice 2 :

a) Pour $\alpha \in \mathbb{R}$, soit $f_\alpha : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|^\alpha$. Démontrer que la famille de fonctions $(f_\alpha)_{\alpha \in \mathbb{R}}$ est \mathbb{R} -libre dans le \mathbb{R} -ev $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Indication : Penser à dériver $\sum_{i=1}^n \lambda_i f_{\alpha_i}(x)$ pour $x > 0$; procéder par récurrence sur n .

b) Soit $I = \mathbb{R} \setminus \mathbb{N}$. Démontrer que la famille de fonctions $(g_{a,\alpha})_{(a,\alpha) \in \mathbb{R} \times I}$, où $g_{a,\alpha}(x) = f_\alpha(x-a)$ pour $x \in \mathbb{R}$, est \mathbb{R} -libre dans le \mathbb{R} -ev $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Exercice 3 : Démontrer que le \mathbb{Q} -ev \mathbb{R} n'admet pas de partie génératrice finie.

Exercice 4 : On considère le \mathbb{R} -ev $E = \mathbb{R}^{\mathbb{N}}$. Si $x \in \mathbb{R}_+^*$ soit S_x la suite $(x^n)_{n \in \mathbb{N}}$. En utilisant le fait que $x^n \rightarrow 0$ quand $n \rightarrow \infty$ pour $x \in]0, 1[$, démontrer que la famille $(S_x)_{x > 0}$ est \mathbb{R} -libre dans E .

Exercice 5 : Soit E le \mathbb{R} -ev $\mathcal{F}(\mathbb{R}_+, \mathbb{R})$. Si $\alpha \in \mathbb{R}$, soit \mathcal{E}_α la fonction $\mathbb{R}_+ \rightarrow \mathbb{R}, t \mapsto \exp(\alpha t)$. Prouver que la famille $(\mathcal{E}_\alpha)_{\alpha \in \mathbb{R}}$ est \mathbb{R} -libre dans E .

Exercice 6 : Dans un K -ev E on considère une famille $(L_i)_{i \in I}$ de parties libres telles que :

$$\forall i \in I, \forall j \in I, \exists k \in I \mid L_i \cup L_j \subset L_k.$$

Démontrer que $\bigcup_{i \in I} L_i$ est encore une partie libre de E .

Exercice 7 : Soit deux réels α, β ($\alpha < \beta$). On considère le \mathbb{R} -ev E des fonctions $[\alpha, \beta] \rightarrow \mathbb{R}$ continues et affines par morceaux. Si $a \in [\alpha, \beta]$, soit

$$f_a(x) = |x - a| \quad (x \in [\alpha, \beta]).$$

Montrer que la famille $(f_a)_{a \in [\alpha, \beta]}$ est une base du \mathbb{R} -ev E .

Exercice 8 : On considère \mathbb{R} comme un \mathbb{Q} -ev. Montrer que les vecteurs $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ sont linéairement indépendants. Même question avec $1, \sqrt[3]{2}, \sqrt[3]{4}$.

Exercice 9 : Soit K un corps de caractéristique $\neq 2$ et E un K -ev. Si (v_1, v_2, \dots, v_n) est une famille libre, en est-il de même de $(u_1 + u_2, u_2 + u_3, \dots, u_n + u_1)$?

Exercice 10 : On donne $(p, q) \in \mathbb{R}^2$ et le \mathbb{R} -ev $\mathcal{F}(\mathbb{R}, \mathbb{R})$. Trouver une relation de dépendance linéaire entre les 3 fonctions

$$f : x \mapsto \sin x, \quad g : x \mapsto \sin(x + p) \quad \text{et} \quad h : x \mapsto \cos(x + q).$$

§ VI.4. STRUCTURE D'ALGÈBRE SUR UN CORPS COMMUTATIF

DÉFINITION VI.4.1

Soit K un corps commutatif. On appelle K -algèbre tout K -ev \mathcal{A} muni d'une application **K -bilinéaire**, c'est-à-dire d'une application $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$, $(x, y) \mapsto x * y$, telle que, pour tous éléments x, x_1, x_2, y_1, y_2, y de \mathcal{A} et tous éléments $\lambda_1, \lambda_2, \mu_1, \mu_2$ de K on ait :

$$(\lambda_1 x_1 + \lambda_2 x_2) * y = \lambda_1 (x_1 * y) + \lambda_2 (x_2 * y)$$

et

$$x * (\mu_1 y_1 + \mu_2 y_2) = \mu_1 (x * y_1) + \mu_2 (x * y_2).$$

L'application K -bilinéaire notée $*$ est alors appelée le **produit** de cette K -algèbre. Une K -algèbre \mathcal{A} est dite à **élément unité** (ou **unifère**) ssi son produit admet un élément neutre $e_{\mathcal{A}}$, qui est alors appelé **unité** de \mathcal{A} .

Elle est dite **associative** (resp. **commutative**) ssi son produit est associatif (resp. commutatif).

Exemple 1 : Si L est un anneau, et si K est un sous-anneau de L qui soit un corps commutatif, alors L est de manière naturelle une K -algèbre \hat{L} : le produit de \hat{L} est le produit dans l'anneau L ; l'addition de \hat{L} est celle de l'anneau L ; si $x \in L$ et $\lambda \in K$ le produit du scalaire λ et du vecteur x est le produit λx dans L . Cette K -algèbre, notée simplement L , sera appelée *la K -algèbre L* . Elle est associative et à élément unité 1_L .

De manière générale, lorsqu'une K -algèbre \mathcal{A} est *associative et unifère*, l'addition et le produit de \mathcal{A} définissent sur \mathcal{A} une **structure d'anneau**, dite **sous-jacente** à la structure de K -algèbre de \mathcal{A} .

Une K -algèbre \mathcal{A} est dite **nulle** ssi $\mathcal{A} = \{0_{\mathcal{A}}\}$. Si \mathcal{A} est unifère, d'unité e , alors \mathcal{A} est nulle ssi $e = 0_{\mathcal{A}}$.

Exemple 2 : Soit E un espace vectoriel sur un corps commutatif K . Muni de sa structure de K -ev, l'anneau $\text{Hom}_K(E)$ devient une K -algèbre unifère et associative. L'unité est Id_E . Cette algèbre est non nulle ssi $E \neq \{0_E\}$. La catégorie d'algèbres ainsi obtenue joue un rôle essentiel dans la théorie des algèbres.

Exemple 3 : Soit E un espace vectoriel non nul sur un corps commutatif K . Considérons $\text{Hom}_K(E)$ comme un K -ev, et munissons-le du produit suivant, appelé *crochet* :

$$(x, y) \mapsto [x, y] = x \circ y - y \circ x.$$

Alors $\text{Hom}_K(E)$ devient une K -algèbre. Cette algèbre n'est ni commutative, ni associative, ni unifère, et pourtant elle présente un intérêt majeur pour l'étude approfondie des espaces vectoriels (cf. [3]).

Exemple 4 : Soit \mathcal{A} une K -algèbre et soit C un sous-corps de K ; en restreignant à C les scalaires, on sait qu'on obtient sur \mathcal{A} une structure de C -ev ; le produit de \mathcal{A} et cette structure de C -ev, définissent sur \mathcal{A} une structure de C -algèbre, dite **obtenue par restriction des scalaires à C** à partir de \mathcal{A} . Si aucune confusion n'est à craindre, on la notera encore \mathcal{A} , sinon, on peut la noter $\mathcal{A}_{(C)}$.

Convention. Dans la suite de cet ouvrage, et sauf mention explicite du contraire, le terme « K -algèbre » devra être pris au sens restreint de K -algèbre **unifère et associative**. Avec cette convention toute K -algèbre pourra être considérée (en « oubliant » la multiplication par les scalaires du corps commutatif K) comme un anneau.

Exemple 5 : Soit I un ensemble non vide et K un corps commutatif. Munissons le K -ev $\mathcal{F}(I, K) = K^I$ du produit suivant : si $f \in K^I$ et $g \in K^I$, $f \cdot g$ est l'application $I \rightarrow K, i \mapsto f(i) g(i)$. Alors $\mathcal{F}(I, K)$ devient une K -algèbre commutative. Son élément unité e est la fonction $I \rightarrow K, i \mapsto 1_K$. Si I n'est pas un singleton, cette K -algèbre n'est pas intègre (une K -algèbre est dite **intègre** ssi en tant qu'anneau, c'est un anneau intègre ; elle est dite **sans diviseur de zéro** ssi en tant qu'anneau, elle n'a pas de diviseur de zéro). Une attention particulière peut être accordée au cas où le corps K est le corps à deux éléments $\mathbb{Z}/2\mathbb{Z} = \{\bar{0}, \bar{1}\}$.

Exemple 6 : Soit M un monoïde, d'élément neutre ε , et soit K un corps commutatif. Notons $(e_\mu)_{\mu \in M}$ la base canonique du K -ev $K^{(M)}$. Définissons sur $K^{(M)}$ un produit, noté $*$, de la manière suivante : si $x = \sum_{\mu \in M} x_\mu e_\mu$ et si

$$y = \sum_{\mu \in M} y_\mu e_\mu, \text{ on pose } x * y = \sum_{\mu \in M} z_\mu e_\mu, \text{ où, pour chaque } \mu \in M :$$

$$z_\mu = \sum_{\alpha\beta = \mu} x_\alpha y_\beta. \text{ Cela a bien un sens car :}$$

a) pour μ fixé, la famille $(x_\alpha y_\beta)_{(\alpha, \beta) \in M \times M}$ est à support fini (contenu dans $\text{supp}((x_\alpha)) \times \text{supp}((y_\beta))$), donc z_μ est bien défini, et :

b) la famille $(z_\mu)_{\mu \in M}$ est à support fini (contenu dans $\{\alpha\beta\}_{(\alpha, \beta) \in \text{supp}((x_\alpha)) \times \text{supp}((y_\beta))}$). Le produit $(x, y) \mapsto x * y$ ainsi défini sur $K^{(M)}$ munit ce K -ev d'une structure de K -algèbre, dite **K -algèbre du monoïde M** , et généralement notée $K^{[M]}$. Lorsque M est un groupe, la K -algèbre $K^{[M]}$ est appelée **algèbre du groupe M** .

DÉFINITION VI.4.2

Soit K un corps commutatif et \mathcal{A} et \mathcal{B} deux K -algèbres. On appelle **homomorphisme de K -algèbres** de \mathcal{A} dans \mathcal{B} toute application K -linéaire $f : \mathcal{A} \rightarrow \mathcal{B}$ qui est aussi un **homomorphisme d'anneaux** de \mathcal{A} dans \mathcal{B} .

On appelle **isomorphisme de K -algèbres** de \mathcal{A} sur \mathcal{B} toute bijection $f : \mathcal{A} \rightarrow \mathcal{B}$ telle que f et f^{-1} soient des homomorphismes de K -algèbres. Les K -algèbres \mathcal{A} et \mathcal{B} sont dites **isomorphes** ssi il existe au moins un isomorphisme de K -algèbres de \mathcal{A} sur \mathcal{B} . Un isomorphisme de K -algèbres de \mathcal{A} sur \mathcal{A} est appelé un **automorphisme** de la K -algèbre \mathcal{A} .

Pour toute K -algèbre \mathcal{A} , $\text{Id}_{\mathcal{A}}$ est un automorphisme de la K -algèbre \mathcal{A} . La composée de deux homomorphismes de K -algèbres en est un.

Si $f : \mathcal{A} \rightarrow \mathcal{B}$ est un homomorphisme *bijectif* de K -algèbres, alors c'est un isomorphisme de K -algèbres, et $f^{-1} : \mathcal{B} \rightarrow \mathcal{A}$ en est aussi un.

La relation « Les K -algèbres \mathcal{A} et \mathcal{B} sont isomorphes » est réflexive, symétrique et transitive.

Exemple 7 : Si \mathcal{A} est une K -algèbre non nulle, d'élément unité e , l'application $K \rightarrow \mathcal{A}$, $\lambda \mapsto \lambda e$ est un homomorphisme de K -algèbres (dit *canonique*). En effet, elle est K -linéaire, envoie 1_K sur e , et si $\lambda \in K$ et $\mu \in K$, on a : $(\lambda e) * (\mu e) = (\lambda \mu) e$ par K -bilinearité du produit de \mathcal{A} , et parce que $e * e = e$. Puisque K est un corps, cet homomorphisme canonique est injectif.

DÉFINITION VI.4.3

Soit K un corps commutatif et \mathcal{A} une K -algèbre. On appelle **sous- K -algèbre** de \mathcal{A} tout sous- K -ev \mathcal{B} de \mathcal{A} qui est en même temps un sous-anneau de \mathcal{A} .

L'intersection d'une famille de sous- K -algèbres de \mathcal{A} est une sous- K -algèbre de \mathcal{A} . Si $f : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ est un homomorphisme de K -algèbres, alors pour toute sous- K -algèbre \mathcal{B}_1 de \mathcal{A}_1 (resp. \mathcal{B}_2 de \mathcal{A}_2), $f(\mathcal{B}_1)$ (resp. $f^{-1}(\mathcal{B}_2)$) est une sous- K -algèbre de \mathcal{A}_2 (resp. de \mathcal{A}_1).

Exemple 8 : Soit \mathcal{A} une K -algèbre non nulle, d'élément unité e . L'ensemble, noté Ke , image de l'homomorphisme canonique $K \rightarrow \mathcal{A}$, est une sous- K -algèbre de \mathcal{A} ; c'est l'intersection de toutes les sous- K -algèbres de \mathcal{A} , autrement dit, la plus petite (au sens de l'inclusion) sous- K -algèbre de \mathcal{A} . Cette K -algèbre, en tant qu'anneau, est un corps isomorphe à K , l'application $K \rightarrow Ke$, $\lambda \mapsto \lambda e$ étant un isomorphisme de corps, qui permet, si on veut, d'identifier K et Ke .

Exemple 9 : Soit $(x_i)_{i \in I}$ une famille non vide d'éléments d'une K -algèbre \mathcal{A} . L'intersection des sous- K -algèbres de \mathcal{A} qui contiennent

x_i est une **sous- K -algèbre de \mathcal{A}** dite **engendrée par les x_i** . (En appliquant cette notion à la famille associée à une partie de \mathcal{A} , on obtient la notion de sous- K -algèbre **engendrée par une partie de \mathcal{A}**). On vérifie aisément que la sous- K -algèbre engendrée par $(x_i)_{i \in I}$ est le sous- K -ev de \mathcal{A} engendré par les éléments suivants : e , et $x_{i_1} * x_{i_2} * \dots * x_{i_n}$, où $n \in \mathbb{N}^*$ est arbitraire, et où (i_1, i_2, \dots, i_n) est une n -suite arbitraire à valeurs dans I . En particulier, si $x \in \mathcal{A}$ est fixé, la sous- K -algèbre de \mathcal{A} engendrée par x est le sous- K -ev engendré par la famille $(x^n)_{n \in \mathbb{N}}$ (où $x^0 = e$, et $x^n = x^{n-1} * x$ pour $n \in \mathbb{N}^*$). Cette sous- K -algèbre sera notée $K[x]$: c'est donc l'ensemble des combinaisons linéaires $\sum_{n \in \mathbb{N}} \lambda_n x^n$, où (λ_n) parcourt $K^{(\mathbb{N})}$. C'est une sous- K -algèbre *commutative* de \mathcal{A} .

DÉFINITION VI.4.4

Une K -algèbre \mathcal{A} est dite **de type fini** ssi il existe une famille finie $(x_i)_{i \in I}$ qui engendre \mathcal{A} comme K -algèbre, i.e. telle que la sous- K -algèbre engendrée par $(x_i)_{i \in I}$ soit égale à \mathcal{A} .

Idéaux d'une K -algèbre

On a parfois besoin en Algèbre linéaire de la notion d'idéal présentée ci-dessous.

DÉFINITION VI.4.5

Soit A un anneau, non nécessairement commutatif. Une partie \mathfrak{a} de A est appelée un **idéal à gauche** (resp. **à droite**) ssi c'est un sous-groupe additif de A , et

$$\forall \lambda \in A, \quad \forall x \in \mathfrak{a}, \quad \lambda x \in \mathfrak{a} \quad (\text{resp. } x\lambda \in \mathfrak{a}).$$

On dit que \mathfrak{a} est un **idéal bilatère** de A ssi c'est à la fois un idéal à droite et un idéal à gauche.

Soit K un corps commutatif et \mathcal{A} une K -algèbre : un **idéal à gauche** de \mathcal{A} (resp. idéal à droite, idéal bilatère) est un idéal à gauche (resp. à droite, bilatère) de l'anneau \mathcal{A} . Il est clair que tout idéal de \mathcal{A} (à gauche, à droite ou bilatère) est un sous- K -ev de \mathcal{A} .

Revenant à un anneau A , la famille des idéaux à gauche (resp. à droite, bilatères) de A est stable par intersection quelconque ; $\{0_A\}$ et A sont toujours des idéaux bilatères de A .

Soit $\rho : A \rightarrow B$ un homomorphisme d'anneaux *surjectif* : pour tout idéal \mathfrak{a} de A (à gauche, à droite ou bilatère), $\rho(\mathfrak{a})$ est un idéal de B , de même nature ; ρ n'étant plus nécessairement surjectif, pour tout idéal \mathfrak{b} de B (à gauche, à droite ou bilatère), $\rho^{-1}(\mathfrak{b})$ est un idéal de même nature dans A . En particulier $\text{Ker}(\rho)$ est un idéal bilatère de A .

Exemple 10 : Soit E un espace vectoriel non nul sur un corps commutatif K , et soit F un sous- K -ev de E . L'ensemble $\{u \in \text{Hom}_K(E) \mid \text{Im}(u) \subset F\}$ est un idéal à droite de la K -algèbre $\text{Hom}_K(E)$. L'ensemble $\{u \in \text{Hom}_K(E) \mid F \subset \text{Ker}(u)\}$ est un idéal à gauche de cette K -algèbre.

Bien entendu dans le cas d'un anneau commutatif on retrouve la notion d'idéal déjà présentée au § III.7.

Exercice 1 : Soit K un corps commutatif et \mathcal{A} une K -algèbre (le produit est noté sans signe).

a) Montrer que, si E est une partie non vide de \mathcal{A} , l'ensemble

$$\mathcal{C}(E) = \{x \in \mathcal{A} \mid \forall y \in E \quad xy = yx\}$$

est une sous- K -algèbre de \mathcal{A} . Si $E_1 \subset E_2 \subset \mathcal{A}$, comparer $\mathcal{C}(E_1)$ et $\mathcal{C}(E_2)$. Si $(E_i)_{i \in I}$ est une famille de parties de \mathcal{A} , quelle est la sous- K -algèbre $\mathcal{C}\left(\bigcup_{i \in I} E_i\right)$? Enfin, si $E \subset \mathcal{A}$, comparer

$\mathcal{C}(\mathcal{C}(E))$ et la sous- K -algèbre \hat{E} engendrée par E .

b) On appelle **centre** de \mathcal{A} la sous- K -algèbre $\mathcal{C}(\mathcal{A})$. Si e est l'unité de \mathcal{A} , on a : $Ke \subset \mathcal{C}(\mathcal{A})$. Si M est un groupe, déterminer le centre de la K -algèbre $K^{[M]}$ définie dans l'exemple 6. Condition sur M pour que cette K -algèbre soit commutative.

Exercice 2 : Soit I un ensemble infini. Démontrer que la \mathbb{R} -algèbre $\mathcal{F}(I, \mathbb{R})$ n'est pas de type fini. Reprendre la question en remplaçant \mathbb{R} par un corps commutatif arbitraire.

Exercice 3 : Trouver les idéaux à droite, à gauche, et bilatères, dans la K -algèbre $K^{[\mathbb{S}_3]}$, puis $K^{[u_4]}$, où K est un corps commutatif.

Exercice 4 : a) Soit K un corps commutatif et \mathcal{A} une K -algèbre ; pour tout élément inversible σ de \mathcal{A} , vérifier que $f_\sigma : \mathcal{A} \rightarrow \mathcal{A}$, $x \mapsto \sigma x \sigma^{-1}$, est un automorphisme de la K -algèbre \mathcal{A} (dit *intérieur*).

b) On suppose que $\mathcal{A} = K^{[M]}$, où M est un groupe. Pour tout automorphisme φ du groupe M , on définit

$$f_\varphi : \mathcal{A} \rightarrow \mathcal{A}, x = \sum_{g \in M} x_g \cdot g \mapsto f_\varphi(x) = \sum_{g \in M} x_{\varphi^{-1}(g)} \cdot g.$$

Vérifier que f_φ est un automorphisme de la K -algèbre \mathcal{A} , et qu'on a : $f_{\varphi \circ \psi} = f_\varphi \circ f_\psi$ pour tous $\varphi, \psi \in \text{Aut}(M)$. Montrer que si φ est un automorphisme intérieur du groupe M , alors f_φ est un automorphisme intérieur de la K -algèbre \mathcal{A} . Montrer que l'application $f \mapsto f_\varphi$ est injective.

Exercice 5 : Soit K un corps commutatif et E un ensemble fini, de cardinal $n \geq 1$. On note \mathcal{A} la K -algèbre $\mathcal{F}(E, K)$.

a) Soit \mathcal{R} une relation d'équivalence sur E ; montrer que l'ensemble $\mathcal{B}_{\mathcal{R}}$ des $f \in \mathcal{A}$ constantes sur chaque classe (mod \mathcal{R}) est une sous- K -algèbre de \mathcal{A} .

b) Montrer que $\mathcal{R} \mapsto \mathcal{B}_{\mathcal{R}}$ est une bijection de l'ensemble des relations d'équivalence de E sur l'ensemble des sous- K -algèbres de \mathcal{A} .

Exercice 6 : Soit K un corps commutatif et E un ensemble fini de cardinal $n \geq 1$. On considère l'ensemble \mathcal{M} des *homomorphismes de K -algèbres* $\mathcal{A} \rightarrow K$, où \mathcal{A} est la K -algèbre $\mathcal{F}(E, K)$. Pour chaque $a \in E$, l'application $\chi_a : \mathcal{A} \rightarrow K$, $f \mapsto f(a)$ appartient à \mathcal{M} . Montrer que $E \rightarrow \mathcal{M}$, $a \mapsto \chi_a$, est une bijection. Si u est un automorphisme de la K -algèbre \mathcal{A} , que peut-on dire de $\chi_a \circ u$? En déduire que le groupe des automorphismes de la K -algèbre \mathcal{A} est isomorphe à \mathbb{S}_n .

Exercice 7 : On sait qu'il existe un seul groupe G (à un isomorphisme près) e_0, e_1, e_2 . Considérons e_0, e_1, e_2 comme la base d'un espace vectoriel E sur le

(E est l'ensemble des vecteurs $x = x_0 e_0 + x_1 e_1 + x_2 e_2$). Munissons E du produit

$$x * y = \sum_{\substack{i \in \coprod_{j \in \coprod_{k \in \{0,2\}} \mathbb{Z}}} x_i y_j e_i \cdot e_j$$

(où $e_i \cdot e_j$ s'effectue dans le groupe G). Montrer qu'on obtient ainsi une \mathbb{R} -algèbre \mathcal{A} de type fini qui est associative, commutative et unifère. Caractériser le groupe U des éléments inversibles de cette algèbre ainsi que l'ensemble des diviseurs de zéro.

Peut-on construire une \mathbb{R} -algèbre engendrée par trois éléments linéairement indépendants qui soit sans diviseur de zéro ?

Exercice 8 : Soit $(1, \varepsilon)$ la base d'un espace vectoriel E sur \mathbb{R} . On définit une multiplication dans E d'élément neutre 1 en posant

$$\varepsilon * \varepsilon = q + p\varepsilon \quad (p \in \mathbb{R}, q \in \mathbb{R}),$$

et en étendant par linéarité la multiplication à deux vecteurs quelconques de E .

a) Etudier les cas particuliers $\varepsilon^2 = -1$; $\varepsilon^2 = 0$; $\varepsilon^2 = +1$.

b) Montrer que pour p et q quelconques on obtient une algèbre isomorphe à l'une des trois étudiées au a).

§ VI.5 LE CORPS DES NOMBRES COMPLEXES

Construction de \mathbb{C}

Nous désignons par \mathbb{C} le \mathbb{R} -ev \mathbb{R}^2 , muni de la multiplication suivante :

si $z_1 = (x_1, y_1) \in \mathbb{C}$ et $z_2 = (x_2, y_2) \in \mathbb{C}$ on pose :

$$(1) \quad z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1).$$

Notons provisoirement (e, u) la base canonique du \mathbb{R} -ev \mathbb{R}^2 : $e = (1, 0)$, $u = (0, 1)$, et notons $\underline{0}$ l'élément nul $(0, 0)$ de \mathbb{C} . Il est clair que l'application $\mathbb{C} \rightarrow \mathbb{C}$, $(z_1, z_2) \mapsto z_1 z_2$ est \mathbb{R} -bilinéaire, que e est élément neutre pour cette multiplication qui est par ailleurs commutative. Pour vérifier l'associativité de cette multiplication, fixons z_1 et z_2 dans \mathbb{C} ; l'application $\mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto (z_1 z_2) z - z_1 (z_2 z)$ est \mathbb{R} -linéaire, prend la valeur $\underline{0}$ pour $z = e$ et $z = u$, donc elle est nulle pour tout z puisque (e, u) est une base du \mathbb{R} -ev \mathbb{C} .

En résumé, muni de la multiplication définie par (1), le \mathbb{R} -ev \mathbb{C} est une \mathbb{R} -algèbre commutative, d'élément unité e , et $e \neq \underline{0}$.

Si $z = (x, y) \in \mathbb{C}$, nous appellerons **conjugué** de z , et nous noterons \bar{z} , l'élément $(x, -y)$. On vérifie immédiatement que l'application $z \mapsto \bar{z}$ est \mathbb{R} -linéaire, involutive, et que $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$, autrement dit c'est un **automorphisme involutif de la \mathbb{R} -algèbre \mathbb{C}** .

Si $z = (x, y) \in \mathbb{C}$, on a $z\bar{z} = (x^2 + y^2, 0)$. On appelle **module** de z (noté $\text{mod } z$ ou $|z|$) le réel positif $(x^2 + y^2)^{1/2}$. Il est clair que : $|z| = 0$ ssi $z = \underline{0}$, et que $z\bar{z} = |z|^2 e$. De plus $|z_1 z_2| = |z_1| |z_2|$ pour tous $z_1, z_2 \in \mathbb{C}$. En effet : $(z_1 z_2) \times (\overline{z_1 z_2}) = (\bar{z}_1 \bar{z}_2) \times (z_1 z_2) = (z_1 \bar{z}_1) \times (z_2 \bar{z}_2)$, d'où le résultat.

THÉORÈME VI.5.1

|| La \mathbb{R} -algèbre \mathbb{C} est un corps commutatif.

Démonstration :

Il suffit de prouver que tout $z \in \mathbb{C} \setminus \{0\}$ est inversible dans la \mathbb{R} -algèbre \mathbb{C} . Or, si $z \in \mathbb{C} \setminus \{0\}$, $z\bar{z} = |z|^2 e$ avec $|z|^2 \neq 0$, d'où $\frac{1}{|z|^2} (z\bar{z}) = e$, c'est-à-dire $z \left(\frac{1}{|z|^2} \bar{z} \right) = e$, ce qui prouve bien que z est inversible, et son inverse est $\frac{1}{|z|^2} \bar{z}$. ■

L'application $\mathbb{R} \rightarrow \mathbb{C}$, $x \mapsto xe = (x, 0)$ définit un isomorphisme du corps \mathbb{R} sur un sous-corps de \mathbb{C} . Nous identifierons désormais, à l'aide de cet isomorphisme, \mathbb{R} et ce sous-corps. En particulier e sera identifié au nombre réel 1, 0 à 0, d'où $z\bar{z} = |z|^2$ et, si $z \neq 0$, $\frac{1}{z} = \frac{1}{|z|^2} \bar{z}$, produit qui peut s'interpréter comme le produit du scalaire $\frac{1}{|z|^2}$ par l'élément \bar{z} du \mathbb{R} -ev \mathbb{C} ou comme le produit de deux éléments de la \mathbb{R} -algèbre \mathbb{C} ; \mathbb{C} apparaît ainsi comme une **extension du corps** \mathbb{R} . On remarque que $u \in \mathbb{C} \setminus \mathbb{R}$ et que $u^2 = -1$. On pose désormais $u = i$ et la base canonique de \mathbb{C} considéré comme \mathbb{R} -ev devient $(1, i)$. Un élément $z = (x, y) \in \mathbb{C}$ s'écrit alors de manière unique $z = x + iy$, avec $x \in \mathbb{R}$ et $y \in \mathbb{R}$.

DÉFINITION VI.5.1

§ Le corps commutatif \mathbb{C} construit ci-dessus s'appelle **corps des nombres complexes**. Si $z = x + iy \in \mathbb{C}$ avec $x \in \mathbb{R}$ et $y \in \mathbb{R}$, x s'appelle la **partie réelle** de z et se note $\text{Re}(z)$, le réel y s'appelle la **partie imaginaire** de z et se note $\text{Im}(z)$.

Les applications $\text{Re} : \mathbb{C} \rightarrow \mathbb{R}$ et $\text{Im} : \mathbb{C} \rightarrow \mathbb{R}$ sont \mathbb{R} -linéaires. Ce sont en fait les *formes coordonnées* relatives à la base $(1, i)$ du \mathbb{R} -ev \mathbb{C} .

Les nombres complexes z tels que $\text{Re}(z) = 0$ sont les λi , $\lambda \in \mathbb{R}$. On les appelle nombres **imaginaires purs**. (Ceux tels que $\text{Im}(z) = 0$ sont les nombres réels.)

THÉORÈME VI.5.2

|| Il y a exactement deux automorphismes de la \mathbb{R} -algèbre \mathbb{C} :
|| l'application identique $\text{Id}_{\mathbb{C}}$, et la conjugaison $\Gamma : z \mapsto \bar{z}$.⁽¹⁾

Démonstration :

On sait déjà que $\text{Id}_{\mathbb{C}}$ et Γ sont des automorphismes de la

⁽¹⁾ Il ne faudrait surtout pas croire que \mathbb{C} , en tant que **corps**, ne possède que deux automorphismes. En fait, les automorphismes de corps de \mathbb{C} forment un ensemble **infini** qui n'est même pas dénombrable.

\mathbb{R} -algèbre \mathbb{C} . Réciproquement soit f un tel automorphisme. Alors nécessairement $f(x) = x$ pour tout $x \in \mathbb{R}$, d'où en particulier

$$f(-1) = -1 = f(i^2) = (f(i))^2 = i^2.$$

Donc $[f(i) - i][f(i) + i] = 0$ et, puisque \mathbb{C} est intègre, $f(i) = i$ ou $f(i) = -i$. Si $f(i) = i$, pour $z = x + iy \in \mathbb{C}$, on a :

$$f(z) = x + yf(i) = x + iy = z,$$

donc $f = \text{Id}_{\mathbb{C}}$. Si $f(i) = -i$, pour $z = x + iy \in \mathbb{C}$, on a :

$$f(z) = x + yf(i) = x - iy = \bar{z},$$

donc $f = \Gamma$. ■

PROPOSITION VI.5.1

|| Pour tous z, t dans $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, on a $|z + t| \leq |z| + |t|$, et l'égalité a lieu ssi il existe $\rho \in \mathbb{R}_+^*$ tel que $z = \rho t$.

Démonstration :

On a

$$|z + t|^2 = (z + t)(\bar{z} + \bar{t}) = |z|^2 + |t|^2 + z\bar{t} + \bar{z}t$$

et

$$(|z| + |t|)^2 = |z|^2 + |t|^2 + 2|z||t|.$$

Mais $|z\bar{t}| = |z||\bar{t}| = |z||t|$, d'où $|z\bar{t} + \bar{z}t| \leq 2|z||t|$, et par suite $|z + t| \leq |z| + |t|$. Si $z = \rho t$, avec $\rho \in \mathbb{R}_+^*$,

$$|z + t| = (1 + \rho)|t| = |z| + |t|.$$

Enfin supposons $|z + t| = |z| + |t|$; alors $z\bar{t} + \bar{z}t = 2|z||t|$, d'où en posant $z\bar{t} = u$, $u + \bar{u} = 2|u|$, c'est-à-dire $\text{Re}(u) = |u|$, d'où $u = |u|$, donc $u \in \mathbb{R}_+^*$ et $z\bar{t} = u$, d'où $z = \frac{1}{|t|^2} ut = \rho t$ avec $\rho = \frac{u}{|t|^2} \in \mathbb{R}_+^*$. ■

Cette proposition VI.5.1, que nous appellerons inégalité triangulaire, n'est évidemment qu'un cas particulier de l'*inégalité de Cauchy-Schwarz* des espaces euclidiens. Munissons \mathbb{C} , considéré comme \mathbb{R} -ev, de la **structure affine euclidienne orientée canonique** de \mathbb{R}^2 (c'est-à-dire celle pour laquelle $(0, (1, i))$ est un *repère orthonormé direct*). On obtient alors, sur le plan affine euclidien ainsi défini, une structure de corps qui est celle de \mathbb{C} . Ce plan affine euclidien, muni de cette structure de corps, s'appelle **plan d'Argand-Cauchy** ⁽¹⁾. Grâce à la structure de corps de \mathbb{C} , il est possible

⁽¹⁾ Jean Robert *Argand* (1768-1822) a écrit en 1806 un « Essai sur une manière de représenter les quantités imaginaires dans les constructions géométriques ».

Le baron Augustin-Louis *Cauchy* (1789-1857) a laissé une œuvre mathématique considérable (plus de 700 mémoires) et a exercé une influence profonde en son temps. Il est seulement dommage qu'il n'ait pas reconnu le génie d'Abel et de Galois.

d'étudier de nombreuses propriétés de géométrie euclidienne plane, comme nous le verrons au § VI.9.

Exercice 1 : Trouver la partie réelle et la partie imaginaire des nombres complexes

$$z_1 = \frac{(1+2i)^2 - (1-i)^3}{(3+2i)^3 - (2+i)^2} \quad \text{et} \quad z_2 = \frac{(2+i)^3 + (1-i)^2}{1+i + (2i-1)^2}.$$

Exercice 2 : Montrer que si deux des trois nombres $\frac{d-a}{b-c}$, $\frac{d-b}{c-a}$, $\frac{d-c}{a-b}$ sont imaginaires purs, le troisième l'est aussi. Donner une interprétation géométrique de ce résultat.

Exercice 3 : Soit u et v deux nombres complexes. Montrer qu'on a la relation :

$$|u+v|^2 + |u-v|^2 = 2(|u|^2 + |v|^2).$$

Interprétation géométrique.

Exercice 4 : Montrer que les nombres complexes distincts a et b ont même module ssi $\exists k \in \mathbb{R}$ tel que $a+b = ki(a-b)$. En déduire que $|a| = |b| \Leftrightarrow \exists \lambda \in \mathbb{R}_+$ tel que $(a+b)^2 + \lambda(a-b)^2 = 0$ ou $a=b$.

Exercice 5 : A quelle condition le module de la somme de deux nombres complexes est-il égal à la différence des deux modules ?

Exercice 6 : Montrer que si l'entier naturel A est somme de deux carrés ainsi que l'entier naturel B , il en va de même de leur produit AB et plus généralement de $A^m B^n$ ($m \in \mathbb{N}$, $n \in \mathbb{N}$).

Exercice 7 : Soit $z = p + iq$ donné. Peut-il être le produit de deux complexes dont l'un a une partie réelle donnée a et l'autre une partie imaginaire donnée b ? Discuter.

Exercice 8 : Entiers de Gauss. On considère l'ensemble E des nombres complexes de la forme $a + ib$ ($a \in \mathbb{Z}$, $b \in \mathbb{Z}$) et à chaque $z = a + ib \in E$ on associe l'entier naturel

$$N(z) = z\bar{z} = a^2 + b^2.$$

a) Montrer que $N(zt) = N(z)N(t)$ et que E est un anneau intègre dont on cherchera les éléments inversibles.

b) Montrer que si $N(z)$ est un entier premier, alors z est un élément *irréductible* de E (c'est-à-dire que les seules décompositions de z sont du type εu où ε est un élément inversible et u associé à z). Donner un exemple prouvant que la réciproque est fautive.

c) Soit $x \in E$ et $y \in E \setminus \{0\}$. Le quotient dans \mathbb{C} de x par y est de la forme $u + iv$, où $u \in \mathbb{Q}$, $v \in \mathbb{Q}$. Soit u_0 l'entier rationnel le plus voisin de u (ou l'un de ces entiers en cas d'ambiguïté), et de même v_0 un entier le plus voisin de v . Montrer que l'on peut définir dans E une *division euclidienne* $x = y(u_0 + iv_0) + r$ avec $N(r) < N(y)$. En déduire que tout idéal de E est principal et que la décomposition d'un élément de E en produit de facteurs irréductibles est unique (à l'ordre près des facteurs et à des facteurs inversibles près).

Exercice 9 : On pose

$$x = \frac{1+uv}{u+v}, \quad y = i \frac{1-uv}{u+v}, \quad z = \frac{u-v}{u+v}.$$

Comment faut-il choisir u et v dans \mathbb{C} pour que x , y et z soient réels ? Trouver une relation indépendante de u et v entre x , y et z .

Exercice 10 : Déterminer $z \in \mathbb{C}$ pour que z , $\frac{1}{z}$ et $1-z$ aient le même module.

§ VI.6 RACINES CARRÉES D'UN NOMBRE COMPLEXE

THÉORÈME VI.6.1

|| Soit $a = \alpha + i\beta \in \mathbb{C}^*$ ($\alpha \in \mathbb{R}, \beta \in \mathbb{R}$). Il existe exactement deux nombres complexes z tels que $z^2 = a$, et ils sont opposés.

Démonstration :

Supposons trouvé z_0 tel que $z_0^2 = a$. Alors $z^2 = a$ équivaut à $(z - z_0)(z + z_0) = 0$, donc à $z \in \{z_0, -z_0\}$, ce qui montre qu'il y a exactement deux nombres z vérifiant $z^2 = a$. Il reste donc à trouver un tel z_0 .

Si $\beta = 0$ et $\alpha > 0$, on prend $z_0 = \sqrt{\alpha}$. Si $\beta = 0$ et $\alpha < 0$, on prend $z_0 = i\sqrt{-\alpha}$. Si $\beta \neq 0$, posons $z = x + iy$ avec $x \in \mathbb{R}$ et $y \in \mathbb{R}$. De $(x + iy)^2 = \alpha + i\beta$ on déduit, entre autres, $x^2 - y^2 = \alpha$, $2xy = \beta$, et aussi $x^2 + y^2 = |a| = (\alpha^2 + \beta^2)^{1/2}$ d'où

$$x^2 = \frac{1}{2} (\alpha + (\alpha^2 + \beta^2)^{1/2}) \quad \text{et} \quad y^2 = \frac{1}{2} (-\alpha + (\alpha^2 + \beta^2)^{1/2}).$$

Tenant compte du signe de xy qui doit être le même que le signe de β , il suffit de poser $\varepsilon = \frac{\beta}{|\beta|}$ pour constater que le nombre $z_0 = x_0 + iy_0$ vérifie $z_0^2 = a$ si l'on prend

$$x_0 = \left[\frac{1}{2} (\alpha + (\alpha^2 + \beta^2)^{1/2}) \right]^{1/2} \quad \text{et} \quad y_0 = \varepsilon \left[\frac{1}{2} (-\alpha + (\alpha^2 + \beta^2)^{1/2}) \right]^{1/2}. \quad \blacksquare$$

Exemple 1 : Soit à trouver les racines carrées de $2 + i$. Ce sont z_0 et $-z_0$, où z_0 , en utilisant la méthode ci-dessus, vaut

$$\left(\frac{2 + \sqrt{5}}{2} \right)^{1/2} + \left(\frac{\sqrt{5} - 2}{2} \right)^{1/2} i.$$

On obtient facilement, comme application du théorème VI.6.1, la *résolution par radicaux des équations du second degré dans \mathbb{C}* . En effet, soit a et b deux nombres complexes. La recherche des $z \in \mathbb{C}$ solutions de l'équation $z^2 + 2az + b = 0$ équivaut à celle des $Z = z + a$ tels que $Z^2 = a^2 - b$. On voit donc que l'équation $z^2 + 2az + b = 0$, $z \in \mathbb{C}$, admet exactement deux racines complexes si $a^2 - b \neq 0$, une et une seule (égale à $-a$) si $a^2 - b = 0$.

On remarquera que, dans le cas où a et b sont réels, les deux racines dans \mathbb{C} de l'équation $z^2 + 2az + b = 0$ sont réelles si $a^2 - b = \Delta' > 0$, non réelles et complexes conjuguées si $a^2 - b = \Delta' < 0$.

Si la résolution des équations du second degré à coefficients dans \mathbb{C} ne présente aucune difficulté, le corps des complexes ayant pra

construit sur mesure pour cela ⁽¹⁾, il est beaucoup plus remarquable que tout polynôme non constant à coefficients dans \mathbb{C} admette toujours au moins une racine dans \mathbb{C} , ce que l'on traduit en disant que **le corps \mathbb{C} est algébriquement clos**, résultat connu sous le nom de théorème de D'Alembert ⁽²⁾, ou théorème fondamental de l'Algèbre.

Exemple 2 : La méthode ci-dessus, appliquée à l'équation $z^2 + z + 1 = 0$ donne pour racines

$$j = \frac{-1}{2} + i \frac{\sqrt{3}}{2} \quad \text{et} \quad \bar{j} = \frac{-1}{2} - i \frac{\sqrt{3}}{2}.$$

En tenant compte que $j^2 = -j - 1$ on s'aperçoit que la seconde racine n'est autre que j^2 . Comme $z^3 - 1 = (z - 1)(z^2 + z + 1)$ on en déduit que les racines complexes de l'équation $z^3 - 1 = 0$ sont 1, j et j^2 .

Exercice 1 : Soit l'équation $z^2 - (2 + i\omega)z + i\omega + 2 - \omega = 0$ où $\omega \in \mathbb{C}$.

a) Pour quelle valeur de ω cette équation admet-elle deux racines complexes conjuguées. Les calculer.

b) Résoudre l'équation pour $\omega = 1 + i$, puis pour ω quelconque.

Exercice 2 : Calculer les racines carrées de $5 + 12i$; de $4i - 3$; de $8i - 15$; de $9 + 40i$.

Exercice 3 : Résoudre les équations

$$iz^2 + (1 - 5i)z + 6i - 2 = 0; \quad 2z^2 - (20 + 9i)z + 50 = 0.$$

Exercice 4 : Résoudre les équations $z^4 = 24i - 7$; $z^4 = 2 - i\sqrt{12}$; $z^4 - 3iz^2 + 4 = 0$.

Exercice 5 : Résoudre les équations $z^2 = -(\bar{z})^2$; $8z^2 = \bar{z}$; $8z^2 = \bar{z} - 1$.

Exercice 6 : Condition sur le paramètre a pour que l'équation $z + \bar{z}^2 = a$ admette une racine réelle. Si cette condition est réalisée, résoudre l'équation.

Exercice 7 : L'équation $z^2 - 2(\alpha + i\beta)z - 2 - 2i = 0$ a deux racines d'images M_1 et M_2 dans le plan d'Argand-Cauchy. Trouver l'ensemble des points M_1 et M_2 sachant que la droite M_1M_2 a pour pente 1.

Exercice 8 : Soit $z \in \mathbb{C}$ d'image M et soit P et Q les images de ses deux racines carrées. Quel est l'ensemble des points M pour lesquels $MP \perp MQ$? Quels sont les lieux géométriques correspondants pour P et Q .

Exercice 9 : Comment faut-il choisir z pour que les images des nombres z , z^2 et z^4 soient alignées?

⁽¹⁾ La vérité historique oblige à reconnaître que ce n'est pas pour résoudre les équations du second degré à coefficients réels que les nombres complexes ont été inventés, mais à l'occasion de la résolution des équations du troisième degré sous le nom de « quantités imaginaires », et cela au milieu du XVI^e siècle de notre ère.

⁽²⁾ Jean Le Rond D'Alembert (1717-1783), mathématicien et philosophe français, qui a collaboré avec Denis Diderot (1713-1784) pour sa célèbre Encyclopédie.

§ VI.7 NOMBRES COMPLEXES DE MODULE 1

Le groupe \mathbb{U}

L'application $z \rightarrow |z|$ envoie \mathbb{C}^* dans \mathbb{R}_+^* , et vérifie $|zt| = |z||t|$ pour tous $z \in \mathbb{C}$, $t \in \mathbb{C}$. Elle définit donc un *homomorphisme du groupe multiplicatif \mathbb{C}^* dans le groupe multiplicatif \mathbb{R}_+^** . Le noyau de cet homomorphisme est donc un *sous-groupe du groupe multiplicatif \mathbb{C}^** , noté \mathbb{U} . Par définition \mathbb{U} est donc l'ensemble des $z \in \mathbb{C}$ tels que $|z| = 1$, muni de la multiplication des nombres complexes. On l'appelle **groupe des nombres complexes de module 1**. Dans le plan d'Argand-Cauchy, \mathbb{U} n'est autre que le **cercle-unité de \mathbb{C}** , c'est-à-dire de centre 0 et de rayon 1, d'équation $x^2 + y^2 = 1$. Les nombres 1, -1 , i et $-i$ sont quatre points remarquables de ce cercle.

THÉORÈME VI.7.1

|| L'application $f : \mathbb{R}_+^* \times \mathbb{U} \rightarrow \mathbb{C}^*$, $(r, u) \mapsto ru$ définit un **isomorphisme du groupe produit $\mathbb{R}_+^* \times \mathbb{U}$ sur le groupe multiplicatif \mathbb{C}^*** .

Démonstration :

Soit (r_1, u_1) et (r_2, u_2) deux éléments quelconques de l'ensemble de départ. Alors

$$\begin{aligned} f((r_1, u_1)(r_2, u_2)) &= f(r_1 r_2, u_1 u_2) = r_1 r_2 u_1 u_2 = \\ &= (r_1 u_1)(r_2 u_2) = f(r_1, u_1) \times f(r_2, u_2), \end{aligned}$$

donc f est bien un homomorphisme de groupes.

Le noyau de f est l'ensemble des $(r, u) \in \mathbb{R}_+^* \times \mathbb{U}$ tels que $ru = 1$. Mais si $ru = 1$, alors $|ru| = |r||u| = 1 = |r| = r$, donc $r = 1$ et en conséquence $u = 1$. Donc f est injectif. Enfin, soit $z \in \mathbb{C}^*$, on a : $z = f\left(\left(|z|, \frac{z}{|z|}\right)\right)$, donc f est surjectif. ■

Homomorphisme exponentiel

Nous introduirons dans le Cours d'Analyse (tome 2) l'application exponentielle $\exp : \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto \exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$. Nous verrons que l'application

$\mathcal{E}_1 : \mathbb{R} \rightarrow \mathbb{C}$, $x \mapsto \exp(ix)$, prend ses valeurs dans \mathbb{U} , et que si l'on note \mathcal{E} la corestriction de \mathcal{E}_1 à \mathbb{U} , on a le résultat :

THÉORÈME VI.7.2

|| L'application \mathcal{E} est un **homomorphisme surjectif** du groupe additif \mathbb{R} sur le groupe multiplicatif \mathbb{U} . Le noyau de cet homomorphisme est l'ensemble $2\pi\mathbb{Z} = \{2\pi k\}_{k \in \mathbb{Z}}$, où le nombre π est défini comme étant le double du **plus petit réel $\alpha > 0$ tel que Re**

Les fonctions $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \operatorname{Re}(\mathcal{E}(x))$ et $x \mapsto \operatorname{Im}(\mathcal{E}(x))$ se nomment respectivement **fonction cosinus** et **fonction sinus** et se notent \cos et \sin . Ces fonctions sont donc définies comme sommes de séries entières, et c'est à partir de là qu'on démontre l'existence du nombre $\alpha = \frac{\pi}{2}$ tel que $\cos \alpha = 0$. Bien entendu, le nombre π ainsi que les fonctions \cos et \sin sont identiques au nombre π bien connu du lecteur et aux fonctions déjà étudiées dans les classes précédentes et dont les propriétés élémentaires (sens de variation, périodicité, parité, formules d'addition et toutes les formules usuelles de trigonométrie qui s'en déduisent) sont également supposées connues, mais seront retrouvées dans le tome 2 à partir de leur définition ci-dessus.

Par définition, on a donc, pour tout $x \in \mathbb{R}$:

$$(1) \quad \exp(ix) = \cos x + i \sin x ,$$

ce que l'on note aussi e^{ix} . Le théorème VI.7.2 entraîne les conséquences fondamentales suivantes :

$$(2) \quad (\forall n \in \mathbb{Z}) \quad \exp(inx) = \cos nx + i \sin nx = (e^{ix})^n .$$

Ce résultat est connu sous le nom de **formule de De Moivre** ⁽¹⁾.

(3) Pour tout réel a , l'application $[a, a + 2\pi[\rightarrow \mathbb{U}$, $x \mapsto \exp(ix)$ est bijective.

En remarquant que $\exp(-ix) = \cos x - i \sin x$, on obtient les **formules d'Euler** ⁽²⁾.

$$(4) \quad \boxed{\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}}$$

signalons enfin quelques égalités remarquables : $e^{i\pi} = -1$, $e^{i\frac{\pi}{2}} = i$, $e^{\frac{2i\pi}{3}} = j$, $e^{-\frac{2i\pi}{3}} = j^2$ et notons que $(x \in \mathbb{R} \text{ et } e^{ix} = 1) \Leftrightarrow (x \in 2\pi\mathbb{Z})$.

Si nous appliquons le théorème V.7.3, nous voyons, d'après le théorème VI.7.2, que par passage au quotient, l'homomorphisme $\mathcal{E} : \mathbb{R} \rightarrow \mathbb{U}$ définit un *isomorphisme* du groupe quotient $\mathbb{R}/2\pi\mathbb{Z}$ sur le groupe \mathbb{U} . Cet isomorphisme associe, à toute classe \mathcal{C} de $\mathbb{R} \bmod (2\pi\mathbb{Z})$, la valeur constante prise par \mathcal{E} sur les éléments de \mathcal{C} .

⁽¹⁾ Abraham *de Moivre* (1667-1754) est un mathématicien anglais d'origine française.

⁽²⁾ Leonhard *Euler* (1707-1783), déjà signalé comme l'un des mathématiciens les plus prolifiques de tous les temps, est à l'origine de l'adoption de la lettre e pour le nombre $\sum_{n=0}^{+\infty} \frac{1}{n!}$ et de la lettre π pour le nombre d'Archimède.

Applications des formules de De Moivre et d'Euler

Exemple 1 : Somme des termes d'une progression géométrique.

Soit à évaluer, pour $n \in \mathbb{N}$ et $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, la somme $S_n = \sum_{k=0}^n e^{ik\theta}$.

Puisque $e^{i\theta} \neq 1$, on trouve

$$S_n = \sum_{k=0}^n (e^{i\theta})^k = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} = \frac{e^{i\frac{n+1}{2}\theta}}{e^{i\frac{\theta}{2}}} \times \frac{e^{i\frac{n+1}{2}\theta} - e^{-i\frac{n+1}{2}\theta}}{e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}}},$$

et grâce aux formules d'Euler

$$S_n = e^{in\frac{\theta}{2}} \times \frac{\sin \frac{n+1}{2} \theta}{\sin \frac{\theta}{2}}.$$

En prenant les parties réelle et imaginaire, on obtient :

$$\sum_{k=0}^n \cos k\theta = \cos \frac{n\theta}{2} \times \frac{\sin \frac{n+1}{2} \theta}{\sin \frac{\theta}{2}}, \quad \sum_{k=0}^n \sin k\theta = \sin \frac{n\theta}{2} \times \frac{\sin \frac{n+1}{2} \theta}{\sin \frac{\theta}{2}}.$$

La première de ces égalités peut se lire :

$$(5) \quad \frac{1}{2} + \sum_{k=1}^n \cos k\theta = \frac{\sin \frac{2n+1}{2} \theta}{2 \sin \frac{\theta}{2}},$$

façon élégante de linéariser le quotient du second membre et que l'on peut retrouver en utilisant les formules d'Euler et en effectuant la division.

Exemple 2 : linéarisation de $\cos^n x$.

Soit à exprimer, pour $n \in \mathbb{N}$ et $x \in \mathbb{R}$, $\cos^n x$ comme combinaison linéaire de sinus ou cosinus d'arcs multiples de x (par exemple, en vue d'un calcul intégral).

En appliquant les formules d'Euler et du binôme, il vient : $\cos^n x =$

$$= \frac{1}{2^n} (e^{ix} + e^{-ix})^n = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{ikx} \times e^{-i(n-k)x} = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{i(2k-n)x}.$$

En regroupant les termes équidistants des extrêmes, on obtient :
pour n pair ($n = 2p$),

$$\cos^{2p} x = \frac{1}{2^{2p}} \left[\binom{2p}{p} + 2 \sum_{k=0}^{p-1} \binom{2p}{k} \cos(2p - 2k)x \right]$$

pour n impair ($n = 2p + 1$),

$$\cos^{2p+1} x = \frac{1}{2^{2p}} \sum_{k=0}^p \binom{2p+1}{k} \cos(2p+1-2k)x.$$

On procédera de même pour $\sin^n x$ (somme de sinus si n impair, de cosinus si n est pair) et pour $\cos^m x \times \sin^n x$.

Exemple 3 : Expression de $\cos nx$ comme polynôme en $\cos x$.

$$\begin{aligned} \text{On a } \cos nx &= \operatorname{Re}(e^{inx}) = \operatorname{Re}(\cos x + i \sin x)^n \\ &= \sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k} \cos^{n-2k} x (i \sin x)^{2k} \\ &= \sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k} \cos^{n-2k} x (-1)^k (1 - \cos^2 x)^k \end{aligned}$$

qui s'écrit bien sous la forme d'un polynôme T_n en $\cos x$ dont le coefficient du terme de plus haut degré (terme en $\cos^n x$) est :

$$\sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k} = 2^{n-1}.$$

Nous donnerons au chapitre suivant le développement complet de T_n .

Pour n impair, $\sin nx$ est également un polynôme en $\sin x$. Pour $n \in \mathbb{N}^*$, $\frac{\sin nx}{\sin x}$ peut s'exprimer comme polynôme en $\cos x$.

Exemple 4 : Expression de $\operatorname{tg} nx$ en fonction rationnelle de $\operatorname{tg} x$.

On sait que, pour tout réel $x \in \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \right\}_{k \in \mathbb{Z}}$, on définit $\operatorname{tg} x$ comme étant égal à $\frac{\sin x}{\cos x}$. En séparant partie réelle et partie imaginaire dans $e^{inx} = (\cos x + i \sin x)^n$ on obtient

$$\cos nx = \sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k} \cos^{n-2k} x (-\sin^2 x)^k$$

$$\text{et } \sin nx = \sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k+1} \cos^{n-2k-1} x (-1)^k \sin^{2k+1} x.$$

d'où, après division du numérateur et du dénominateur par $\cos^n x$:

$$\operatorname{tg} nx = \frac{\sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k+1} (-1)^k \operatorname{tg}^{2k+1} x}{\sum_{k=0}^{E\left(\frac{n}{2}\right)} \binom{n}{2k} (-1)^k \operatorname{tg}^{2k} x},$$

soit

$$\operatorname{tg} nx = \frac{\binom{n}{1} \operatorname{tg} x - \binom{n}{3} \operatorname{tg}^3 x + \dots}{1 - \binom{n}{2} \operatorname{tg}^2 x + \dots}.$$

Exemple 5 : Calcul de certaines fonctions symétriques.

L'exemple précédent nous conduit naturellement à considérer le polynôme

$$f(X) = \sum_{2k+1 \leq n} (-1)^k \binom{n}{2k+1} X^{2k+1}.$$

Si n est impair ($n = 2p + 1$), ce polynôme est de degré n , et, en anticipant sur la théorie des polynômes d'une variable traitée au chapitre VII, il est clair qu'en posant $\theta_k = \frac{k\pi}{2p+1}$ ($-p \leq k \leq p$), les nombres $\operatorname{tg} \theta_k$ sont des zéros du polynôme $f(X)$ puisque $\operatorname{tg} n\theta_k = \operatorname{tg} k\pi = 0$. Ils sont en nombre $2p + 1$, tous distincts, ce qui permet de factoriser $f(X)$, compte tenu du coefficient de X^n , en

$$f(X) = (-1)^p \prod_{k=-p}^p (X - \operatorname{tg} \theta_k).$$

Compte tenu de : $\operatorname{tg}(\theta_{-k}) = -\operatorname{tg} \theta_k$, on peut écrire plus simplement :

$$f(X) = (-1)^p X \prod_{k=1}^p (X^2 - \operatorname{tg}^2 \theta_k),$$

ou, en posant :

$$g(T) = \sum_{2k+1 \leq n} (-1)^k \binom{2p+1}{2k+1} T^k,$$

$$(6) \quad g(T) = (-1)^p \prod_{k=1}^p (T - \operatorname{tg}^2 \theta_k)$$

relation qui est équivalente, en introduisant le nouveau polynôme

$$h(U) = \sum_{2k+1 \leq n} (-1)^k \binom{2p+1}{2k+1} U^{p-k},$$

à :

$$(7) \quad h(U) = (2p + 1) \prod_{k=1}^p (U - \cotg^2 \theta_k).$$

En prenant la somme des racines, on déduit de (6) : $\sum_{k=1}^p \tg^2 \theta_k = p(2p + 1)$,
et de (7) :

$$\sum_{k=1}^p \cotg^2 \theta_k = \frac{1}{3} p(2p - 1).$$

A partir de cette dernière relation il est facile, par des encadrements convenables, de démontrer, après Euler, que $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Exercice 1 : Vérifier (cf. exemples 1 et 3) que, pour $p \in \mathbb{N}^*$,

$$\frac{\sin 2p\theta}{\sin \theta} = \sum_{k=1}^p 2 \cos (2k - 1) \theta \quad \text{et} \quad \frac{\sin (2p + 1) \theta}{\sin \theta} = 1 + \sum_{k=1}^p 2 \cos 2k\theta.$$

Exercice 2 : Soit des réels a_1, a_2, \dots, a_n tels que tous les $\tg a_k$ ($1 \leq k \leq n$) et $\tg (a_1 + \dots + a_n)$ soient définis. On pose $u_k = \tg a_k$. Pour $p \in \llbracket 1, n \rrbracket$, soit

$$\sigma_p = \sum_{j_1 < j_2 < \dots < j_p} u_{j_1} u_{j_2} \dots u_{j_p}.$$

$$\text{Démontrer :} \quad \tg (a_1 + a_2 + \dots + a_n) = \frac{\sigma_1 - \sigma_3 + \dots + (-1)^q \sigma_{2q+1} + \dots}{1 - \sigma_2 + \dots + (-1)^r \sigma_{2r} + \dots}.$$

Exercice 3 : Soit $x \in \mathbb{R}$ et $n \in \mathbb{N}^*$ tels que $\cos (2^k x) \neq 0$ pour $k \in \llbracket 0, n - 1 \rrbracket$. Exprimer sous forme simple

$$S_n = \sum_{k=1}^n \frac{1}{2^k \cos x \cos 2x \dots \cos (2^{k-1} x)}.$$

Exercice 4 : Soit $p \in \mathbb{N}^*$. Montrer : $(\forall x \in \mathbb{R}), (\forall n \geq 2p + 1)$

$$\sum_{q=0}^{2n-1} \cos^{2p} \left(x + \frac{q\pi}{2n} \right) = 2n \frac{\binom{2p}{p}}{2^{2p}}.$$

Exercice 5 : Exprimer sous une forme simple

$$U_n(x) = \sum_{k=0}^n \binom{n}{k} \cos kx$$

et

$$V_n(x) = \sum_{k=1}^n \binom{n}{k} \sin kx \quad (x \in \mathbb{R}, n \in \mathbb{N}^*).$$

Exercice 6 : Donner une expression simple des sommes suivantes :

$$S_n = \sum_{k=0}^n \cos^3 kx \quad \text{et} \quad T_n = 1 + \sum_{k=1}^{n-1} \frac{\cos kx}{\cos^k x}.$$

Exercice 7 : Pour $x \in \mathbb{R}$, $n \in \mathbb{N}^*$, et $p \in \mathbb{N}^*$ trouver une méthode permettant de calculer simplement

$$S_{n,p}(x) = \sum_{k=0}^n k^p \cos kx \quad \text{et} \quad T_{n,p}(x) = \sum_{k=0}^n k^p \sin kx.$$

Application : $p = 2, p = 3$.

Exercice 8 : Exprimer simplement $S_n = \sum_{k=1}^n \frac{1}{\cos kx \cos (k+1)x}$.

Indication : penser à $\operatorname{tg} p - \operatorname{tg} q$.

Exercice 9 : Prouver que $1 + \frac{1}{\cos 2x} = \frac{\operatorname{tg} 2x}{\operatorname{tg} x}$ et en déduire

$$P = \prod_{k=0}^n \left(1 + \frac{1}{\cos (2^k x)} \right).$$

Exercice 10 : Pour $x \in \mathbb{R}$ et $\theta \in \mathbb{R}$, exprimer simplement

$$S_n(x) = \sum_{k=0}^n x^k \cos k\theta \quad \text{et} \quad T_n(x) = \sum_{k=1}^n x^k \sin k\theta.$$

Exercice 11 : Exprimer simplement les sommes :

$$S_n(\alpha) = \sum_{k=0}^{n-1} 3^k \sin^3 \frac{\alpha}{3^{k+1}} \quad \text{et} \quad T_n(\alpha) = \sum_{k=0}^n \frac{1}{2^k} \operatorname{tg} \frac{\alpha}{2^k}.$$

Exercice 12 : On démontre en Analyse que si $x \in \left] 0, \frac{\pi}{2} \right[$ on a : $\sin x < x < \operatorname{tg} x$, d'où

$$\cotg^2 x < \frac{1}{x^2} < 1 + \cotg^2 x.$$

En appliquant ces inégalités à $x = \theta_k = \frac{k\pi}{2p+1}$ (cf. exemple 5) et en ajoutant, en déduire un encadrement de $S_p = 1 + \frac{1}{2^2} + \dots + \frac{1}{p^2}$. Qu'obtient-on si $p \rightarrow +\infty$?

Exercice 13 : Exprimer $\operatorname{tg} 6\theta$ en fonction de $\operatorname{tg} \theta$. En déduire une équation algébrique que vérifie le nombre $x = \operatorname{tg} \frac{\pi}{24}$.

§ VI.8 ARGUMENTS D'UN NOMBRE COMPLEXE ; RACINES n -ièmes

On a vu que l'application $\mathbb{R} \rightarrow \mathbb{U}$, $\theta \mapsto e^{i\theta}$ est surjective, et que, si $\theta \in \mathbb{R}$, on a : $e^{i\theta} = 1$ ssi $\theta \in 2\pi\mathbb{Z}$. Cela conduit à poser :

DÉFINITION VI.8.1

Etant donné un nombre complexe $z \neq 0$, on appelle **argument de z** tout réel θ tel que $e^{i\theta} = \frac{z}{|z|}$. L'ensemble des arguments de z sera noté $\arg(z)$.

Pour z fixé dans \mathbb{C}^* , $\arg(z)$ est toujours non vide. Soit $\theta_0 \in \arg(z)$. Si $\theta \in \mathbb{R}$, pour que l'on ait : $\theta \in \arg(z)$, il faut et il suffit que $e^{i\theta} = e^{i\theta_0}$, c'est-à-dire : $e^{i(\theta - \theta_0)} = 1$. Par suite :

THÉORÈME VI.8.1

Si $z \in \mathbb{C}^*$, l'ensemble $\arg(z)$ est non vide, et pour tout θ_0 de $\arg(z)$, on a :

$$\arg(z) = \{\theta_0 + 2k\pi\}_{k \in \mathbb{Z}}$$

(ce que l'on note aussi : $\arg(z) = \theta_0 + 2\pi\mathbb{Z}$).

Soit alors $z \in \mathbb{C}^*$. On appelle **forme trigonométrique de z** tout couple $(r, \theta) \in \mathbb{R}_+^* \times \mathbb{R}$ tel que $z = r e^{i\theta}$. Les formes trigonométriques de z sont évidemment les couples $(|z|, \theta)_{\theta \in \arg(z)}$. On reconnaît donc l'égalité de deux nombres complexes non nuls z_1 et z_2 au fait que $|z_1| = |z_2|$ et $\theta_1 - \theta_2 \in 2\pi\mathbb{Z}$. Cette forme trigonométrique est particulièrement bien adaptée pour calculer le produit de nombres complexes, le quotient de deux complexes et, bien entendu, la puissance (*entière*) d'un nombre complexe. On savait déjà que le module d'un produit est le produit des modules. On peut maintenant ajouter les règles :

- (1) Si $z_i \in \mathbb{C}^* (i \in [1, n])$ et si $\theta_i \in \arg(z_i)$, alors $\sum_{i=1}^n \theta_i \in \arg\left(\prod_{i=1}^n z_i\right)$.
- (2) Si $z \in \mathbb{C}^*$ et si $\theta \in \arg(z)$, alors $-\theta \in \arg\left(\frac{1}{z}\right)$, d'où $\frac{1}{z} = \frac{1}{|z|} e^{-i\theta}$.

La propriété (3) du § VI.7 entraîne la suivante :

- (3) Soit $a \in \mathbb{R}$. Pour tout $z \in \mathbb{C}^*$, il existe un et un seul réel θ dans $[a, a + 2\pi[\cap \arg(z)$.

De même qu'on peut définir sur \mathbb{C} un **ordre total** en utilisant la forme algébrique $z = x + iy$ et en posant : $z_1 \leq z_2$ ssi $x_1 < x_2$ ou $(x_1 = x_2$ et $y_1 \leq y_2)$, on peut définir sur \mathbb{C}^* un autre ordre total en désignant par $\arg_0(z)$ l'unique réel de $\arg(z) \cap [0, 2\pi[$ et en posant : $z_1 \leq z_2$ ssi $|z_1| < |z_2|$ ou $(|z_1| = |z_2|$ et $\arg(z_1) \leq \arg(z_2)$). Chacun de ces ordres est un ordre *lexicographique*, mais, s'ils peuvent rendre de grands services, aucun d'eux n'est véritablement satisfaisant en ce sens que le premier n'est pas compatible avec la structure de corps de \mathbb{C} , le second avec la structure de groupe de \mathbb{C}^* .

Un calcul d'arguments doit toujours être mené avec soin :

Exemple 1 : Soit $\theta \in \mathbb{R}$. On se propose de mettre sous forme trigonométrique le nombre $z = 1 + \cos \theta + i \sin \theta$. Il est immédiat que

$$z = 2 \cos \frac{\theta}{2} e^{i\frac{\theta}{2}}.$$

On doit donc déjà supposer $\theta \notin \pi + 2\pi\mathbb{Z}$, sinon $z = 0$. Ce

étant remplie il est clair que $|z| = 2 \left| \cos \frac{\theta}{2} \right|$. Il ne reste plus qu'à trouver un argument de z . Si $\cos \frac{\theta}{2} > 0$, alors $\frac{\theta}{2} \in \arg(z)$ et si $\cos \frac{\theta}{2} < 0$, alors $\frac{\theta}{2} + \pi \in \arg(z)$. Or, dire que $\cos \frac{\theta}{2} > 0$ revient à dire qu'il existe un entier k tel que

$$-\frac{\pi}{2} + 2k\pi < \frac{\theta}{2} < \frac{\pi}{2} + 2k\pi,$$

c'est-à-dire tel que $-\frac{1}{4} + k < \frac{\theta}{4\pi} < \frac{1}{4} + k$, et $\cos \frac{\theta}{2} < 0$ s'il existe $k \in \mathbb{Z}$ tel que $\frac{1}{4} + k < \frac{\theta}{4\pi} < \frac{3}{4} + k$. Le second cas se présentera donc si la partie décimale de $\frac{\theta}{4\pi}$ (c'est-à-dire $\frac{\theta}{4\pi} - E\left(\frac{\theta}{4\pi}\right)$) est dans l'intervalle $\left] \frac{1}{4}, \frac{3}{4} \right[$.

DÉFINITION VI.8.2

$\left\{ \begin{array}{l} \text{On appelle } \textbf{argument principal} \text{ d'un nombre } z \in \mathbb{C} \setminus \mathbb{R}_-, \text{ et on note} \\ \text{Arg}(z), \text{ l'unique réel élément de } \arg(z) \cap]-\pi, +\pi[. \end{array} \right.$

En particulier si $z \in \mathbb{R}_+^*$, $\text{Arg}(z)$ qu'on écrit plus brièvement $\text{Arg } z$, vaut 0.

Racines n -ièmes d'un nombre complexe

THÉORÈME VI.8.2

$\left\| \begin{array}{l} \text{Soit } n \text{ un entier } \geq 2 \text{ et } A \text{ un nombre complexe non nul. L'ensemble} \\ \text{noté } \mathcal{R}_n(A) \text{ des racines } n\text{-ièmes de } A, \text{ c'est-à-dire l'ensemble} \\ \{z \in \mathbb{C} \mid z^n = A\} \text{ est de } \textbf{cardinal } n. \text{ Pour tout } \theta \in \arg(A), \text{ on a :} \\ \\ \mathcal{R}_n(A) = \left\{ |A|^{1/n} \exp \left(i \left(\frac{\theta}{n} + \frac{2k\pi}{n} \right) \right) \right\}_{k \in \llbracket 0, n-1 \rrbracket}. \end{array} \right.$

Démonstration :

Un $z \in \mathbb{C}$ tel que $z^n = A$ est nécessairement non nul. Cherchons-le donc sous sa forme trigonométrique

$$z = r e^{i\varphi} \quad (r \in \mathbb{R}_+^*, \varphi \in \mathbb{R}).$$

L'équation binôme $z^n = A$ équivaut à $r^n e^{in\varphi} = A$, c'est-à-dire, en choisissant un $\theta \in \arg(A)$, à $r^n e^{in\varphi} = |A| e^{i\theta}$, soit encore $r^n = |A|$ et $n\varphi - \theta \in 2\pi\mathbb{Z}$.

Pour $k \in \llbracket 0, n-1 \rrbracket$, si nous posons

$$z_k = |A|^{1/n} \exp \left(\frac{i\theta}{n} + \frac{2ik\pi}{n} \right),$$

il est clair que $z_k \in \mathcal{R}_n(A)$. Mais si k et l sont distincts dans $\llbracket 0, n-1 \rrbracket$, il est non moins clair que

$$\frac{\theta}{n} + \frac{2k\pi}{n} - \left(\frac{\theta}{n} + \frac{2l\pi}{n} \right) = \frac{2(k-l)\pi}{n} \notin 2\pi\mathbb{Z},$$

ce qui prouve que z_k et z_l sont distincts, ce qui donne déjà n racines distinctes, autant que de z_k distincts. Il reste à voir qu'on les a toutes. Soit $z = |A|^{1/n} e^{i\varphi} \in \mathcal{R}_n(A)$, où nécessairement $\varphi = \frac{\theta}{n} + \frac{2m\pi}{n}$, avec $m \in \mathbb{Z}$.

La division euclidienne de m par n donne $m = qn + r$ avec $r \in \llbracket 0, n-1 \rrbracket$, d'où $\frac{\theta}{n} + \frac{2r\pi}{n} \in \arg(z)$, et par suite $z = z_r$. En conclusion

$$\mathcal{R}_n(A) = \{z_k\}_{k \in \llbracket 0, n-1 \rrbracket} \quad \text{et} \quad \text{card}(\mathcal{R}_n(A)) = n. \quad \blacksquare$$

Lorsque $n = 2$, on retrouve donc par une autre voie le théorème VI.6.1.

Racines de l'unité

Donnons-nous un entier $n \geq 2$. L'application $f_n: \mathbb{U} \rightarrow \mathbb{U}, z \mapsto z^n$ est un homomorphisme de groupes, surjectif d'après le théorème VI.8.2. Son noyau est noté μ_n : c'est un sous-groupe de \mathbb{U} , et nous venons de voir qu'il est de cardinal n . Le groupe μ_n s'appelle **groupe des racines n -ièmes de l'unité** dans \mathbb{C} . Par définition $\mu_n = \{z \in \mathbb{C} \mid z^n = 1\}$.

THÉORÈME VI.8.3

$$\left\| \begin{array}{l} \text{Le groupe } \mu_n \text{ est } \textbf{cyclique}. \text{ Ses générateurs sont les nombres} \\ \zeta_k = \exp\left(\frac{2ik\pi}{n}\right), \quad k \in \llbracket 0, n-1 \rrbracket, \quad k \vee n = 1. \end{array} \right.$$

Démonstration :

$0 \in \arg(1)$. La démonstration du théorème précédent montre donc que $\mu_n = \{\zeta_k\}_{0 \leq k \leq n-1}$. Mais, pour tout $k \in \mathbb{Z}$, on a $\zeta_k = (\zeta_1)^k$, d'où il résulte que ζ_1 engendre le groupe μ_n , qui est bien cyclique. Puisque ζ_1 est un générateur, et que $\text{card}(\mu_n) = n$, c'est que ζ_1 est d'ordre n . Pour trouver tous les éléments d'ordre n de μ_n , il suffit de se reporter au théorème V.2.4. \blacksquare

DÉFINITION VI.8.3

*On appelle **racine n -ième primitive** de l'unité dans \mathbb{C} tout générateur du groupe μ_n , c'est-à-dire chacun des nombres*

$$\zeta_k = \exp\left(\frac{2ik\pi}{n}\right), \quad k \in \llbracket 0, n-1 \rrbracket, \quad k \vee n = 1$$

Le nombre de racines n -ièmes *primitives* de 1 est $\varphi(n)$, où φ est l'indicateur d'Euler défini au § IV.3. Parmi ces racines, ζ_1 joue un rôle particulier du fait que $\arg_0(\zeta_1)$ est le plus petit parmi les $\arg_0(\zeta_k)_{k \in \llbracket 0, n-1 \rrbracket}$. On dit parfois que ζ_1 est le **générateur principal** de μ_n .

Exemple 2 : Notons \mathcal{P}_n l'ensemble des racines primitives n -ièmes de 1. Le lecteur vérifiera facilement que

$$\mathcal{P}_2 = \{-1\}, \mathcal{P}_3 = \{j, j^2\}, \mathcal{P}_4 = \{i, -i\}, \mathcal{P}_5 = \mu_5 \setminus \{1\}$$

$$\mathcal{P}_6 = \{-j, -j^2\}, \mathcal{P}_7 = \mu_7 \setminus \{1\}, \dots, \mathcal{P}_{12} = \{e^{i\pi/6}, e^{-i\pi/6}, e^{5i\pi/6}, e^{-5i\pi/6}\}.$$

THÉORÈME VI.8.4

|| Soit n un entier ≥ 2 . Le seul sous-groupe fini de cardinal n de (\mathbb{C}^*, \times) est μ_n .

Démonstration :

Soit G un sous-groupe fini de \mathbb{C}^* de cardinal n . D'abord G est un sous-groupe de \mathbb{U} , car si $z \in G$, $z^n = 1$ (l'ordre de chaque élément divise le cardinal du groupe : théorème de Lagrange vu au § V.2). Puisque $z^n = 1$, c'est que $G \subset \mu_n$. Mais de $\text{card}(G) = n = \text{card}(\mu_n)$ il résulte que $G = \mu_n$. ■

Ce théorème sera revu plus loin dans un cadre plus général.

Images des μ_n dans le plan d'Argand-Cauchy

Rappelons que, pour $z \in \mathbb{C}^*$ on note $\arg_0(z)$ l'unique élément de $\arg(z) \cap [0, 2\pi[$.

Posons $\zeta_1 = \exp\left(\frac{2i\pi}{n}\right)$ et $\zeta_k = (\zeta_1)^k$. On a donc

$$\theta_k = \arg_0(\zeta_k) = \frac{2k\pi}{n} \quad \text{pour } 0 \leq k \leq n-1.$$

Les points ζ_k du plan d'Argand-Cauchy sont les sommets d'une ligne polygonale \mathcal{L}_1 dite régulière, convexe, inscrite dans le cercle \mathbb{U} ; la convexité est équivalente au fait que

$$0 = \theta_0 < \theta_1 < \dots < \theta_{n-1} < 2\pi.$$

La longueur du côté de ce polygone est $l_{n,1} = 2 \sin \frac{\pi}{n}$, l'apothème $\cos \frac{\pi}{n}$.

Considérons maintenant une racine n -ième primitive ζ_k ($k \in \llbracket 0, n-1 \rrbracket$ et k premier avec n). En joignant par un segment de droite les sommets 1 à ζ_k , ζ_k à $(\zeta_k)^2$, ..., $(\zeta_k)^{n-1}$ à 1, on obtient une ligne polygonale régulière \mathcal{L}_k inscrite dans \mathbb{U} . Les sommets de \mathcal{L}_k sont évidemment les éléments de μ_n puisque l'application

$$\llbracket 0, n-1 \rrbracket \rightarrow \mu_n, q \mapsto (\zeta_k)^q$$

est une bijection. La longueur du côté du polygone \mathcal{L}_k est

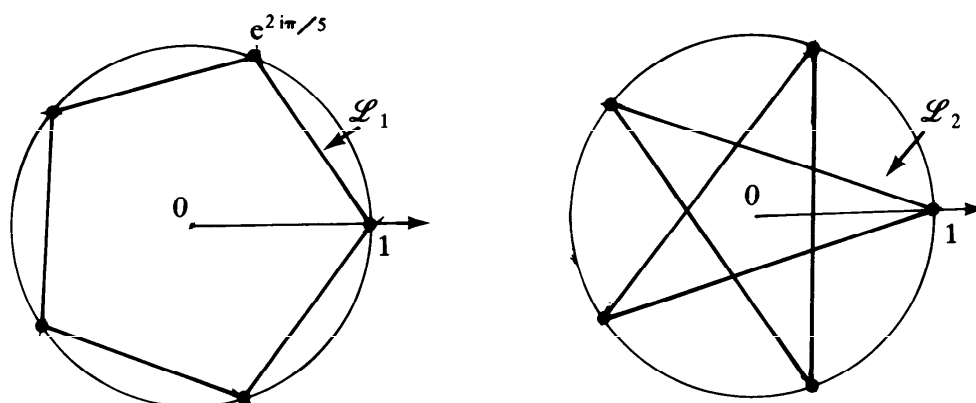
$$l_{n,k} = 2 \sin \frac{k\pi}{n} = |\zeta_k - 1| ,$$

mais il est clair que si $k \notin \{1, n-1\}$, en posant $\varphi_q = \arg_0((\zeta_k)^q)$, les arguments successifs ne vérifient plus

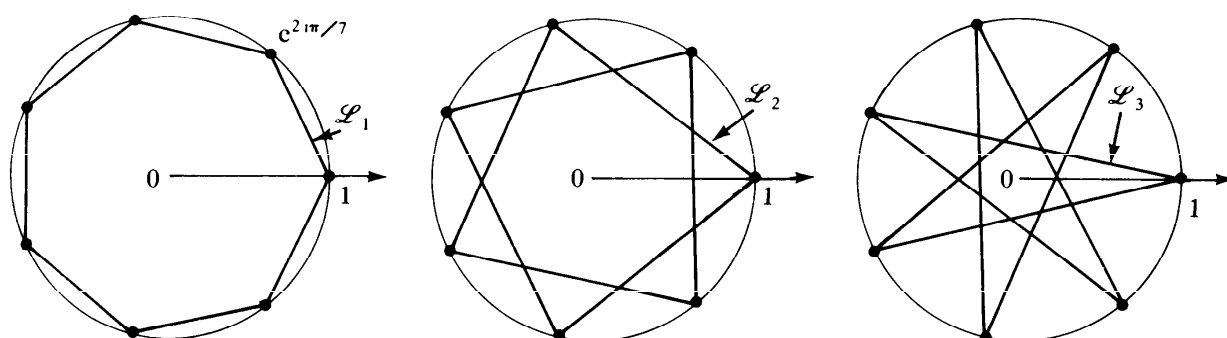
$$0 = \varphi_0 < \varphi_1 < \dots < \varphi_{n-1} < 2\pi \quad \text{ou} \quad \varphi_n = 0 < \varphi_{n-1} < \dots < \varphi_1 < 2\pi ,$$

et la ligne \mathcal{L}_k n'est plus convexe. Notons que si $k_1 + k_2 = n$ les lignes \mathcal{L}_{k_1} et \mathcal{L}_{k_2} sont égales car $\zeta_{k_1} = \bar{\zeta}_{k_2}$; si $k_1 \neq k_2$ et $k_1 + k_2 \neq n$ elles sont distinctes puisque non isométriques. Finalement nous obtenons $\frac{1}{2} \varphi(n)$ polygones \mathcal{L}_k distincts lorsque k décrit $\llbracket 0, n-1 \rrbracket$ en restant premier avec n . Un et un seul de ces polygones est convexe, les autres sont *étoilés*. Voici deux illustrations graphiques de cette situation, et le lecteur pourra multiplier à loisir ces exemples.

Exemple 3 : $n = 5$ (un pentagone convexe et un étoilé).



Exemple 4 : $n = 7$ (un heptagone convexe et deux étoilés).



Exercice 1 : Montrer que si $z \in \mathbb{C} \setminus \mathbb{R}$, l'argument principal de z vérifie selon le cas :

$$\begin{aligned} \text{si } \operatorname{Im}(z) > 0 \quad \operatorname{Arg} z &= \frac{\pi}{2} - \operatorname{Arctg} \frac{\operatorname{Re}(z)}{\operatorname{Im}(z)} \quad \text{et} \quad \operatorname{Arg} z > 0 \\ \text{si } \operatorname{Im}(z) < 0 \quad \operatorname{Arg} z &= \frac{-\pi}{2} - \operatorname{Arctg} \frac{\operatorname{Re}(z)}{\operatorname{Im}(z)} \quad \text{et} \quad \operatorname{Arg} z < 0 \end{aligned}$$

Vérifier également que pour $z \in \mathbb{C} \setminus \mathbb{R}_-$,

$$\operatorname{Arg} z = 2 \operatorname{Arctg} \frac{\operatorname{Im}(z)}{|z| + \operatorname{Re}(z)}.$$

Exercice 2 : Expliquer pourquoi, contrairement à \mathbb{R} , il n'existe pas de relation d'ordre total dans \mathbb{C} compatible avec l'addition et la multiplication de \mathbb{C} , ce qui en ferait un *corps ordonné*. Montrer de même que \mathbb{U} (resp. μ_n) ne sont pas des groupes ordonnables.

Exercice 3 : a) Soit z_1 et z_2 dans $\mathbb{C} \setminus \mathbb{R}_-$ tels que $z_1 z_2 \notin \mathbb{R}_-$. Calculer $\operatorname{Arg}(z_1 + z_2)$ en fonction de $\operatorname{Arg} z_1$ et de $\operatorname{Arg} z_2$.

b) Généraliser au produit de n facteurs.

Exercice 4 : Calculer les racines cubiques de $\frac{1+i}{\sqrt{2}}$ et mettre le résultat sous forme algébrique.

Exercice 5 : a) Soit $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}$. Calculer $\sum_{\zeta \in \mu_n} \zeta^k$.

b) Soit $\zeta_1 = e^{2i\pi/n}$. On pose $S = \sum_{p=0}^{n-1} \zeta_1^{p^2}$. Calculer $|S|^2$.

Exercice 6 : Si $a \in \mathbb{R}$, résoudre dans \mathbb{C} l'équation $\left(\frac{1+iz}{1-iz}\right)^n = \frac{1+ia}{1-ia}$.

Exercice 7 : Montrer que tout nombre complexe de module 1 peut s'écrire sous la forme $\frac{1+ia}{1-ia}$ où a est un réel dépendant de $\operatorname{Arg} z$.

Exercice 8 : Calculer le module et l'argument principal du complexe $\left(\frac{1+i\sqrt{3}}{1-i}\right)^{20}$; de $\left(\frac{1+\sqrt{2}+i}{1+\sqrt{2}-i}\right)^{20}$; de $\frac{1+\sin \theta + i \cos \theta}{1+\sin \theta - i \cos \theta}$.

Exercice 9 : Simplifier $\left(\frac{3\sqrt{3}-i}{\sqrt{3}+2i}\right)^{25}$; $(1+i)^n + (1-i)^n$; $\frac{(1+i)^4}{(1-i)^3} + \frac{(1-i)^4}{(1+i)^3}$.

Exercice 10 : Trouver z tel que $|z| = |z-4|$ et $\operatorname{Arg} z = \operatorname{Arg}(z+1+i)$.

Exercice 11 : Trouver 3 nombres de module 1 dont la somme et le produit sont égaux à 1.

Exercice 12 : Résoudre le système à 2 inconnues réelles :

$$x^6 - 15x^4y^2 + 15x^2y^4 - y^6 = 1 \quad \text{et} \quad 3x^5y - 10x^3y^3 + 3xy^5 = 0.$$

Exercice 13 : Résoudre le système à deux inconnues complexes : $(1+t)^n = (1-t)^n$ et $z^2 + t^2 = 1$, où $n \in \mathbb{N}^*$ est donné.

Exercice 14 : Déterminer A sachant que ses trois racines cubiques vérifient la relation $z_1 - z_2 = \frac{1}{z_3}$.

Exercice 15 : Soit α le côté de l'heptagone régulier convexe, β et γ les côtés des heptagones étoilés inscrits dans \mathbb{U} . Montrer que α^2 , β^2 , γ^2 sont racines de l'équation $X^3 - 7X^2 + 14X - 7 = 0$. Vérifier que $\frac{1}{\alpha} = \frac{1}{\beta} + \frac{1}{\gamma}$.

Exercice 16 : On admet ici que tout sous-groupe additif de \mathbb{R} non réduit à $\{0\}$ est soit du type $a\mathbb{Z}$, avec $a > 0$, soit *partout dense* dans \mathbb{R} (c'est-à-dire d'intersection non vide avec tout intervalle qui n'est ni vide ni réduit à un point). On appelle *partie discrète* de \mathbb{C} tout $A \subset \mathbb{C}$ tel que $\forall a \in A \exists r > 0 \mid \mathcal{B}(a, r) \cap A = \{a\}$, la notation $\mathcal{B}(a, r)$ désignant la boule fermée de centre a et de rayon r (c'est-à-dire $\{z \in \mathbb{C} \mid |z-a| \leq r\}$).

Montrer que si G est un sous-groupe discret de \mathbb{U} , alors G est fini (donc égal à l'un des μ_n , $n \in \mathbb{N}^*$).

Indication : Utiliser l'homomorphisme $\mathcal{E} : \mathbb{R} \rightarrow \mathbb{U}$, $x \mapsto e^{ix}$.

Exercice 17 : Caractères d'un groupe. Pour chaque groupe G , on munit l'ensemble $\widehat{G} = \text{Hom}(G, \mathbb{U})$ des homomorphismes de groupe de G dans \mathbb{U} de sa structure naturelle de groupe (si $f \in \widehat{G}$ et $g \in \widehat{G}$, $(\forall x \in G) (fg)(x) = f(x)g(x)$).

a) Si G est cyclique, démontrer que \widehat{G} est cyclique et de même cardinal.

b) On suppose $G = G_1 \times G_2 \times \cdots \times G_p$ où G_k est cyclique, de cardinal n_k . Démontrer que \widehat{G} est isomorphe à $\widehat{G}_1 \times \widehat{G}_2 \times \cdots \times \widehat{G}_p$ et en déduire que G est isomorphe à $\widehat{\widehat{G}}$.

c) Trouver \widehat{G} lorsque G est l'un des groupes suivants rencontrés au chapitre V : 1) le groupe du carré 2) le groupe quaternionique 3) le groupe \mathbb{U}_4 4) le groupe diédral D_n .

N.B. Le groupe \widehat{G} est appelé le groupe des *caractères* de G .

Exercice 18 : On reprend les notations de l'exercice 17. Pour chaque $\lambda \in \mathbb{R}$, on considère l'élément f_λ de $\widehat{\mathbb{Q}}$ (groupe des caractères du groupe additif \mathbb{Q}) défini par : $(\forall x \in \mathbb{Q}) f_\lambda(x) = \exp(i\lambda x)$. Étudier l'homomorphisme du groupe $(\mathbb{R}, +)$ dans le groupe $\widehat{\widehat{G}}$.

Exercice 19 : Montrer que les groupes abéliens $\widehat{\mathbb{Q}}_+^*$ et $(\mathbb{R}/2\pi\mathbb{Z})^{\mathbb{N}}$ sont isomorphes.

Exercice 20 : Soit $a = e^{2i\alpha}$, $b = e^{2i\beta}$, $c = e^{2i\gamma}$. Calculer le module et un argument de $\frac{c^2 + ab}{ab}$, de $\frac{a+b}{ab+bc+ca+c^2}$.

§ VI.9 NOMBRES COMPLEXES ET GÉOMÉTRIE

Soit \mathcal{P} un plan affine euclidien orienté. Munissons \mathcal{P} d'un repère orthonormé direct $\mathcal{R} = (O, \vec{e}_1, \vec{e}_2)$ et faisons correspondre à tout élément $x + iy$ de \mathbb{C} , considéré lui-même comme un plan affine euclidien orienté (plan d'Argand-Cauchy), le point de \mathcal{P} : $O + x\vec{e}_1 + y\vec{e}_2$. Il s'agit de toute évidence d'une bijection isométrique qui préserve l'orientation de \mathbb{C} sur \mathcal{P} . Dans cette application on dit que $z = x + iy$ est l'**affixe** du point $M = O + x\vec{e}_1 + y\vec{e}_2$ (relativement au repère \mathcal{R}), et que M est l'**image** (relative à \mathcal{R}) du nombre complexe z . On peut donc limiter l'étude des **applications des nombres complexes à la géométrie euclidienne** du plan d'Argand-Cauchy, puisqu'on pourra ensuite transporter les propriétés au plan \mathcal{P} par une bijection du type précédent.

De même qu'il existe sur \mathbb{R} un grand nombre de structures (algébriques, d'ordre, topologiques), de même \mathbb{C} est considéré tantôt sous son aspect algébrique (groupe additif, **corps**, algèbre sur \mathbb{R}), tantôt sous son aspect géométrique (**plan affine euclidien**, **\mathbb{R} -ev euclidien**, espace *directeur* du précédent, avec sa métrique associée).

Si, par exemple, z et t sont deux **nombres complexes**, on peut aussi les considérer comme des **points** du plan d'Argand-Cauchy ; le vecteur \vec{zt} n'est autre que le nombre complexe $t - z$ (et nous n'hésiterons pas à écrire

$$\boxed{\vec{zt} = t - z},$$

et c'est aussi le vecteur d'une **translation** opérant sur le plan affine \mathbb{C} ; la *norme* euclidienne du vecteur \overrightarrow{zt} n'est autre que le *module* de $t - z$, et ces remarques très simples peuvent déjà être exploitées.

Mais, pour étudier la notion d'angle, qui est plus délicate, quelques préliminaires nous paraissent indispensables, et d'abord la présentation des classes de réels mod (a) .

Groupe-quotients $\mathbb{R}/a\mathbb{Z}$, $a > 0$

Soit $a > 0$ un réel fixé. L'ensemble $a\mathbb{Z} = \{ka\}_{k \in \mathbb{Z}}$ est un sous-groupe additif de \mathbb{R} . La relation binaire définie sur \mathbb{R} par (1) $x - y \in a\mathbb{Z}$ est d'équivalence. On la note $x \equiv y \text{ mod } (a)$. Si $x_0 \in \mathbb{R}$, la classe d'équivalence de x_0 est $\{x_0 + ka\}_{k \in \mathbb{Z}}$, notée en abrégé $x_0 + a\mathbb{Z}$. Il est clair que $x'_0 \equiv x_0 \text{ mod } (a)$ et $y'_0 \equiv y_0 \text{ mod } (a)$ entraînent $x'_0 + y'_0 \equiv x_0 + y_0 \text{ mod } (a)$. On peut donc définir, sur l'ensemble quotient de \mathbb{R} par la relation (1), ensemble noté $\mathbb{R}/a\mathbb{Z}$, une addition par la condition : si $X \in \mathbb{R}/a\mathbb{Z}$ et $Y \in \mathbb{R}/a\mathbb{Z}$, $X + Y$ est la classe d'équivalence de tous les réels $x + y$, pour $x \in X$ et $y \in Y$. Muni de cette loi $+$, $\mathbb{R}/a\mathbb{Z}$ est un **groupe abélien** d'élément neutre $\text{Cl}(0) = a\mathbb{Z}$, où l'opposé de $X = x + a\mathbb{Z}$ est $-x + a\mathbb{Z}$. Ce groupe noté encore $\mathbb{R}/a\mathbb{Z}$ est appelé **groupe-quotient de $(\mathbb{R}, +)$ par le sous-groupe $a\mathbb{Z}$** .

L'application canonique $\mathcal{C} : \mathbb{R} \rightarrow \mathbb{R}/a\mathbb{Z}$, $x \mapsto \text{Cl}(x) = x + a\mathbb{Z}$, est un **homomorphisme de groupes surjectif, de noyau $a\mathbb{Z}$** . Sa restriction à $[\alpha, \alpha + a[$ est une bijection de $[\alpha, \alpha + a[$ sur $\mathbb{R}/a\mathbb{Z}$, pour tout réel α . En particulier, pour tout $x \in \mathbb{R}$, il existe un unique élément $E_a(x) \in \mathbb{Z}$ tel que $x - E_a(x) \times a \in [0, a[$. En appelant E_1 la fonction **partie entière**, on voit que $E_a(x)$ n'est autre que $E_1\left(\frac{x}{a}\right)$. Considérons en particulier l'homomorphisme exponentiel $\mathcal{E} : (\mathbb{R}, +) \rightarrow \mathbb{U}$, $x \mapsto \exp(ix)$ qui est surjectif et de noyau $2\pi\mathbb{Z}$, et sa décomposition canonique (cf. théorème I.5.4)

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\mathcal{E}} & \mathbb{U} \\ \mathcal{C} = \downarrow \text{can} & & \uparrow \text{Id}_{\mathbb{U}} \\ \mathbb{R}/2\pi\mathbb{Z} & \xrightarrow{\overline{\mathcal{E}}} & \mathbb{U} = \text{Im}(\mathcal{E}) \end{array}$$

On construit ainsi une application $\overline{\mathcal{E}}$ qui est bijective et dont on voit aisément que c'est un **isomorphisme de groupes** dès que $\mathbb{R}/2\pi\mathbb{Z}$ est muni de sa structure de groupe-quotient. Nous dirons que $\overline{\mathcal{E}}$ est l'**isomorphisme canoniquement déduit de \mathcal{E}** . Si $X = x + 2\pi\mathbb{Z}$, son image $\overline{\mathcal{E}}(X)$ est $\exp(ix)$. L'isomorphisme réciproque de \mathcal{E} transforme le produit de deux nombres complexes de module 1 en la somme mod (2π) de deux réels. Si $u \in \mathbb{U}$, les réels θ tels que $\overline{\mathcal{E}}(\text{Cl}(\theta)) = u$ ne sont autres que les **déterminations de l'argument de u** , et on a : $\overline{\mathcal{E}}^{-1}(u) = \arg(u)$.

Angles orientés de vecteurs et de demi-droites

Nous étudierons au Tome 3 le groupe des matrices orthogonales directes réelles d'ordre 2, noté $\Omega_2^+(\mathbb{R})$. Retenons simplement que ce groupe est constitué des matrices de la forme :

$$(2) \quad \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad (a, b) \in \mathbb{R}^2, \quad a^2 + b^2 = 1.$$

On constate immédiatement que la multiplication des matrices de la forme (2) correspond exactement au produit des nombres complexes associés $a + ib$. De manière précise :

THÉORÈME VI.9.1

$$\left\| \begin{array}{l} \text{L'application } \mathcal{M} : \\ \mathbb{U} \rightarrow \Omega_2^+(\mathbb{R}), \quad u = c + is \mapsto \begin{pmatrix} c & -s \\ s & c \end{pmatrix}, \\ \text{où } (c, s) \in \mathbb{R}^2 \text{ et } c^2 + s^2 = 1, \text{ est un isomorphisme de groupes.} \end{array} \right.$$

Or nous verrons que dans toutes les bases orthonormées directes du plan vectoriel d'Argand-Cauchy, et en particulier dans la base $(1, i)$, à chaque matrice $R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \in \Omega_2^+(\mathbb{R})$ est attachée une seule et même **rotation** $\rho(R)$. Plus précisément, l'application $\rho : \Omega_2^+(\mathbb{R}) \rightarrow \text{SO}(\mathbb{C})$, $R \mapsto \rho(R)$ est un isomorphisme de groupes, $\text{SO}(\mathbb{C})$ désignant le groupe des rotations du plan vectoriel d'Argand-Cauchy.

En effet, pour chaque « matrice de rotation » $R = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$, le vecteur $z = x + iy \in \mathbb{C}$ ($(x, y) \in \mathbb{R}^2$) est transformé en le vecteur

$$\rho(R)(z) = cx - sy + i(sx + cy)$$

et on reconnaît bien l'image du nombre $(c + is)(x + iy)$, d'où

THÉORÈME VI.9.2

$$\left\| \begin{array}{l} \text{A chaque } u \in \mathbb{U}, \text{ faisons correspondre l'application } \rho_u : \mathbb{C} \rightarrow \mathbb{C}, \\ z \mapsto uz. \text{ Alors l'application } u \mapsto \rho_u \text{ est un isomorphisme du groupe} \\ \mathbb{U} \text{ sur le groupe des rotations vectorielles } \text{SO}(\mathbb{C}) \text{ du plan d'Argand-} \\ \text{Cauchy.} \end{array} \right.$$

Avec cette interprétation, à i correspond une rotation de $\frac{+\pi}{2}$, et à i^2 correspond une rotation de π c'est-à-dire un demi-tour, ce qui enlève tout mystère à l'égalité $i^2 = -1$. De manière générale, en appelant **angle orienté** d'une rotation vectorielle ρ toute détermination de l'argument de $u \in \mathbb{U}$ tel que $\rho = \rho_u$, la composée $\rho_1 \circ \rho_2$ des deux rotations d'angle orienté θ_1 et θ_2 est la rotation d'angle $\theta_1 + \theta_2$.

Grâce à l'identification des rotations vectorielles du plan et des éléments du groupe \mathbb{U} , il devient tout à fait évident que, étant donnés deux vecteurs unitaires quelconques de \mathbb{C} , **il existe une et une seule rotation $\rho \in \text{SO}(\mathbb{C})$ telle que $v = \rho(u)$** (celle correspondant à $w = \frac{v}{u}$).

DÉFINITION VI.9.1

Si $z \in \mathbb{C}^*$ et $t \in \mathbb{C}^*$, on appelle **angle orienté de z et de t** tout angle orienté de l'unique rotation vectorielle ρ qui envoie $u = \frac{z}{|z|}$ sur $v = \frac{t}{|t|}$. On appelle **angle orienté de deux demi-droites vectorielles \vec{D}_1 et \vec{D}_2 de \mathbb{C}** tout angle orienté de z_1 et de z_2 , où z_1 est un vecteur directeur de \vec{D}_1 et z_2 de \vec{D}_2 (c'est-à-dire que

$$\vec{D}_i = \{ \lambda z_i \}_{\lambda \in \mathbb{R}_+} \quad (i = 1, 2)).$$

D'après la structure de groupe de (\mathbb{U}, \times) on a immédiatement :

THÉORÈME VI.9.3

Soit $z_0, z_1, \dots, z_n \in \mathbb{C}$ et soit θ_k un angle orienté de z_{k-1} et de z_k ($1 \leq k \leq n$). Alors $\theta_1 + \theta_2 + \dots + \theta_n$ est un angle orienté de z_0 et de z_n .

C'est la *relation de Chasles* ⁽¹⁾ pour les angles de demi-droites (vectorielles) de \mathbb{C} . Nous noterons indifféremment $(\vec{D}_1, \vec{D}_2) = \theta \bmod (2\pi)$ ou $(z_1, z_2) = \theta \bmod (2\pi)$.

Angles orientés de droites vectorielles

Considérons maintenant l'ensemble \mathcal{G}_1 des droites vectorielles de \mathbb{C} . Sur chaque droite D on peut choisir un vecteur directeur unitaire u tel que $D = \{ \lambda u \}_{\lambda \in \mathbb{R}}$, mais il est clair que l'on a aussi $D = \{ \lambda (-u) \}_{\lambda \in \mathbb{R}}$. Chaque droite est donc caractérisée par une **paire** $\{u, -u\}$ d'éléments de \mathbb{U} . Nous noterons \mathbb{V} l'ensemble de ces paires $\{u, -u\}_{u \in \mathbb{U}}$. Il est très facile de définir dans \mathbb{V} une multiplication à partir de celle de \mathbb{U} et l'on obtient un **groupe abélien** (de manière précise \mathbb{V} est le **groupe-quotient** de \mathbb{U} par son sous-groupe $\mu_2 = \{-1, +1\}$). L'application canonique $g: \mathbb{U} \rightarrow \mathbb{V}$, $u \mapsto \{u, -u\}$ est un homomorphisme de groupes surjectif, de noyau μ_2 . Pratiquement on reconnaît que **deux éléments u_1 et u_2 de \mathbb{U} ont même image dans \mathbb{V} au fait que $u_1^2 = u_2^2$** .

⁽¹⁾ Michel Chasles (1793-1880), mathématicien français qui a introduit en géométrie les grandeurs « orientées », auteur d'un « Aperçu historique sur l'origine et le développement des méthodes en géométrie ».

En composant g avec l'homomorphisme exponentiel \mathcal{E} on obtient l'homomorphisme $\mathbb{R} \rightarrow \mathbb{V}$, $\theta \mapsto \{e^{i\theta}, -e^{i\theta}\} = \{e^{i\theta}, e^{i(\theta+\pi)}\}$, de noyau $\pi\mathbb{Z}$, ce qui donne le diagramme suivant :

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\mathcal{E}} & \mathbb{U} \\ \mathcal{D} = \text{can} \downarrow & & \downarrow g \\ \mathbb{R}/\pi\mathbb{Z} & \xrightarrow{\mathcal{F}} & \mathbb{V} \end{array}$$

Il est clair que, pour que deux réels θ_1 et θ_2 définissent le même élément de \mathbb{V} , il faut et il suffit que $\theta_1 \equiv \theta_2 \pmod{\pi}$.

De même qu'à chaque $u \in \mathbb{U}$ nous avons fait correspondre une rotation vectorielle ρ_u , à chaque $v = \{u, -u\} \in \mathbb{V}$ on peut associer la paire de rotations $\{\rho_u, \rho_{-u}\}$. Si $D \in \mathcal{G}_1$, $\rho_u(D)$ et $\rho_{-u}(D)$ sont une seule et même droite que l'on peut noter $v * D$. Cela signifie que l'application $\mathbb{V} * \mathcal{G}_1 \rightarrow \mathcal{G}_1$, $(v, D) \mapsto v * D$ est une **action à gauche du groupe \mathbb{V} sur l'ensemble \mathcal{G}_1** , visiblement **fidèle et transitive** ($\forall D_1 \in \mathcal{G}_1$, $\forall D_2 \in \mathcal{G}$, il existe un et un seul $v \in \mathbb{V}$ tel que $D_2 = v * D_1$).

DÉFINITION VI.9.2

$\left\{ \begin{array}{l} \text{Soit } D_1 \text{ et } D_2 \text{ deux droites vectorielles de } \mathbb{C}. \text{ On appelle } \textbf{angle orienté} \\ \textbf{de } D_1 \text{ et } D_2 \text{ tout réel } \theta \text{ tel que } v = \{e^{i\theta}, -e^{i\theta}\}, v \text{ désignant l'unique} \\ v \in \mathbb{V} \text{ tel que } v * D_1 = D_2. \end{array} \right.$

THÉORÈME VI.9.4

$\left\| \begin{array}{l} \text{Soit } D_0, D_1, \dots, D_n \text{ des droites vectorielles de } \mathbb{C} \text{ et soit } \theta_k \text{ un angle} \\ \text{orienté de } D_{k-1} \text{ et de } D_k \text{ (} 1 \leq k \leq n \text{)}. \text{ Alors } \theta_1 + \theta_2 + \dots + \theta_n \text{ est} \\ \text{un angle orienté de } D_0 \text{ et de } D_n. \end{array} \right.$

C'est la relation de Chasles pour les angles de droites ; nous noterons un tel angle $\widehat{(D_1, D_2)} = \theta \pmod{\pi}$, puisque l'ensemble des angles orientés de D_1 et de D_2 est une classe de réels mod (π) .

Il est facile de voir que les groupes \mathbb{U} et \mathbb{V} sont isomorphes. En effet il suffit de considérer l'application $f: \mathbb{V} \rightarrow \mathbb{U}$, $v = \{u, -u\} \mapsto u^2$. f est visiblement un isomorphisme de groupes (en se reportant au § V.7, cet isomorphisme s'obtient par passage au quotient à partir de l'homomorphisme surjectif $\mathbb{U} \rightarrow \mathbb{U}$, $u \mapsto u^2$ dont le noyau est μ_2).

Il en résulte que les groupes $\mathbb{R}/2\pi\mathbb{Z}$ et $\mathbb{R}/\pi\mathbb{Z}$ sont eux aussi isomorphes, ce qui n'a évidemment rien d'étonnant, si l'on utilise l'application : $\mathbb{R}/\pi\mathbb{Z} \rightarrow \mathbb{R}/2\pi\mathbb{Z}$, $\theta \pmod{\pi} \mapsto 2\theta \pmod{2\pi}$ qui est bien un isomorphisme.

Exercice 1 : Soit E l'ensemble des entiers de Gauss $\{a + ib\}_{(a,b) \in \mathbb{Z}^2}$ et G le sous-groupe du groupe $\text{Is}(\mathbb{C})$ formé des isométries f du plan d'Argand-Cauchy qui

$f(E) = E$. On notera $G_+ = G \cap \text{Is}_+(\mathbb{C})$ et on utilisera l'opération à gauche naturelle de ces groupes sur E .

- Trouver les stabilisateurs $(G_+)_z$ et G_z d'un point $z \in E$.
- Trouver tous les sous-groupes finis de G et de G_+ .
- Trouver tous les sous-groupes distingués de G_+ .

Exercice 2 : Mêmes questions qu'à l'exercice 1 en remplaçant l'ensemble E par l'ensemble $J = \{a + bj\}_{(a,b) \in \mathbb{Z}^2}$.

Exercice 3 : Soit $n \in \mathbb{N}^*$. En utilisant l'homomorphisme $\mathbb{U} \rightarrow \mathbb{U}$, $z \mapsto z^n$, démontrer que le groupe \mathbb{U} est isomorphe au groupe-quotient \mathbb{U}/μ_n .

Exercice 4 : Soit l'équation $z^2 - 2z e^{i\theta} + 1 = 0$ où θ est un paramètre réel, et où l'inconnue est $z \in \mathbb{C}$. Résoudre cette équation. Quel est le lieu géométrique des racines dans le plan d'Argand-Cauchy lorsque θ décrit \mathbb{R} .

Exercice 5 : Soit a, b, c trois points non alignés dans \mathbb{C} . Calculer en fonction de a, b, c le centre du cercle circonscrit, l'orthocentre, le centre du cercle inscrit, et les centres des cercles exinscrits au triangle a, b, c .

Exercice 6 : Montrer que la condition nécessaire et suffisante pour que le triangle (a, b, c) soit équilatéral de sens direct est : $(1) a + bj + cj^2 = 0$. En déduire pour quelles valeurs de z les points d'affixes az^2, a^2z et z^3 sont les sommets d'un triangle équilatéral.

Exercice 7 : Soit a, b, c, d tels que $a + c = b + d$ et $a + ib = c + id$. Quelle est la figure formée par les 4 images de a, b, c, d dans le plan d'Argand-Cauchy ? En déduire l'existence d'un complexe z tel que

$$(z - a)^4 = (z - b)^4 = (z - c)^4 = (z - d)^4.$$

Exercice 8 : Déterminer z pour que les images de z, z^2 et z^3 soient les sommets d'un triangle rectangle.

Exercice 9 : Soit p et q dans \mathbb{C} , z_1, z_2, z_3 les racines de l'équation $z^3 + pz + q = 0$ et a_1, a_2 celles de $3z^2 + p = 0$. On suppose tous ces points distincts. Démontrer

$$(\widehat{z_3 - z_1, z_3 - a_1}) = (\widehat{z_3 - a_2, z_3 - z_2}) \bmod (2\pi)$$

et les relations analogues.

Exercice 10 :

a) Soient $n \in \mathbb{N}^*$ et $v \in \mathbb{V}$. Trouver les $w \in \mathbb{V}$ tels que $w^n = v$.

b) Soit alors deux droites vectorielles D_1, D_2 dans \mathbb{C} . Trouver les droites vectorielles D telles que $n(\widehat{D_1, D}) = (\widehat{D_1, D_2}) \bmod (\pi)$ (bissectrices de (D_1, D_2) pour $n = 2$, trisectrices pour $n = 3$, n -sectrices du couple (D_1, D_2) pour n quelconque).

c) Soit a, b, c trois points non alignés dans \mathbb{C} . On note $\alpha = \text{Arg} \frac{c-a}{b-a}$, $\beta = \text{Arg} \frac{a-b}{c-b}$, $\gamma = \text{Arg} \frac{b-c}{a-c}$ qu'on suppose > 0 , de telle sorte que $\alpha + \beta + \gamma = \pi$. Soit $D_1(a)$ (resp. $\Delta_1(a)$) la trisectrice des droites $D(a, b)$ et $D(a, c)$ telle que

$$(D(a, b), D_1(a)) = \frac{\alpha}{3} \bmod (\pi) \text{ (resp. } (D(a, b), \Delta_1(a)) = \frac{2\alpha}{3} \bmod (\pi)).$$

On définit de même $D_1(b), \Delta_1(b), D_1(c)$ et $\Delta_1(c)$. Les droites $D_1(b)$ et $\Delta_2(c)$ se rencontrent en un point noté s_a . Démontrer le théorème de Morley affirmant que le triangle (s_a, s_b, s_c) est équilatéral, s_b et s_c étant définis de façon analogue. De même $\Delta_1(b)$ et $D_1(c)$ se rencontrent en un point noté t_a , et on définit de même t_b et t_c . Alors le triangle (t_a, t_b, t_c) est équilatéral.

d) On considère maintenant les trois trisectrices $D_1(a), D_2(a), D_3(a)$ du couple $(D(a, b), D(a, c))$ et les trois trisectrices $\Delta_1(a), \Delta_2(a), \Delta_3(a)$ du couple $(D(a, c), D(a, b))$. On définit de même $D_1(b), D_2(b), \dots, \Delta_3(c)$. Combien y a-t-il de poin

18 droites ainsi obtenues, autres que a, b, c ? Avec ces points communs, peut-on former des triangles équilatéraux autres que les deux déjà rencontrés au c) ? Il est recommandé de faire une figure.

§ VI.10 NOMBRES COMPLEXES ET SIMILITUDES

Notations

Outre les notations utilisées au § VI.9 nous désignerons par σ la bijection $\mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$ et nous conviendrons que, dans le plan d'Argand-Cauchy, $\overrightarrow{\mathcal{P}}$ représente le groupe des *similitudes vectorielles*, $\overrightarrow{\mathcal{P}}_+$ celui des *similitudes vectorielles directes* ; \mathcal{S} le groupe des *similitudes*, \mathcal{S}_+ celui des *similitudes directes* ; $O(\mathbb{C})$ le groupe des *isométries vectorielles*, $SO(\mathbb{C})$ celui des *isométries vectorielles directes* ; $Is(\mathbb{C})$ le groupe des *isométries*, $Is_+(\mathbb{C})$ celui des *isométries directes* que l'on appelle encore des *déplacements* ; enfin \mathcal{T} le groupe des *translations* de \mathbb{C} .

On a par exemple $\mathcal{T} \subset Is_+(\mathbb{C})$, $\sigma \in O(\mathbb{C}) \setminus SO(\mathbb{C})$.

Pour toute **bijection affine** f de \mathbb{C} , nous noterons $L(f)$ sa *partie linéaire*, qui est donc un élément de $GL_{\mathbb{R}}(\mathbb{C})$, et par $\det(f)$ le déterminant de $L(f)$.

Pour tout sous-groupe G du groupe des bijections affines de \mathbb{C} (resp. du groupe linéaire de \mathbb{C}), nous considérerons toujours l'action à gauche naturelle de G sur \mathbb{C} définie par $(f, z) \mapsto f(z)$. Le stabilisateur de $z \in \mathbb{C}$ pour cette action sera noté G_z (rappelons que

$$G_z = \{f \in G \mid f(z) = z\}.$$

Groupe des similitudes vectorielles

Quand nous étudierons en Algèbre linéaire le groupe des matrices orthogonales réelles d'ordre 2, noté $\Omega_2(\mathbb{R})$, nous verrons qu'en plus des « matrices de rotation » de déterminant $+1$, qui constituent le groupe $\Omega_2^+(\mathbb{R})$, il contient aussi des matrices orthogonales indirectes, de déterminant -1 , de la forme (1) $\begin{pmatrix} c & s \\ s & -c \end{pmatrix}$, où $(c, s) \in \mathbb{R}^2$ et $c^2 + s^2 = 1$. A chaque $u = c + is \in \mathbb{U}((c, s) \in \mathbb{R}^2)$ on peut faire correspondre la matrice $\mathcal{N}(u)$ définie par (1) et se demander, le \mathbb{R} -ev \mathbb{C} étant rapporté à la base $(1, i)$, ce que représente cette matrice $\mathcal{N}(u)$. On constate d'abord que chaque point de la droite D_u d'équation cartésienne $(c-1)x + sy = 0$ est invariant dans la transformation de matrice $\mathcal{N}(u)$ et on reconnaît ensuite qu'il s'agit d'une **symétrie orthogonale** s_u autour de la droite vectorielle D_u . Si $u = e^{i\theta}$ ($\theta \in \mathbb{R}$), un vecteur directeur de D_u est $e^{i\theta/2}$. Or, en appliquant s_u au vecteur $z = x + iy$, on vérifie immédiatement que

$$s_u(z) = u\bar{z} = \rho_u \circ \sigma(z),$$

d'où $s_u = \rho_u \circ \sigma$. Compte tenu du théorème VI.9.2 et du fait que $\sigma \in O(\mathbb{C}) \setminus SO(\mathbb{C})$, il vient :

THÉORÈME VI.10.1

|| L'application $\mathbb{U} \rightarrow O(\mathbb{C}) \setminus SO(\mathbb{C})$, $u \mapsto \rho_u \circ \sigma$ est bijective. Si $u = e^{i\theta}$, l'axe de la symétrie orthogonale $s_u = \rho_u \circ \sigma$ est la droite dirigée par $e^{i\theta/2}$.

Rappelons qu'on appelle **similitude vectorielle** le produit (commutatif) d'une isométrie vectorielle par une homothétie (que l'on peut toujours supposer de rapport > 0). En particulier le produit d'une isométrie vectorielle directe par une homothétie est une similitude directe : si

$$\varphi = \lambda \omega \quad (\lambda > 0, \omega \in SO(\mathbb{C}))$$

est une telle similitude directe, λ est appelé le **rapport** de similitude (et l'on a $\det(\varphi) = \lambda^2 \det(\omega)$) et on appelle **angle** de la similitude φ tout angle orienté de ω . Il suffit de munir $\mathbb{R}_+^* \times O(\mathbb{C})$ de sa structure de groupe produit pour s'apercevoir que l'application

$$\mathbb{R}_+^* \times O(\mathbb{C}) \rightarrow \overrightarrow{\mathcal{P}}, \quad (\lambda, \omega) \mapsto \lambda \omega = (\lambda \text{Id}_{\mathbb{C}}) \circ \omega = \omega \circ (\lambda \text{Id}_{\mathbb{C}})$$

est un *isomorphisme de groupes*.

Si $\varphi = \lambda \omega$ ($\lambda > 0, \omega \in O(\mathbb{C}) \setminus SO(\mathbb{C})$) on dit que φ est une similitude indirecte et l'axe de la symétrie orthogonale ω s'appelle **droite principale** de la similitude φ .

THÉORÈME VI.10.2

|| Soit $a \in \mathbb{C}^*$; l'application $\varphi_a : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto az$ est une **similitude directe**, de rapport $|a|$, d'angle tout réel $\theta \in \arg(a)$, et de déterminant $|a|^2$. L'application $\mathbb{C}^* \rightarrow \overrightarrow{\mathcal{P}}_+, a \mapsto \varphi_a$ est un *isomorphisme de groupes*.

Démonstration :

Posons $\frac{a}{|a|} = u$; alors $\varphi_a = (|a| \text{Id}_{\mathbb{C}}) \circ \rho_u$, où ρ_u est la rotation vectorielle associée à u . Donc $\varphi_a \in \overrightarrow{\mathcal{P}}_+$, et

$$\det(\varphi_a) = |a|^2 \det(\rho_u) = |a|^2.$$

De $\varphi_{ab} = \varphi_a \circ \varphi_b$ on déduit bien que l'application $a \mapsto \varphi_a$ est un homomorphisme de groupes. Cet homomorphisme est injectif car $\varphi_a(1) = a$ pour $a \in \mathbb{C}^*$; et il est surjectif, car la similitude directe $\lambda \omega$, où $\lambda \in \mathbb{R}_+^*$ et $\omega \in SO(\mathbb{C})$, n'est autre que φ_a , avec $a = \lambda u$ et où $u \in \mathbb{U}$ vérifie $\rho_u = \omega$. ■

Ainsi chaque nombre complexe $\neq 0$ apparaît maintenant aussi comme un opérateur de similitude.

Les théorèmes VI.10.1 et VI.10.2 entraînent immédiatement :

THÉORÈME VI.10.3

|| Soit $a \in \mathbb{C}^*$; l'application $\psi_a = \varphi_a \circ \sigma : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto a\bar{z}$, est une **similitude indirecte**, de rapport $|a|$, et de droite principale D dirigée par $e^{i\theta/2}$ si $\frac{a}{|a|} = e^{i\theta}$. L'application $a \mapsto \psi_a$ définit une bijection de l'ensemble \mathbb{C}^* sur l'ensemble $\overrightarrow{\mathcal{P}} \setminus \overrightarrow{\mathcal{P}}_+$.

On peut passer maintenant à l'étude des similitudes dans le plan affine \mathbb{C} .

Groupe des similitudes directes de \mathbb{C}

THÉORÈME VI.10.4

|| Si $(a, b) \in \mathbb{C}^* \times \mathbb{C}$, l'application $f_{a,b} : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto az + b$ est une **similitude directe**, de partie linéaire $\varphi_a : z \mapsto az$.

Démonstration :

Soit $(a, b) \in \mathbb{C}^* \times \mathbb{C}$. Si $z \in \mathbb{C}$ on a :

$$\overrightarrow{f_{a,b}(0) f_{a,b}(z)} = az = a(z - 0) = \varphi_a(\overrightarrow{Oz}),$$

ce qui montre bien que la partie linéaire de $f_{a,b}$ est bien φ_a . Il est clair par ailleurs que $f_{a,b}$, produit d'une similitude directe vectorielle et d'une translation est bien une similitude directe dans le plan affine \mathbb{C} . ■

On sait déjà que les similitudes directes forment un groupe, mais il est très facile de le retrouver en calculant la composée $f_{a,b} \circ f_{a',b'}$: c'est $f_{c,d}$ tel que l'image de $z \in \mathbb{C}$ soit

$$a(a'z + b') + b = aa'z + ab' + b,$$

et l'application inverse de $f_{a,b}$: c'est $f_{a^{-1}, -a^{-1}b}$. De plus l'application $(a, b) \mapsto f_{a,b}$ est injective. On obtient donc un groupe isomorphe à \mathcal{S}_+ en munissant $\mathbb{C}^* \times \mathbb{C}$ de la multiplication

$$(a, b) * (a', b') = (aa', ab' + b).$$

On voit tout de suite que ce groupe n'est pas abélien.

En particulier **pour que $f_{a,b}$ soit un déplacement, il faut et il suffit que $|a| = 1$** . Dans ce cas, pour $a = +1$, $f_{a,b}$ est une **translation**, et pour $a \neq 1$, c'est une **rotation** dont le centre est le point invariant $\frac{1}{1-a}$ et d'angle tout réel $\theta \in \arg(a)$.

De toute façon $f_{a,b}$ est une similitude directe dont le centre est l'unique point fixe $\frac{b}{1-a}$ (le cas $a = 1$ a déjà été vu), dont le rapport est $|a|$ et d'angle tout réel $\theta \in \arg(a)$.

On sait qu'un sous-groupe remarquable de \mathcal{S}_+ est le sous-groupe $\mathcal{HT}(\mathbb{C})$ des **homothéties et translations**. Un élément $f_{a,b}$ de $\mathcal{HT}(\mathbb{C})$ se reconnaît au fait que $a \in \mathbb{R}^*$.

Remarquons enfin que les similitudes directes de centre donné z_0 forment un sous-groupe de \mathcal{S}_+ isomorphe au groupe des similitudes vectorielles $\overline{\mathcal{S}}_+$ (cf. exercice 2).

THÉORÈME VI.10.5

|| Soit $(a, b) \neq (1, 0)$ et $(a', b') \neq (1, 0)$ donnés dans $\mathbb{C}^* \times \mathbb{C}$. Pour que les similitudes directes $f_{a,b}$ et $f_{a',b'}$ soient **conjuguées** dans le groupe \mathcal{S}_+ , il faut et il suffit que $a = a'$.

Démonstration :

Supposons d'abord que $a = a' \neq 1$. Pour pouvoir écrire :

$$f_{a',b'} = \tau \circ f_{a,b} \circ \tau^{-1},$$

il suffit de choisir pour τ la translation $z \mapsto z + \frac{b' - b}{1 - a}$, donc $f_{a,b}$ et $f_{a,b'}$ sont conjuguées dans \mathcal{S}_+ .

Notons que si $a = a' = 1$, alors par hypothèse $b \neq 0$ et $b' \neq 0$ et il est facile d'écrire :

$$f_{1,b'} = f_{a,0} \circ f_{1,b} \circ f_{1/\alpha,0}$$

en prenant $\alpha = \frac{b'}{b}$ et $f_{1,b'}$ et $f_{1,b}$ sont encore conjuguées.

Réciproquement, supposons $f_{a',b'} = g \circ f_{a,b} \circ g^{-1}$ avec $g \in \mathcal{S}_+$, c'est-à-dire $g = g_{\alpha,\beta}$ avec $(\alpha, \beta) \in \mathbb{C}^* \times \mathbb{C}$. En écrivant $f_{a',b'} \circ g = g \circ f_{a,b}$, on obtient $\alpha(a - a') = 0$ d'où $a = a'$. ■

Remarque. Le groupe \mathcal{S}_+ est *doublement transitif* sur \mathbb{C} . En effet soit $z_1 \neq z_2$, $z'_1 \neq z'_2$ quatre complexes. Le système $az_1 + b = z'_1$, $az_2 + b = z'_2$ aux inconnues a et b (dans \mathbb{C}) admet la solution unique

$$a = \frac{z'_1 - z'_2}{z_1 - z_2}, \quad b = \frac{z_1 z'_2 - z_2 z'_1}{z_1 - z_2}$$

et on voit que $a \neq 0$.

Similitudes indirectes de \mathbb{C} , groupe des similitudes

Rappelons qu'une similitude indirecte de \mathbb{C} est le produit (d'ailleurs commutatif) d'une similitude vectorielle indirecte par une translation. Or on a déjà vu qu'une similitude vectorielle indirecte pouvait toujours s'écrire sous la forme $\psi_a = \varphi_a \circ \sigma : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto a\bar{z}$. Il en résulte que toute similitude indirecte de \mathbb{C} pourra s'écrire, et de manière unique, sous la forme $g_{a,b} = f_{a,b} \circ \sigma$, c'est-à-dire de telle sorte que l'image de z par $g_{a,b}$ soit $a\bar{z} + b$ (on a également $g_{a,b}(z) = \overline{a\bar{z} + b}$, c'est-à-dire $g_{a,b} = \sigma \circ f_{\bar{a},\bar{b}}$), et la partie linéaire de $g_{a,b}$ est ψ_a . Plaçons-nous d'abord dans le cas où le rapport de similitude est tel que $|a| \neq 1$.

THÉORÈME VI.10.6

Soit $(a, b) \in \mathbb{C}^* \times \mathbb{C}$ et $g_{a,b}$ la similitude indirecte $z \mapsto a\bar{z} + b$ de rapport $|a| \neq 1$. Alors $g_{a,b}$ admet un point fixe unique $z_0 = \frac{a\bar{b} + b}{1 - |a|^2}$ appelé **centre de similitude**. Soit Δ sa **droite principale**, passant par z_0 et dirigée par $e^{i\theta/2}$ pour $\theta \in \arg(a)$. Alors $g_{a,b}$ est le produit commutatif de l'homothétie de centre z_0 et de rapport $|a|$, et de la symétrie orthogonale par rapport à la droite Δ .

Démonstration :

Pour que $g_{a,b}(z) = z$ il faut et il suffit que $a\bar{z} - z = -b$, d'où par conjugaison $\bar{a}z - \bar{z} = -\bar{b}$. Si $a\bar{a} = |a|^2 \neq 1$ ce système aux deux inconnues complexes z et \bar{z} admet une solution unique et on trouve

$$z_0 = \frac{a\bar{b} + b}{1 - |a|^2}.$$

Utilisant ce point z_0 , on a :

$$\overrightarrow{g_{a,b}(z_0)g_{a,b}(z)} = \overrightarrow{z_0g_{a,b}(z)} = \psi_a(\overrightarrow{z_0z}).$$

Si $\theta \in \arg(a)$, ψ_a est le produit commutatif de l'homothétie de rapport $|a|$ et de la symétrie orthogonale d'axe dirigé par $e^{i\theta/2}$ (cf. théorèmes VI.10.1 et VI.10.3). ■

Étudions maintenant le cas où $|a| = 1$.

THÉORÈME VI.10.7

Soit $(a, b) \in \mathbb{C}^* \times \mathbb{C}$ et $g_{a,b}$ la similitude indirecte $z \mapsto a\bar{z} + b$ de rapport $|a| = 1$. Si $a\bar{b} + b \neq 0$, $g_{a,b}$ n'admet aucun point fixe. C'est le produit commutatif d'une symétrie orthogonale s et d'une translation τ parallèle à l'axe de la symétrie, le couple (s, τ) étant défini de manière unique par ces conditions, et dit associé à $g_{a,b}$.
Si $a\bar{b} + b = 0$, alors $g_{a,b}$ est une symétrie orthogonale.

Démonstration :

Le cas $b = 0$ a déjà été étudié (voir le théorème VI.10.1). Si $b \neq 0$ et $a\bar{b} + b = 0$ il est évident que $a\bar{z} - z = -b$ est vérifiée pour $z_0 = \frac{b}{2}$ qui est donc un point fixe de $g_{a,b}$. ψ_a est alors la symétrie orthogonale d'axe la droite Δ dirigée par $e^{i\theta/2}$ et en prenant $\frac{b}{2}$ comme origine, on voit que $g_{a,b}$ n'est autre que la symétrie orthogonale d'axe, la parallèle à Δ passant par $z_0 = \frac{b}{2}$ et dirigée par $e^{i\theta/2}$.

Si maintenant $a\bar{b} + b \neq 0$, $g_{a,b}$ n'admet aucun point fixe. Posons $a = e^{i\theta}$ et cherchons z_0 tel que $g_{a,b}(z) - z_0$ puisse se mettre sous la forme $e^{i\theta}(\overline{z - z_0}) + \lambda e^{i\theta/2}$ ce qui donne

$$b e^{-i\theta/2} - z_0 e^{-i\theta/2} = -e^{i\theta/2} \bar{z}_0 + \lambda \quad (\lambda \text{ réel}).$$

Visiblement $\lambda = \operatorname{Re}(b e^{-i\theta/2})$, d'où en posant $b e^{-i\theta/2} = \lambda + i\mu$ on trouve facilement $z_0 = \frac{i\mu}{2} e^{i\theta/2}$. Par suite $g_{a,b}$ est le produit (commutatif) de la symétrie orthogonale s par rapport à la droite passant par z_0 et dirigée par $e^{i\theta/2}$ et de la translation de vecteur $\lambda e^{i\theta/2}$ qui est bien parallèle à l'axe de la symétrie. Il reste à prouver l'unicité du couple (s, τ) mais elle provient du fait que, dans $g_{a,b}$, l'axe de s est la seule droite globalement invariante. ■

Remarquons que $g_{a,b}$ est une isométrie de \mathbb{C} ssi $|a| = 1$.

Etudions enfin la conjugaison dans le groupe \mathcal{S} des similitudes.

THÉORÈME VI.10.8

Soit (a, b) et (a', b') dans $\mathbb{C}^* \times \mathbb{C}$. Pour que $g_{a,b}$ et $g_{a',b'}$ soient des similitudes indirectes conjuguées dans le groupe \mathcal{S} , il faut et il suffit : ou bien que $|a| = |a'| \neq 1$, ou bien que l'on ait $|a| = |a'| = 1$ avec $a\bar{b} + b$ et $a'\bar{b}' + b'$ ou bien tous deux nuls ou bien tous deux non nuls.

Démonstration :

Si $g_{a,b}$ et $g_{a',b'}$ sont conjugués dans \mathcal{S} , pour tout $h \in \mathcal{S}$ tel que $g_{a',b'} = h \circ g_{a,b} \circ h^{-1}$, h définit une bijection de l'ensemble des points fixes de $g_{a,b}$ sur celui de $g_{a',b'}$, donc si $|a| = |a'| = 1$, les nombres $a\bar{b} + b$ et $a'\bar{b}' + b'$ sont nuls ou non nuls ensemble. De plus, en passant aux parties linéaires : $\psi_{a'} = L(h) \circ \psi_a \circ L(h^{-1})$ et en utilisant les rapports de ces similitudes vectorielles on voit que $|a| = |a'|$.

Réciproquement supposons d'abord $|a| = |a'| \neq 1$. $g_{a,b}$ a pour centre z_0 et $g_{a',b'}$ a pour centre z'_0 . Choisissons $z_1 \neq z_0$ sur la droite principale de $g_{a,b}$ et $z'_1 \neq z'_0$ sur la droite principale de $g_{a',b'}$. D'après la remarque qui suit l'étude de la conjugaison dans \mathcal{S}_+ , il existe une $h \in \mathcal{S}_+$ telle que $h(z_0) = z'_0$ et $h(z_1) = z'_1$. Alors $g_{a',b'} = h \circ g_{a,b} \circ h^{-1}$. Supposons ensuite $|a| = |a'| = 1$ et $a\bar{b} + b, a'\bar{b}' + b'$ tous deux nuls. On recommence le même raisonnement avec z_0 et z_1 distincts choisis sur l'axe de la symétrie orthogonale $g_{a,b}$ (resp. z'_0 et z'_1 sur l'axe de $g_{a',b'}$).

Supposons enfin $|a| = |a'| = 1$ et $a\bar{b} + b = a'\bar{b}' + b' \neq 0$. Choisissons un point z_0 (resp. z'_0) sur l'unique droite invariante de $g_{a,b}$ (resp. $g_{a',b'}$). Posons $a = e^{i\theta}$, $a' = e^{i\theta'}$, et soit $\lambda e^{i\theta/2}$, $\lambda' e^{i\theta'/2}$ ($\lambda \in \mathbb{R}^*$, $\lambda' \in \mathbb{R}^*$) les vecteurs des translations τ (resp. τ') associées à $g_{a,b}$ (resp. $g_{a',b'}$). Prenons $h \in \mathcal{S}_+$ telle que $h(z_0) = z'_0$ et $L(h)(\lambda e^{i\theta/2}) = \lambda' e^{i\theta'/2}$. Il est clair que $h \circ g_{a,b} \circ h^{-1} = g_{a',b'}$. ■

Exercice 1 : Quels sont les sous-groupes finis du groupe \mathcal{S}_+ ? du groupe \mathcal{S} ?

Exercice 2 : Soit G un sous-groupe du groupe affine de \mathbb{C} (voir la définition dans le tome de Géométrie : un élément du groupe affine est le produit d'une transformation linéaire et d'une translation) et soit \mathcal{T} le groupe des translations de \mathbb{C} .

a) Montrer que si $G \cap \mathcal{T} = \{\operatorname{Id}_{\mathbb{C}}\}$, alors G est abélien.

b) Soit G_{z_0} le stabilisateur de z_0 dans G (c'est-à-dire $G_{z_0} = \{f \in G \mid f(z_0) = z_0\}$). Vérifier que $(\mathcal{S}_+)_z$ est isomorphe à $\overline{\mathcal{P}}_+$, et que \mathcal{S}_{z_0} est isomorphe à $\overline{\mathcal{P}}$, pour tout $z_0 \in \mathbb{C}$.

c) Si G est un sous-groupe abélien du groupe \mathcal{S}_+ , prouver qu'il existe $z_0 \in \mathbb{C}$ tel que $G \subset (\mathcal{S}_+)_z$, ou bien que $G \subset \mathcal{T}$.

d) Chercher les couples d'isométries de \mathbb{C} *permutables*. En déduire les sous-groupes abéliens de $\text{Is}(\mathbb{C})$, puis de \mathcal{S} .

Exercice 3 : Soit G un sous-groupe du groupe \mathcal{S}_+ possédant la propriété suivante :
(1) Toute G -orbite dans \mathbb{C} est un ensemble *discret* (cf. § VI.8, exercice 16). On suppose en outre que $G \cap \mathcal{T} = \{\text{Id}_{\mathbb{C}}\}$.

Démontrer que G est monogène. Préciser dans quel cas G est fini, et donner une description des G -orbites dans \mathbb{C} .

Exercice 4 : Nous admettons ici la propriété suivante de nature topologique : (T) Les sous-groupes *discrets*, non réduits à $\{0\}$, de $(\mathbb{C}, +)$ sont soit des sous-groupes (dits de rang 1) de la forme $a\mathbb{Z}$, où $a \in \mathbb{C}^*$ et $a\mathbb{Z} = \{ka\}_{k \in \mathbb{Z}}$, soit les sous-groupes (dits de rang 2) de la forme $a\mathbb{Z} + b\mathbb{Z} = \{ma + nb\}_{(m,n) \in \mathbb{Z}^2}$ où (a, b) est une base du \mathbb{R} -ev \mathbb{C} .

Le groupe \mathcal{T} des translations de \mathbb{C} sera identifié au groupe additif \mathbb{C} . Soit G un sous-groupe de $\text{Is}_+(\mathbb{C})$ vérifiant la propriété (1) de l'exercice 3, et tel que $G \cap \mathcal{T} = H \neq \{\text{Id}_{\mathbb{C}}\}$.

a) Montrer que H est un sous-groupe discret de $(\mathbb{C}, +)$.

b) On suppose que H est de rang 1. Donner la structure de G .

c) On suppose que H est de rang 2, que $G_0 \neq \{\text{Id}_{\mathbb{C}}, -\text{Id}_{\mathbb{C}}\}$ et que $1 \in H$, $\min_{z \in H} \{|z|\} = 1$.

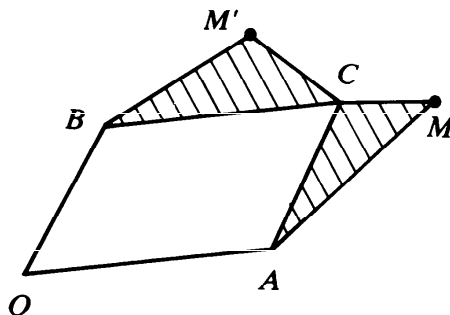
$H = \{m + n\tau\}_{(m,n) \in \mathbb{Z}^2}$ où $\tau \in \mathbb{C} \setminus \mathbb{R}$ et $\text{Re}(\tau) \geq 0$. Démontrer que H se réduit alors soit à l'ensemble des entiers de Gauss soit à l'ensemble $J = \{a + bj\}_{(a,b) \in \mathbb{Z}^2}$ (cf. exercices 1 et 2 du § VI.9), et que G est un sous-groupe du groupe \mathcal{G} des isométries f de \mathbb{C} telles que $f(H) = H$.

Exercice 5 : Étudier la transformation $z \mapsto (2 + i)\bar{z} + 1 - 3i$.

Exercice 6 : Montrer que la transformation $z \mapsto iz + (1 - i)\bar{z}$ est un endomorphisme du \mathbb{R} -ev \mathbb{C} . Noyau. Image. Description géométrique.

Exercice 7 : On compose la translation de vecteur a et la rotation de centre O , d'angle $\pi/4$. Quel est le centre de la rotation produit ?

Exercice 8 : *Appareil à similitude.* Soit un parallélogramme articulé $OACB$. Aux tiges AC et BC sont liés deux triangles tels que AMC et BCM' soient directement semblables par construction.



Montrer que le lieu de M' se déduit du lieu de M par une similitude plane directe de centre O .

Exercice 9 : Soit abc un triangle du plan d'Argand-Cauchy. Sur les côtés on construit 3 triangles abc' , bca' , cab' directement semblables entre eux.

a) Montrer que abc et $a'b'c'$ ont le même centre de gravité.

b) Si les triangles abc' , ... sont équilatéraux, leurs centres forment un tria

§ VI.11 NOMBRES COMPLEXES, DROITES ET CERCLES

Tout point $z = x + iy$ (où $x = \operatorname{Re}(z)$ et $y = \operatorname{Im}(z)$) du plan d'Argand-Cauchy peut être identifié au point de *coordonnées cartésiennes* (x, y) dans le repère canonique $(O, 1, i)$ de \mathbb{C} . Si l'on remarque que $x = \frac{z + \bar{z}}{2}$ et $y = \frac{z - \bar{z}}{2i}$ on comprend qu'à toute relation entre les coordonnées de points M_k correspondra bijectivement une relation entre les affixes des points M_k et les affixes conjugués.

Exemple 1 : Soit deux vecteurs de \mathbb{C} : \vec{v}_1 d'affixe z_1 et \vec{v}_2 d'affixe z_2 . Exprimer en fonction de z_1 et z_2 le produit scalaire $(\vec{v}_1 | \vec{v}_2)$.

On connaît l'expression analytique du produit scalaire :

$$(\vec{v}_1 | \vec{v}_2) = x_1 x_2 + y_1 y_2.$$

On en déduit immédiatement

$$(\vec{v}_1 | \vec{v}_2) = \frac{1}{2} (z_1 \bar{z}_2 + \bar{z}_1 z_2).$$

Droites de \mathbb{C}

Soit Δ l'ensemble des classes d'équivalence des triplets $(u, v, w) \in \mathbb{R}^3$ tels que $(u, v) \neq (0, 0)$ pour la relation d'équivalence (dite de *proportionnalité*) $\exists \lambda \in \mathbb{R}^* \mid (u', v', w') = (\lambda u, \lambda v, \lambda w)$. A la classe γ de représentant (u, v, w) associons la droite $D(\gamma)$, qui ne dépend que de γ , dont une équation cartésienne est $ux + vy + w = 0$. On sait que $\gamma \mapsto D(\gamma)$ est une bijection de Δ sur l'ensemble des droites de \mathbb{C} .

En utilisant $z = x + iy$ on obtient l'équation

$$z \frac{u - iv}{2} + \bar{z} \frac{u + iv}{2} + w = 0,$$

soit en posant $\frac{u + iv}{2} = a$, l'équation

$$(1) \quad \boxed{\bar{a}z + a\bar{z} + w = 0},$$

où bien entendu les coefficients a et w sont définis à une constante multiplicative près de \mathbb{R}^* .

Exemple 2 : Soit α et β dans \mathbb{C} ($\alpha \neq \beta$). Une équation de la droite passant par α et β est

$$(\bar{\alpha} - \bar{\beta})z - (\alpha - \beta)\bar{z} + \alpha\bar{\beta} - \bar{\alpha}\beta = 0$$

qui est bien de type (1) si on divise par i .

Cercles de \mathbb{C}

L'équation du cercle de centre $\omega(a, b)$ et de rayon R est évidemment $(x - a)^2 + (y - b)^2 = R^2$. Réciproquement l'équation

$$(2) \quad x^2 + y^2 - 2\alpha x - 2\beta y + \gamma = 0$$

représente le cercle de centre $\omega(\alpha, \beta)$, de rayon $R = \sqrt{\alpha^2 + \beta^2 - \gamma}$, à condition que $\alpha^2 + \beta^2 - \gamma \geq 0$ (ce cercle se réduit à un point, son centre, lorsque $R = 0$). Si \mathcal{C} désigne l'ensemble

$$\{(\alpha, \beta, \gamma) \in \mathbb{R}^3 \mid \alpha^2 + \beta^2 - \gamma \geq 0\},$$

en associant à chaque élément de \mathcal{C} le cercle d'équation (2) on obtient une bijection de \mathcal{C} sur l'ensemble des cercles du plan d'Argand-Cauchy. En utilisant $z = x + iy$ on obtient l'équation

$$(3) \quad \boxed{z\bar{z} - \bar{c}z - c\bar{z} + \gamma = 0}, \quad \text{avec } c = \alpha + i\beta \quad \text{et} \quad |c|^2 - \gamma \geq 0.$$

Exemple 3 : Soit a et b dans \mathbb{C} , $a \neq b$. Etant donné $\theta \in \mathbb{R}$, trouver l'ensemble E_θ des points $z \in \mathbb{C} \setminus \{a, b\}$ tels que

$$\widehat{(D(a, z), D(b, z))} = \theta \bmod (\pi).$$

Voici une solution possible. Il y a d'abord le cas simple où $\theta = 0 \bmod (\pi)$ car alors a, b, z sont alignés et l'ensemble E_θ est la droite $D(a, b)$ privée de a et b . Supposons maintenant $\theta \notin \pi\mathbb{Z}$. Les droites $D(a, z)$ et $D(b, z)$ admettent, pour $z \in \mathbb{C} \setminus \{a, b\}$, les vecteurs unitaires directeurs $u = \frac{z-a}{|z-a|}$ et $v = \frac{z-b}{|z-b|}$. La relation $\widehat{(D(a, z), D(b, z))} = \theta \bmod \pi$ équivaut à $\frac{v}{u} \in \{e^{i\theta}, -e^{-i\theta}\}$, c'est-à-dire à $v^2 = u^2 e^{2i\theta}$, soit encore $\frac{z-b}{\bar{z}-\bar{b}} = e^{2i\theta} \frac{z-a}{\bar{z}-\bar{a}}$, ou en développant

$$z\bar{z} - \bar{c}z - c\bar{z} + \gamma = 0,$$

à condition de poser

$$c = \frac{i}{2 \sin \theta} (b e^{-i\theta} - a e^{i\theta}) \quad \text{et} \quad \gamma = \frac{i}{2 \sin \theta} (\bar{a}b e^{-i\theta} - a\bar{b} e^{i\theta}).$$

On vérifie bien que γ est réel et on reconnaît l'équation (3) qui est celle d'un cercle si $|c|^2 - \gamma \geq 0$. Or $|c|^2 - \gamma = c\bar{c} - \gamma$ se calcule aisément et l'on trouve $\frac{1}{4 \sin^2 \theta} |a - b|^2$. L'ensemble cherché E_θ est donc le cercle de centre c , de rayon $R = \frac{|a - b|}{2|\sin \theta|}$ privé des points a et b .

En donnant à θ toutes les valeurs sur $]0, \pi[$, on obtient un *faisceau linéaire de cercles* de points de base a et b (c'est-à-dire tous les cercles passant par a et b).

Exemple 4 : On donne a et b dans \mathbb{C} ($a \neq b$) et $\lambda \in \mathbb{R}_+^*$. Trouver l'ensemble C_λ des $z \in \mathbb{C}$ tels que $|z - b| = \lambda |z - a|$.

La condition imposée équivaut à $|z - b|^2 - \lambda^2 |z - a|^2 = 0$, soit en développant

$$(1 - \lambda^2) z\bar{z} + (\bar{a}\lambda^2 - \bar{b})z + (a\lambda^2 - b)\bar{z} + |b|^2 - \lambda^2 |a|^2 = 0.$$

Dans le cas particulier $\lambda = 1$ on reconnaît l'équation d'une droite (évidemment la médiatrice de $\{a, b\}$).

Si $\lambda \neq 1$ on peut diviser par $1 - \lambda^2$ et l'équation obtenue est de la forme (3) avec

$$c = \frac{a\lambda^2 - b}{\lambda^2 - 1} \quad \text{et} \quad \gamma = \frac{|b|^2 - \lambda^2 |a|^2}{1 - \lambda^2},$$

ce qui donne $|c|^2 - \gamma = \frac{\lambda^2}{(\lambda^2 - 1)^2} |a - b|^2 > 0$. L'ensemble C_λ est le cercle de centre c et de rayon $\frac{\lambda}{|\lambda^2 - 1|} |a - b|$.

En donnant à λ toutes les valeurs sur $]0, +\infty[$, on obtient un *faisceau linéaire de cercles* à points limites a et b (en considérant c comme barycentre de $a(\lambda^2)$ et de $b(-1)$, ou analytiquement, on voit que les centres des C_λ décrivent la partie de $D(a, b)$ extérieure au segment $[a, b]$).

Exemple 5 : On donne a et b dans \mathbb{C} ($a \neq b$) et $\theta \in \mathbb{R} - \pi\mathbb{Z}$. Trouver l'ensemble A_θ des $z \in \mathbb{C} \setminus \{a, b\}$ tels que $\widehat{(z - a, z - b)} = \theta \bmod (2\pi)$.

Il est d'abord évident que l'hypothèse faite implique que l'angle des droites $D(a, z)$ et $D(b, z)$ est égal à $\theta \bmod (\pi)$, donc que $A_\theta \subset E_\theta$, E_θ étant le cercle (privé de a et b) trouvé dans l'exemple 3. Mais si nous posons $Z = \frac{z - b}{z - a}$, l'hypothèse de l'énoncé est exactement

$$\text{Arg}(Z) = \theta \bmod (2\pi).$$

Si par exemple $\sin \theta > 0$ on aura $\text{Im}(Z) > 0$, et si $\sin \theta < 0$, $\text{Im}(Z) < 0$. Or $\text{Im}(Z) > 0$ s'écrit $\text{Im}((z - b)(\bar{z} - \bar{a})) > 0$, c'est-à-dire

$$\text{Im}(\bar{a}b - \bar{a}z - b\bar{z}) > 0.$$

Mais l'équation $\text{Im}(\bar{a}b - \bar{a}z - b\bar{z}) = 0$ définit la droite $D(a, b)$. L'ensemble A_θ est donc la partie du cercle E_θ contenue dans l'un des demi-plans déterminés par $D(a, b)$, soit en définitive *un arc de cercle*.

Ces quelques exemples suffisent à montrer que, sans préten

la géométrie pure ou la géométrie analytique, l'utilisation des nombres complexes peut rendre de grands services en géométrie plane euclidienne, spécialement quand il s'agit de droites et de cercles. Nous étudierons en Analyse certaines transformations de \mathbb{C} dans \mathbb{C} . Parmi les plus simples de ces transformations, signalons dès maintenant que les transformations « homographiques »

$$z \mapsto \frac{az + b}{cz + d} \quad ((a, b, c, d) \in \mathbb{C}^4, ad - bc \neq 0)$$

du plan complexe préalablement complété par un « point à l'infini » ont la propriété remarquable de transformer une droite ou un cercle en une figure qui est également une droite ou un cercle. L'ensemble de ces bijections de $\tilde{\mathbb{C}} (= \mathbb{C} \cup \{\infty\})$ encore appelé sphère de Riemann) forme un groupe appelé le *groupe circulaire droit* dont \mathcal{S}_+ est évidemment un sous-groupe (voir tome III, Géométrie).

Exercice 1 : Soit a, b, c trois points non alignés dans le plan d'Argand-Cauchy, et $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. Par $z \in \mathbb{C}$, on mène les trois droites $\Delta_1, \Delta_2, \Delta_3$ telles que

$$(\widehat{\Delta_1, D(b, c)}) = (\widehat{\Delta_2, D(c, a)}) = (\widehat{\Delta_3, D(a, b)}) = \theta \bmod (\pi)$$

qui rencontrent respectivement $D(b, c), D(c, a)$ et $D(a, b)$ en α, β, γ . Trouver l'ensemble des $z \in \mathbb{C}$ tels que α, β, γ soient alignés. (*Réponse :* le cercle circonscrit au triangle a, b, c . Si $\theta = \frac{\pi}{2} \bmod (\pi)$, la droite $D(\alpha, \beta, \gamma)$ est appelée *droite de Simson* du point M relative au triangle a, b, c .)

Exercice 2 : On donne deux cercles de rayons distincts, \mathcal{C} et \mathcal{C}' , dans le plan d'Argand-Cauchy. Trouver toutes les similitudes directes $f \in \mathcal{S}_+$ telles que $f(\mathcal{C}) = \mathcal{C}'$. Trouver l'ensemble des centres de ces similitudes.

Exercice 3 : Ensemble des points M d'affixe z tels que les points d'affixe $1, z, \frac{1}{z}, 1 - z$ soient cocycliques.

Exercice 4 : Etudier la figure formée par les quatre points $a, -a, z$ et $\frac{a^2}{z}$.

a) Ils sont cocycliques.

b) Si $m = \frac{1}{2} \left(z + \frac{a^2}{z} \right)$ montrer que la droite $D(m, z)$ est la bissectrice de l'angle des vecteurs $m - a$ et $m - (-a)$ passant par m .

c) Connaissant le point m , trouver et construire le point z .

Exercice 5 : Ensemble des points M d'affixe z tels que les images de $1, z, 1 + z^2$ soient alignées. Même problème avec $z, \frac{1}{z}$ et z^2 .

Exercice 6 : Ensemble des points $z = \frac{i + t}{1 + i + 2t}$ quand t décrit \mathbb{R} .

Exercice 7 : Déterminer a, b, c, d dans \mathbb{C} pour que l'application $f : z \mapsto \frac{az + b}{cz + d}$ transforme le demi-plan d'équation $ux + vy + w \geq 0$ en un disque donné de centre z_0 et de rayon R .

Exercice 8 : Parmi les homographies $z \mapsto \frac{az + b}{cz + d}$ de $\tilde{\mathbb{C}}$, quelles sont celles qui sont involutives ? Peut-il arriver que $f \circ f \circ f = \text{Id}_{\tilde{\mathbb{C}}}$?

Chapitre VII

POLYNÔMES SUR UN CORPS COMMUTATIF

§ VII.1 POLYNÔMES À UNE INDÉTERMINÉE

Soit K un corps commutatif. Considérons l'ensemble $K^{(\mathbb{N})}$ formé des suites $(a_n)_{n \in \mathbb{N}}$ à support fini d'éléments de K . Cet ensemble, que nous notons provisoirement \mathcal{P}_K , devient un K -espace vectoriel si l'on y définit de façon naturelle une addition et une loi externe de domaine K par :

$$\forall a \in \mathcal{P}_K, \quad \forall b \in \mathcal{P}_K, \quad a = (a_n)_{n \in \mathbb{N}}, \quad b = (b_n)_{n \in \mathbb{N}}, \quad \forall \lambda \in K,$$

$$(1) \quad \boxed{a + b = (a_n + b_n)_{n \in \mathbb{N}}; \quad \lambda a = (\lambda a_n)_{n \in \mathbb{N}}}.$$

Si par ailleurs nous considérons la suite $c = (c_n)_{n \in \mathbb{N}}$ d'éléments de K telle que :

$$(2) \quad \boxed{c_n = \sum_{p+q=n} a_p b_q},$$

cette suite est à support fini, car si $a_n = 0_K$ pour $n \geq d$ et $b_n = 0_K$ pour $n \geq d'$, alors certainement $c_n = 0$ pour $n \geq d + d' - 1$. Donc c est un élément de \mathcal{P}_K , ce qui permet de définir un produit de a par b , noté ab , en posant $ab = c$.

THÉORÈME VII.1.1

|| Muni du produit défini ci-dessus, le K -ev \mathcal{P}_K devient une K -algèbre commutative, dont l'élément unité e est la suite $(\delta_{0n})_{n \in \mathbb{N}}$ telle que $\delta_{00} = 1_K$ et $\delta_{0n} = 0_K$ pour $n \geq 1$. L'homomorphisme canonique $K \rightarrow \mathcal{P}_K$, $\lambda \mapsto \lambda e$ est injectif.

Démonstration :

On vérifie immédiatement que le produit défini par la formule (2) est commutatif et qu'il admet e pour élément neutre. On remarque aussi que :

$$\forall a \in \mathcal{P}_K, \quad \forall b \in \mathcal{P}_K, \quad \forall \lambda \in K, \quad \lambda(ab) = (\lambda a)b = a(\lambda b).$$

On démontre ensuite la distributivité du produit par rapport à l'addition en utilisant les règles de calcul dans un anneau (cf. § III.2) : si a, b, c sont 3 éléments de \mathcal{P}_K de termes généraux a_n, b_n, c_n , en notant $s = (s_n)$ le produit $a(b + c)$ et $s' = (s'_n)$ la somme $ab + ac$, nous avons bien, pour tout n ,

$$s_n = \sum_{p+q=n} a_p(b_q + c_q) = \left(\sum_{p+q=n} a_p b_q \right) + \left(\sum_{p+q=n} a_p c_q \right) = s'_n.$$

On démontre enfin l'associativité de ce produit en utilisant encore les règles de calcul dans un anneau : en notant t_n le terme général de $(ab)c$ et t'_n le terme général de $a(bc)$, nous avons bien, pour tout n ,

$$t_n = \sum_{p+q=n} \left(\sum_{i+j=p} a_i b_j \right) c_q = \sum_{p+q+r=n} a_p b_q c_r = \sum_{p+q=n} a_p \left(\sum_{i+j=q} b_i c_j \right) = t'_n.$$

Quant à l'injectivité de l'application canonique $\lambda \mapsto \lambda e$, elle est immédiate. ■

DÉFINITION VII.1.1

La K -algèbre \mathcal{P}_K construite dans le théorème VII.1.1 s'appelle **K -algèbre des polynômes à une indéterminée à coefficients dans K** . Si $a = (a_n)_{n \in \mathbb{N}}$ est élément de cette algèbre, les a_n sont appelés les **coefficients** de a ; les polynômes de la forme λe ($\lambda \in K$) sont dits **constants**.

Remarque 1 : La K -algèbre des polynômes n'est autre, avec les notations du § VI.4, que la K -algèbre du monoïde $(\mathbb{N}, +)$.

Notons provisoirement P_k le polynôme $(\delta_{kn})_{n \in \mathbb{N}}$ où $\delta_{kk} = 1_K$ et $\delta_{kn} = 0$ si $n \neq k$. Comme nous l'avons vu au § VI.3, la suite $(P_k)_{k \in \mathbb{N}}$ constitue une base (dite canonique) du K -ev \mathcal{P}_K . Les coordonnées dans cette base du polynôme $a = (a_k)_{k \in \mathbb{N}}$ ne sont autres que les a_k : $a = \sum_{k \in \mathbb{N}} a_k P_k$. On

a : $P_0 = e$, élément unité de \mathcal{P}_K . A l'aide de l'injection naturelle $K \rightarrow \mathcal{P}_K, \lambda \mapsto \lambda e$, on convient d'identifier K à la sous-algèbre Ke de \mathcal{P}_K . On peut alors écrire λ au lieu de λe ($\lambda \in K$) et en particulier, $1_K = P_0$ au lieu de e . Considérons maintenant l'élément X : $X = P_1 = (0, 1, 0, 0, \dots)$. Un calcul immédiat montre que

$$(\forall k \in \mathbb{N}) \quad P_k = X^k.$$

Par conséquent, la base canonique du K -ev \mathcal{P}_K n'est autre que la suite des **monômes** en X , c'est-à-dire la famille $(X^k)_{k \in \mathbb{N}}$. D'ailleurs le sous- K -ev engendré par les $(X^k)_{k \in \mathbb{N}}$ n'est autre (cf. § VI.4) que la sous- K -algèbre de \mathcal{P}_K engendrée par X . Donnons à l'élément $X = P_1$ de \mathcal{P}_K le nom d'*indéterminée canonique* et résumons ce qui précède :

THÉORÈME VII.1.2

|| La K -algèbre \mathcal{P}_K des polynômes à une indéterminée à coefficients dans K est **engendrée** par l'indéterminée canonique X ; la suite $(X^k)_{k \in \mathbb{N}}$ forme une base du K -ev \mathcal{P}_K .

Cela conduit à la notation définitive $K[X]$ au lieu de \mathcal{P}_K . Mais cette écriture suppose qu'on a désigné par le symbole X l'indéterminée canonique qu'on aurait pu tout aussi bien désigner par un autre symbole. La phrase : « soit la K -algèbre $K[X]$ (ou $K[T]$, ou $K[U]$, etc...) » sous-entend ce choix. Les éléments de $K[X]$ seront, eux aussi, avantageusement désignés le plus souvent par des lettres majuscules. Si P est un tel polynôme, il s'écrit donc, de manière unique, sous la forme $P = \sum_{k \in \mathbb{N}} a_k X^k$, avec les $(a_k)_{k \in \mathbb{N}}$ à support fini.

Pour pouvoir définir le degré et la valuation d'un polynôme nous avons besoin d'adjoindre à l'ensemble \mathbb{N} deux éléments supplémentaires, notés $-\infty$ et $+\infty$, tout en prolongeant l'ordre usuel de \mathbb{N} à l'ensemble $\overline{\mathbb{N}} = \mathbb{N} \cup \{-\infty, +\infty\}$ ainsi obtenu (c'est-à-dire en convenant que $-\infty < n < +\infty$ pour tout $n \in \mathbb{N}$). On convient en outre que

$$(\forall n \in \mathbb{N}) \quad n + (-\infty) = -\infty, \quad n + (+\infty) = +\infty,$$

et aussi que $(-\infty) + (-\infty) = -\infty$ et que $(+\infty) + (+\infty) = +\infty$ (en revanche on ne définit pas la somme $(+\infty) + (-\infty)$). Cela posé :

DÉFINITION VII.1.2

Soit $P = \sum_{k \in \mathbb{N}} a_k X^k$ élément de $K[X]$. Si $P = 0$ on appelle **degré** de P le symbole $-\infty$ de $\overline{\mathbb{N}}$ et **valuation** de P le symbole $+\infty$ de $\overline{\mathbb{N}}$.
Si $P \neq 0$ on appelle **degré** de P l'entier $\text{Max} \{k \in \mathbb{N} \mid a_k \neq 0\}$ et **valuation** de P l'entier $\text{Min} \{k \in \mathbb{N} \mid a_k \neq 0\}$. Le degré et la valuation de P se notent respectivement $\deg(P)$ et $\text{val}(P)$.

Si P est un polynôme non nul de degré d , le coefficient d'indice d de P s'appelle **coefficient dominant** de P (on notera que le terme dominant du produit de deux polynômes en X non nuls est le produit des termes dominants). Un polynôme P est dit **normalisé** (ou **unitaire**) ssi son coefficient dominant est 1_K .

Fixons $d \in \mathbb{N}$; les polynômes de degré $\leq d$ dans $K[X]$ for

ment le sous- K -ev engendré par les $(X^k)_{0 \leq k \leq d}$; la suite finie $(X^k)_{0 \leq k \leq d}$ est donc une **base** de ce sous- K -ev que l'on note en général $K_d[X]$. Remarquons que la suite de terme général $K_d[X]$ est croissante pour l'inclusion et que $K[X] = \bigcup_d K_d[X]$.

THÉORÈME VII.1.3

Dans $K[X]$, le degré et la valuation vérifient les propriétés suivantes :

$$\forall P \in K[X], \quad \forall Q \in K[X].$$

$$(I) \quad \deg(P + Q) \leq \max(\deg(P), \deg(Q)),$$

et si $\deg(P) \neq \deg(Q)$, il y a égalité

$$\deg(P + Q) = \max(\deg(P), \deg(Q)),$$

et si $\deg(P) = \deg(Q)$, il y a égalité.

$$(II) \quad \deg(PQ) = \deg(P) + \deg(Q),$$

$$\text{val}(PQ) = \text{val}(P) + \text{val}(Q).$$

Démonstration :

Les assertions relatives à l'addition se vérifient très simplement. Bornons-nous à prouver l'assertion (II) pour le degré du produit, celle relative à la valuation de PQ se démontrant de manière analogue.

Soit donc $P \in K[X]$, $Q \in K[X]$. Si P ou $Q = 0$ alors $PQ = 0$, de degré $-\infty$ égal à la somme des degrés. Supposons maintenant P de degré $d \geq 0$ et Q de degré $d' \geq 0$, de coefficients dominants respectifs a_d et $b_{d'}$. Dans le produit PQ le coefficient du terme d'indice $d + d'$ est $a_d b_{d'}$ et les coefficients des termes d'indice $> d + d'$ sont tous nuls d'après la loi de formation du produit. Mais par hypothèse $a_d \neq 0$ et $b_{d'} \neq 0$, donc $a_d b_{d'} \neq 0$, donc $\deg(PQ) = d + d'$. ■

Une conséquence immédiate de cette étude est que $(P \neq 0 \text{ et } Q \neq 0) \Rightarrow PQ \neq 0$, d'où :

THÉORÈME VII.1.4

|| La K -algèbre $K[X]$ est **intègre**.

Éléments inversibles de $K[X]$

Dans la K -algèbre $K[X]$, tout élément $\lambda \in K^*$ est inversible, son inverse étant le polynôme constant λ^{-1} (λ^{-1} désignant l'inverse de λ dans K).

Réciproquement, si $P \in K[X]$ est inversible dans $K[X]$, e

inverse, on a :

$$\deg(PQ) = \deg(P) + \deg(Q) = \deg(1) = 0,$$

d'où nécessairement $\deg(P) = \deg(Q) = 0$, ce qui par définition même du degré impose que les polynômes P et Q sont des constantes non nulles. Il s'ensuit :

THÉOREME VII.1.5

|| Le groupe des éléments inversibles de l'anneau $K[X]$ est l'ensemble des polynômes de degré 0, c'est-à-dire K^* .

COROLLAIRE

|| Deux polynômes $P \in K[X]$ et $Q \in K[X]$ sont **associés** ssi $\exists \lambda \in K^*$ tel que $Q = \lambda P$.

L'association est une relation d'équivalence dans $K[X]$ appelée K^* -proportionnalité. On remarque que dans chaque classe d'équivalence il y a un et un seul polynôme *normalisé*, à l'exception du polynôme nul qui est tout seul dans sa classe.

Nous achèverons cette introduction par l'énoncé du théorème suivant, dont la démonstration est immédiate, mais qui est important pour les applications :

THÉOREME VII.1.6

|| Soit L un sous-anneau du corps commutatif K . L'ensemble des polynômes $P \in K[X]$ dont tous les coefficients appartiennent à L est un sous-anneau de $K[X]$.

Ce sous-anneau se note habituellement $L[X]$.

Exercice 1 : Soit $(E_d)_{d \in \mathbb{N}}$ une suite de polynômes dans $K[X]$ tels que $\deg(E_d) = d$ pour tout d .

- Montrer que, pour tout $n \in \mathbb{N}$, (E_0, E_1, \dots, E_n) est une base du K -ev $K_n[X]$.
- Montrer que $(E_d)_{d \in \mathbb{N}}$ est une base du K -ev $K[X]$.

Exercice 2 : Soit P un polynôme de $K[X]$ ordonné suivant les puissances décroissantes de X :

$$P = a_0 X^n + a_1 X^{n-1} + \dots + a_n.$$

Vérifier que P est le n -ième terme de la suite (A_k) définie par récurrence par

$$A_1 = a_0 X + a_1 \quad \text{et} \quad (\forall k \in \llbracket 1, n-1 \rrbracket) \quad A_{k+1} = A_k X + a_{k+1}.$$

(Cette méthode due à Horner est employée pour faire calculer un polynôme par un ordinateur, tout en économisant le nombre de multiplications.)

Exercice 3 : Soit $f : K[X] \rightarrow \mathbb{N} \cup \{-\infty\}$ vérifiant les propriétés suivantes : $f(0) = -\infty$, $f(1_K) = 0$,

$$(\forall P \in K[X]), (\forall Q \in K[X]), f(P + Q) \leq \max(f(P), f(Q))$$

avec égalité si $f(P) \neq f(Q)$ et $f(PQ) = f(P) + f(Q)$. Prouver que f est l'application $P \mapsto m \deg(P)$ pour un certain $m \in \mathbb{N}$ ou bien l'application $P \mapsto 0$ si $\text{val}(P) \geq 1$, $P \mapsto -\infty$ si $\text{val}(P) = 0$. Enoncer et prouver une propriété analogue avec la valuation.

Exercice 4 : Soit \mathcal{A} la K -algèbre $K^{\mathbb{Z}}$ du groupe \mathbb{Z} (cf. § VI.4).

a) Expliquer pourquoi on peut décrire \mathcal{A} comme l'ensemble des expressions $\sum_{k \in \mathbb{Z}} a_k X^k$, où

$(a_k)_{k \in \mathbb{Z}}$ est une suite d'éléments de K à support fini, et où $X = (\dots, 0, 0, 1, 0, 0, \dots)$, l'indice du terme non nul étant 1.

b) Montrer que \mathcal{A} et $K[X]$ ne sont pas des K -algèbres isomorphes.

c) Montrer que \mathcal{A} et $K[X]$ ne sont pas des anneaux isomorphes.

Exercice 5 : Dans la K -algèbre $K[X]$, on considère la sous- K -algèbre \mathcal{A} engendrée par X^2 et X^3 , notée $\mathcal{A} = K[X^2, X^3]$. Montrer que $K[X]$ et \mathcal{A} ne sont pas des K -algèbres isomorphes.

Exercice 6 : Soit K et L deux corps commutatifs. Si les anneaux $K[X]$ et $L[X]$ sont isomorphes, alors les corps K et L sont isomorphes.

Exercice 7 : Soit K et L deux corps commutatifs et soit $\sigma : K \rightarrow L$ un isomorphisme de K dans L . On note X (resp. Y) l'indéterminée canonique de K (resp. L). Montrer qu'il existe un et un seul homomorphisme d'anneaux $f_\sigma : K[X] \rightarrow L[Y]$ tel que

$$f_\sigma(X) = Y \quad \text{et} \quad (\forall \lambda \in K) \quad f_\sigma(\lambda) = \sigma(\lambda).$$

Vérifier que f_σ est injectif. Appliquer à $K = L = \mathbb{C}$, σ étant la conjugaison des nombres complexes.

Exercice 8 : Si L est un sous-anneau de K , montrer que les éléments inversibles de $L[X]$ sont les éléments inversibles de L .

Exercice 9 : Si trois polynômes P, Q, R à coefficients dans \mathbb{R} vérifient $P^2 - XQ^2 = XR^2$, montrer que l'on a $P = Q = R = 0$.

§ VII.2 L'ANNEAU EUCLIDIEN $K[X]$

Dans tout ce paragraphe le corps commutatif K est fixé.

THÉORÈME VII.2.1

Soit A, B deux polynômes éléments de $K[X]$, avec $B \neq 0$. Il existe un, et un seul, couple (Q, R) d'éléments de $K[X]$ tels que

$$(I) \quad A = BQ + R \quad \text{et} \quad \deg(R) < \deg(B).$$

Démonstration :

Unicité : soit (Q_1, R_1) et (Q_2, R_2) deux couples vérifiant les conditions (I). Alors, par différence, on a

$$B(Q_1 - Q_2) + R_1 - R_2 = 0,$$

soit

$$B(Q_1 - Q_2) = R_2 - R_1.$$

Or le degré du second membre est strictement inférieur au degré de B puisque $\deg(R_2) < \deg(B)$ et $\deg(R_1) < \deg(B)$. Si Q_1

degré du premier membre serait $\geq \deg(B)$, d'où une contradiction. Il s'ensuit nécessairement que $Q_1 - Q_2 = 0$, c'est-à-dire $Q_1 = Q_2$ et par suite $R_1 = R_2$.

Existence : Elle est évidente (avec $R = 0$) si B est un polynôme de degré 0. Soit donc $b_0 X^d$ le terme dominant de B , avec $b_0 \neq 0$ et $d \geq 1$. Nous utiliserons la notation

$$K_m[X] = \{P \in K[X] \mid \deg(P) \leq m\}, \quad (m \in \mathbb{N}).$$

Si $A \in K_{d-1}[X]$ on prend $Q = 0$ et $R = A$, couple qui satisfait (I). Supposons alors l'existence prouvée pour tout $A \in K_{n-1}[X]$ pour $n \in \mathbb{N}$, $n \geq d$, et soit $A \in K_n[X]$ de terme dominant $a_0 X^n$. Alors $A - a_0 b_0^{-1} X^{n-d} B$ est un polynôme A_1 élément de $K_{n-1}[X]$. L'hypothèse de récurrence permet d'affirmer qu'il existe un couple (Q_1, R_1) avec $Q_1 \in K[X]$ et $R_1 \in K_{d-1}[X]$ tel que $A_1 = BQ_1 + R_1$. On en déduit que $A = BQ + R$, avec

$$B = a_0 b_0^{-1} X^{n-d} + Q_1 \quad \text{et} \quad R = R_1 \in K_{d-1}[X].$$

Le théorème d'existence est donc démontré par récurrence portant sur le degré de A . ■

Par définition, Q s'appelle le **quotient** et R le **reste**, dans la **division euclidienne de A par B** . On voit que : le reste est nul ssi A est divisible par B dans l'anneau $K[X]$.

Remarque 1 : Soit L un sous-anneau du corps K contenant les coefficients de A et B , et supposons que $b_0^{-1} \in L$ (c'est-à-dire que b_0 est inversible dans L). Alors la récurrence employée dans la démonstration du théorème VII.2.1 montre que les coefficients du quotient et du reste dans la division euclidienne de A par B appartiennent à L . En d'autres termes, avec les notations du théorème VII.1.6, si $A \in L[X]$, $B \in L[X]$ et $b_0^{-1} \in L$, alors $Q \in L[X]$ et $R \in L[X]$.

Il est clair en particulier que la condition $b_0^{-1} \in L$ est satisfaite si B est **normalisé**. Cette remarque s'appliquera par exemple à $L = \mathbb{Z}$, sous-anneau de $K = \mathbb{Q}$.

Le procédé qui a servi à démontrer l'existence de Q et R pour A et B donnés permet en fait leur détermination pratique. On dispose concrètement les calculs sous forme d'une « division », sans oublier de laisser des vides pour les monômes manquants.

Exemple 1 : On donne $K = \mathbb{C}$,

$$A = X^4 + aX^2 + bX + c, \quad B = X^2 + X + 1.$$

$$\begin{array}{r} X^4 \qquad + aX^2 \qquad + bX \qquad + c \\ - X^3 + (a-1)X^2 + bX \qquad + c \\ \hline aX^2 \qquad + (b+1)X \qquad + c \\ \hline (b-a+1)X + c - a \end{array} \left| \begin{array}{l} X^2 + X + 1 \\ X^2 - X + a \end{array} \right.$$

Les dividendes partiels et le reste s'inscrivent au fur et à mesure qu'on avance dans l'écriture de Q . Ici

$$Q = X^2 - X + a \quad \text{et} \quad R = (b - a + 1)X + c - a.$$

En particulier on peut conclure que : A est divisible par B ssi $b - a + 1 = 0$ et $c = a$.

Idéaux de $K[X]$

THÉORÈME VII.2.2

L'anneau $K[X]$ est principal (cf. définition III.7.4). De manière précise, soit \mathfrak{a} un idéal non nul de $K[X]$: il existe un et un seul polynôme **normalisé** G de degré minimum dans \mathfrak{a} , et on a $\mathfrak{a} = GK[X]$.

Démonstration :

Soit \mathfrak{a} un idéal non nul de $K[X]$ et I l'ensemble des degrés de tous les polynômes de $\mathfrak{a} \setminus \{0\}$; I est une partie non vide de \mathbb{N} ; I admet donc un plus petit élément, soit d . Si $g \in \mathfrak{a}$ est un polynôme de degré d , notons $b_0 X^d$ le terme dominant de g , d'où $b_0 \in K^*$. Alors $G = b_0^{-1} g$ est normalisé, $G \in \mathfrak{a}$ et $\deg(G) = d$. Montrons que $\mathfrak{a} = GK[X]$. Il est d'abord évident que tous les polynômes multiples de G sont dans \mathfrak{a} . Réciproquement, soit F un élément quelconque de \mathfrak{a} . La division euclidienne de F par G donne $F = GQ + R$ avec $\deg(R) < \deg(G)$. Mais $R = F - GQ \in \mathfrak{a}$ d'après la définition d'un idéal. La façon dont l'entier d a été choisi prouve alors que $R = 0$, d'où $F = GQ \in GK[X]$.

Montrons enfin que G est le seul polynôme normalisé de degré d dans \mathfrak{a} . Si G_1 est un autre tel polynôme, on a $G_1 = QG$ avec $Q \in K[X]$ d'après ce qui précède. La comparaison des termes dominants donne $Q = 1$, d'où $G_1 = G$. ■

Le polynôme G mis en évidence dans le théorème VII.2.2 s'appelle le **générateur normalisé** de l'idéal principal non nul \mathfrak{a} . Il est clair que les polynômes $g \in K[X]$ tels que $\mathfrak{a} = gK[X]$ sont les λG , où $\lambda \in K^*$. Ces polynômes s'appellent des **générateurs de l'idéal \mathfrak{a}** .

Anneaux euclidiens

DÉFINITION VII.2.1

Soit A un anneau intègre. On appelle **stathme euclidien** sur A toute fonction $\varphi : A \setminus \{0\} \rightarrow \mathbb{N}$ telle que

(I) Si $x \in A \setminus \{0\}$, $y \in A \setminus \{0\}$ et $x|y$, alors $\varphi(x) \leq \varphi(y)$ ⁽¹⁾.

⁽¹⁾ N.B. La propriété (I) ne sert pas dans la démonstration du théorème VII.2.3. C'est pourquoi certains auteurs appellent *anneau euclidien* un anneau intègre dans lequel existe une fonction φ vérifiant seulement l'axiome (II) qui entraîne l'existence d'une division euclidienne et le théorème VII.2.3. Cependant l'axiome (I) a son utilité car il permet d'obtenir algorithmiquement des éléments *irréductibles* (cf. ci-dessous au § VII.3).

} (II) Pour tous x et y dans A , avec $y \neq 0$, il existe au moins un couple
 } $(q, r) \in A^2$ tel que $x = yq + r$ et $r = 0$ ou $\varphi(r) < \varphi(y)$, c'est-à-dire s'il
 } existe une division euclidienne.
 } L'anneau A est dit **euclidien** ssi il possède au moins un stathme euclidien.

Par exemple, l'anneau \mathbb{Z} est euclidien : un stathme est $x \mapsto |x|$. L'anneau $K[X]$ est euclidien : un stathme est $P \mapsto \deg(P)$.

On notera que l'unicité de la division euclidienne dans A n'est pas exigée dans la définition ci-dessus car elle n'a rien d'indispensable (2 divisions possibles si $A = \mathbb{Z}$). Le raisonnement qui a servi à prouver le théorème VII.2.2 n'ayant utilisé que le fait que $P \mapsto \deg(P)$ est un stathme euclidien sur $K[X]$ prouve donc plus généralement :

THÉORÈME VII.2.3

|| Tout anneau euclidien est principal.

Plus petit commun multiple (ppcm)

Soit $n \in \mathbb{N}^*$ et A_1, A_2, \dots, A_n une suite finie d'éléments de $K[X]$. Les multiples communs à A_1, A_2, \dots, A_n dans l'anneau $K[X]$ sont les éléments de l'idéal

$$\mathfrak{a}_1 \cap \mathfrak{a}_2 \cap \dots \cap \mathfrak{a}_n = \mathfrak{a},$$

où $\mathfrak{a}_i = A_i K[X]$ pour tout i . Il existe donc $G \in K[X]$ tel que $\mathfrak{a} = GK[X]$, et les polynômes G qui vérifient cette condition forment une *classe de K^* -proportionnalité* dans $K[X]$. Ces polynômes G possèdent donc, et sont les seuls polynômes à posséder, la propriété suivante :

$(\mathcal{P}_1) \left\{ \begin{array}{l} \bullet \text{ Pour chaque } i \in \llbracket 1, n \rrbracket, G \text{ est multiple de } A_i. \\ \bullet \text{ Tout } F \in K[X] \text{ qui est multiple de chaque } A_i \text{ est multiple de } G. \end{array} \right.$

DÉFINITION VII.2.2

} Les polynômes G qui possèdent la propriété (\mathcal{P}_1) ci-dessus s'appel-
 } lent les **plus petits communs multiples** des A_i . On les note
 } $\text{ppcm}(A_i)_{1 \leq i \leq n}$.

Les ppcm sont nuls ssi l'un au moins des A_i est nul.

Associons alors à deux polynômes A et B le polynôme nul si $A = 0$ ou $B = 0$ et le ppcm *normalisé* de A et B si $A \neq 0$ et $B \neq 0$. On définit ainsi une loi de composition interne $(A, B) \mapsto A \wedge B$ sur l'anneau $K[X]$, qui est **associative** et **commutative** (comme l'est l'intersection des idéaux engendrés respectivement par A, B et C). Une démonstration tout à fait analogue à celle de la proposition IV.2.1 prouve :

PROPOSITION VII.2.1

|| Si $A_1, \dots, A_n \in K[X]$ alors $\bigwedge_{i=1}^n A_i$ est un ppcm des A_i

On peut ainsi calculer un ppcm des A_i de proche en proche.

Notons aussi que pour $F, A, B \in K[X]$, $F \times (A \wedge B)$ est un ppcm de FA et de FB .

Plus grand commun diviseur (pgcd)

Reprenons $n \in \mathbb{N}^*$ et A_1, A_2, \dots, A_n dans $K[X]$. L'idéal

$$\mathfrak{b} = \mathfrak{a}_1 + \mathfrak{a}_2 + \dots + \mathfrak{a}_n = \sum_{i=1}^n \mathfrak{a}_i$$

engendré (cf. définition III.7.2) par les idéaux $\mathfrak{a}_i = A_i K[X]$ étant principal, ses générateurs D forment une classe de K^* -proportionnalité dans $K[X]$. Ces polynômes D possèdent donc, et sont les seuls polynômes à posséder la propriété suivante :

$$(\mathcal{P}_2) \left\{ \begin{array}{l} \bullet \text{ Pour chaque } i \in \llbracket 1, n \rrbracket, D \text{ divise } A_i \text{ dans l'anneau } K[X]. \\ \bullet \text{ Tout polynôme } F \text{ qui divise chaque } A_i \text{ est diviseur de } D. \end{array} \right.$$

DÉFINITION VII.2.3

$\left\{ \begin{array}{l} \text{Les polynômes } D \text{ qui vérifient la propriété } (\mathcal{P}_2) \text{ ci-dessus s'appellent} \\ \text{les **plus grands communs diviseurs** des } A_i. \text{ On les note} \\ \text{pgcd } (A_i)_{1 \leq i \leq n}. \end{array} \right.$

Les pgcd des A_i sont tous nuls ssi *chaque* A_i est nul.

Associons alors à deux polynômes A et B le polynôme nul si $A = B = 0$, le pgcd *normalisé* de A et B si $(A, B) \neq (0, 0)$. On obtient une loi de composition interne $(A, B) \mapsto A \vee B$ sur l'anneau $K[X]$, qui est **associative** et **commutative** (car la somme des idéaux correspondants l'est).

Si D est un pgcd de A et B , pour tout $F \in K[X]$, DF est un pgcd de FA et de FB . Et par une preuve tout analogue à celle de la proposition IV.2.2, il vient :

PROPOSITION VII.2.2

$$\left\| \begin{array}{l} \text{Si } A_1, \dots, A_n \in K[X], \text{ alors } \bigvee_{i=1}^n A_i \text{ est un pgcd des } A_i. \end{array} \right.$$

Cette proposition permet de ramener de proche en proche tout calcul de pgcd au calcul du pgcd de **deux** polynômes. D'autre part, nous verrons plus loin que les calculs de ppcm se ramènent à ceux de pgcd, d'où l'intérêt d'une technique permettant le calcul du pgcd de deux polynômes :

Algorithme d'Euclide

Nous recherchons le pgcd de $A \in K[X]$ et $B \in K[X]$.

Si $B = 0$, alors $A = \text{pgcd}(A, B)$.

Si $B \neq 0$, $\text{pgcd}(A, B) = \text{pgcd}(B, R)$, où R est le reste de la division euclidienne de A par B . On pose $B = R_0$ dans tous les cas, d'où $A = R_0 Q_1 + R_1$ avec $\deg(R_1) < \deg(R_0)$. Si $R_1 = 0$ alors $A = BQ_1$ et le pgcd de A et de B est B . Si $R_1 \neq 0$ on est conduit à effectuer la division euclidienne de R_0 par R_1 :

$$R_0 = R_1 Q_2 + R_2 \quad \text{avec} \quad \deg(R_2) < \deg(R_1).$$

Supposons ainsi définis les couples (Q_i, R_i) pour $i \in \llbracket 1, k \rrbracket$. Si l'on arrive à un R_k nul, on arrête les divisions successives. Si $R_k \neq 0$ on effectue la division euclidienne de R_{k-1} par R_k : $R_{k-1} = R_k Q_{k+1} + R_{k+1}$ avec $\deg(R_{k+1}) < \deg(R_k)$. On définit ainsi par récurrence deux suites $(Q_i)_{i \in \mathbb{N}}$ et $(R_i)_{i \in \mathbb{N}}$, en convenant que si $R_k = 0$ on pose $R_i = 0$ pour $i \geq k$. La suite infinie $(\deg(R_k))_{k \in \mathbb{N}}$ ne peut rester *strictement décroissante* dans \mathbb{N} . L'ensemble \mathcal{E} des $i \in \mathbb{N}$ tels que $R_i = 0$ est donc non vide. En posant $k_0 = \text{Min}(\mathcal{E})$, on a par construction :

$$(1) \quad A = R_0 Q_1 + R_1, R_0 = R_1 Q_2 + R_2, \dots, R_{k_0-2} = R_{k_0-1} Q_{k_0} \quad \text{avec} \quad R_i \neq 0$$

pour $0 \leq i \leq k_0 - 1$ et

$$\deg(R_0) > \deg(R_1) > \dots > \deg(R_{k_0-1}) \geq 0.$$

La suite (1) est appelée suite des **divisions successives** de A par B .

Si \mathcal{D}_{i+1} est l'ensemble des diviseurs communs à R_i et à R_{i+1} (pour $i \leq k_0 - 2$ avec la convention $R_{-1} = A$), on a $\mathcal{D}_0 = \mathcal{D}_1 = \dots = \mathcal{D}_{k_0-1}$. Mais \mathcal{D}_{k_0-1} n'est autre que l'ensemble $\{\lambda R_{k_0-1}\}_{\lambda \in K^*}$ des polynômes proportionnels à R_{k_0-1} puisque R_{k_0-1} divise R_{k_0-2} . Donc le pgcd de A et B n'est autre que R_{k_0-1} : c'est le **dernier reste non nul** dans la suite des divisions successives.

L'algorithme défini par la suite (1) s'appelle **algorithme d'Euclide pour les polynômes**. Il se prête parfaitement à la programmation sur ordinateur.

Exemple 2 : Soit m et n deux entiers naturels non nuls, $A = X^m - 1$ et $B = X^n - 1$. K est un corps commutatif quelconque. Notons d le pgcd de m et n , d'où $m = \mu d$, $n = \nu d$, avec μ et ν premiers entre eux. Alors $X^d - 1$ est un pgcd de A et B dans $K[X]$.

En effet nous remarquons d'abord que

$$X^m - 1 = (X^d)^\mu - 1 = (X^d - 1)(X^{d(\mu-1)} + \dots + 1),$$

donc $X^d - 1$ divise $X^m - 1$, et il divise aussi $X^n - 1$.

D'autre part soit $D \in K[X]$ un diviseur commun à A et B . On peut choisir deux entiers *naturels* a et b tels que $am - bn = d$. On a alors :

$$X^{am} - 1 = (X^m - 1)F \quad \text{avec} \quad F = X^{m(a-1)} + X^{m(a-2)} + \dots + 1$$

et aussi : $X^{am} - 1 = X^{bn+d} - X^d + X^d - 1 = X^d(X^n - 1)G + X^d - 1$, avec $G = X^{n(b-1)} + X^{n(b-2)} + \dots + 1$. Donc $X^d - 1 = AF - BX^d G$, ce qui montre que tout diviseur commun à A et B divise $X^d - 1$. Donc $X^d - 1 = \text{pgcd}(A, B)$.

DÉFINITION VII.2.4

Des éléments A_1, A_2, \dots, A_n de $K[X]$ sont dits **premiers entre eux** ssi 1_K est un de leurs pgcd, c'est-à-dire ssi les seuls diviseurs communs aux A_i sont les $\lambda \in K^*$.

Il revient au même de dire que l'idéal \mathfrak{b} engendré par les A_i dans $K[X]$ est $K[X]$ lui-même.

THÉORÈME VII.2.4 (théorème de Bezout)

Pour que les polynômes A_1, A_2, \dots, A_n de $K[X]$ soient premiers entre eux, il faut et il suffit qu'il existe dans $K[X]$ des polynômes $\Lambda_1, \Lambda_2, \dots, \Lambda_n$ tels que $\sum_{i=1}^n \Lambda_i A_i = 1_K$.

Démonstration :

Si de tels Λ_i existent, il est clair que tout diviseur commun aux A_i divise 1_K , donc appartient à K^* . Réciproquement, si les A_i sont premiers entre eux, 1_K appartient à l'idéal qu'ils engendrent dans $K[X]$, d'où l'existence de tels Λ_i . ■

Les relations du type $\sum_{i=1}^n \Lambda_i A_i = 1$, avec $\Lambda_i \in K[X]$ s'appellent **relations de Bezout** entre les A_i . S'il en existe, il y en a une infinité.

Exemple 3 : Soit m et n deux entiers naturels premiers entre eux. Posons

$$U = X^{m-1} + X^{m-2} + \dots + 1 \quad \text{et} \quad V = X^{n-1} + X^{n-2} + \dots + 1.$$

Alors, d'après l'exemple 2, $(X-1)U$ et $(X-1)V$ ont pour pgcd : $X-1$, donc U et V sont premiers entre eux. Pour obtenir une relation de Bezout entre U et V on peut procéder de la manière suivante : soit a et b dans \mathbb{N} tels que $am - bn = 1$, alors $LU + MV = 1_K$, avec

$$L = X^{m(a-1)} + X^{m(a-2)} + \dots + 1$$

et
$$M = -X(X^{n(b-1)} + X^{n(b-2)} + \dots + 1).$$

Exemple 4 : Supposons A et B premiers entre eux dans $K[X]$. On obtient une relation de Bezout entre A et B en procédant à des reports successifs dans la suite des divisions successives de A et B , en commençant par le dernier reste non nul (élément de K^*) et en remontant. Ainsi, pour $A = X^4 + X^3 + 1$ et $B = X^2 + X + 1$, éléments de $\mathbb{C}[X]$, il y a deux divisions successives : $A = (X^2 - 1)B + X + 2$, puis

$$B = X^2 + X + 1 = (X + 2)(X - 1) + 3.$$

On part de

$$\begin{aligned} R_2 = 3 = B - R_1 Q_2 &= B - (A - BQ_1) Q_2 = \\ &= -Q_2 A + B(1 + Q_1 Q_2) = (1 - X) A + (X^3 - X^2 - X + 1) B. \end{aligned}$$

On remarque que le couple (U, V) de polynômes tels que $UA + VB = 1$ obtenu par ce procédé vérifie : $\deg(U) < \deg(B)$ et $\deg(V) < \deg(A)$, ce que l'on peut prouver par des récurrences faciles à partir de la suite des divisions successives.

Par des démonstrations analogues à celles du chapitre IV (arithmétique dans \mathbb{Z}) on obtient facilement les résultats suivants :

THÉORÈME VII.2.5 (théorème de Gauss)

|| Soit A, B, C dans $K[X]$. Si A divise BC et si A est premier avec B , alors A divise C .

THÉORÈME VII.2.6

|| Soit A, B_1, B_2, \dots, B_n dans $K[X]$ avec $n \geq 2$. Si A est premier avec chaque B_i ($1 \leq i \leq n$), alors A est premier avec $B = B_1 B_2 \dots B_n$.

THÉORÈME VII.2.7

|| Si les éléments A_i de $K[X]$ ($1 \leq i \leq n$ | $n \geq 2$) sont **deux à deux** premiers entre eux, alors $\text{ppcm}(A_1, A_2, \dots, A_m) = A_1 A_2 \dots A_m$.

COROLLAIRE

|| Si des A_i de $K[X]$ ($1 \leq i \leq n, n \geq 2$) sont **deux à deux** premiers entre eux et si chaque A_i divise le polynôme B de $K[X]$, alors $A_1 A_2 \dots A_n$ divise B .

THÉORÈME VII.2.8

|| Soit $D \in K[X] \setminus \{0\}$ un diviseur commun aux polynômes A_1, A_2, \dots, A_n de $K[X]$. Pour que D soit pgcd de (A_1, A_2, \dots, A_n) , il faut et il suffit que les polynômes $\left(\frac{A_i}{D}\right)_{1 \leq i \leq n}$ soient premiers entre eux.

En combinant les théorèmes VII.2.7 et VII.2.8 on s'aperçoit que, si A et B sont dans $K[X]$, si D est un pgcd de A et B , si M est un ppcm de A et B , alors DM et AB sont K^* -proportionnels, ce qui ramène le calcul du ppcm à celui du pgcd, comme annoncé avant l'algorithme d'Euclide.

On peut compléter le théorème de Bezout par le résultat suivant :

THÉORÈME VII.2.9

|| Soit A et B deux polynômes premiers entre eux

|| $A \neq 0, B \neq 0$. Il existe un et un seul couple (U, V) de polynômes
 || tels que $UA + VB = 1$, $\deg(U) < \deg(B)$, $\deg(V) < \deg(A)$.

Démonstration :

Dans l'exemple 4 ci-dessus nous avons donné le principe de la démonstration d'existence d'un tel couple (U, V) . Supposons alors qu'on ait trouvé deux tels couples (U_1, V_1) et (U_2, V_2) . Cela donne par différence $(U_1 - U_2)A = (V_2 - V_1)B$. Comme A est premier avec B , il doit diviser $V_2 - V_1$ (théorème de Gauss). Mais par hypothèse $\deg(V_2 - V_1) < \deg(A)$. Donc $V_2 - V_1 = 0$ et $U_2 - U_1 = 0$, d'où l'unicité. ■

Arithmétique dans un anneau principal

Soit A un *anneau principal*, c'est-à-dire un anneau intègre où tout idéal est principal. Une fois définies la relation de divisibilité, et les notions d'éléments inversibles et d'éléments associés, il est très facile d'étendre les notions de ppcm et de pgcd : pour toute suite finie (a_1, a_2, \dots, a_n) d'éléments de A , les ppcm des a_i sont les générateurs de l'idéal $\bigcap_{i=1}^n (a_i A)$ et les pgcd des a_i sont les générateurs de

l'idéal $\sum_{i=1}^n (a_i A)$. Les ppcm (resp. les pgcd) forment une classe d'éléments associés dans l'anneau A . Ils vérifient respectivement la propriété (\mathcal{P}_1) qui précède la définition VII.2.2 et la propriété (\mathcal{P}_2) précédant la définition VII.2.3.

De même la notion d'éléments *premiers entre eux* s'étend à A : les a_i sont premiers entre eux si l'ensemble de leurs pgcd est le groupe des éléments inversibles de A . Cela dit les propositions VII.2.1 et VII.2.2 s'étendent sans difficulté à l'anneau A , ainsi que le *théorème de Bezout* et toutes ses conséquences.

En revanche l'existence d'une division euclidienne (et l'algorithme des divisions successives pour la recherche d'un pgcd) n'est pas du tout garantie dans un anneau principal arbitraire, alors qu'elle semble naturelle dans \mathbb{Z} et dans $K[X]$ (dans ce dernier cas il y a même unité du quotient et du reste, ce qui est exceptionnel).

Exercice 1 : Soit $P \in K[X]$ et b, c deux naturels premiers entre eux. Démontrer que $(P^b - 1)(P^c - 1)$ divise $(P - 1)(P^{bc} - 1)$.

Exercice 2 : On donne $K = \mathbb{Z}/5\mathbb{Z}$. Effectuer la division euclidienne de $A = \overline{2}X^3 + \overline{3}X^2 + \overline{1}$ par $B = \overline{1}X^2 + \overline{2}X + \overline{3}$.

Exercice 3 : Soit $K = \mathbb{C}$. Effectuer la division de $4X^3 + X^2$ par $X + 1 + i$.

Exercice 4 : Soit b, c, d, α, β des éléments d'un corps K de caractéristique différente de 2, avec $b \neq 0, c \neq 0$. On considère les polynômes A et B donnés par $A = X^4 + bX^2 + cX + d$, $B = X^2 + \alpha X + \beta$. Calculer le quotient et le reste de la division euclidienne de A par B . A étant fixé, démontrer que les couples (α, β) tels que B divise A sont ceux qui vérifient les conditions :

$$\alpha \neq 0, \quad \beta = \frac{1}{2} \left(\alpha^2 + b - \frac{c}{\alpha} \right) \quad \text{et} \quad \alpha^6 + 2b\alpha^4 + (b^2 - 4d)\alpha^2 - c^2 = 0.$$

Exercice 5 : Soit L un anneau du corps K . On suppose que l'anneau $L[X]$ est principal. Démontrer que L est un sous-corps de K .

Exercice 6 : Soit a et b deux éléments distincts du corps K . Montrer que les polynômes $A = (1 - aX)^m$ et $B = (1 - bX)^n$ où $m \in \mathbb{N}^*$, $n \in \mathbb{N}^*$ sont premiers entre eux. Trouver U et V dans $K[X]$ tels que $UA + VB = 1$, $\deg(U) < n$, $\deg(V) < m$.

Exercice 7 : Soit m_1, m_2, \dots, m_k des naturels ≥ 1 . Montrer que le pgcd normalisé des polynômes $(X^{m_i} - 1)_{1 \leq i \leq k}$ est $X^d - 1$, où $d = \text{pgcd}(m_1, m_2, \dots, m_k)$.

Exercice 8 : On prend $K = \mathbb{C}$. Pour $\theta \in \mathbb{R}$ et $m \in \mathbb{N}^*$ on considère le polynôme

$$F_{m, \theta} = X^{2m} - 2X^m \cos m\theta + 1 \in \mathbb{C}[X].$$

a) Vérifier que $F_{m, \theta}$ est divisible par $F_{1, \theta} = X^2 - 2X \cos \theta + 1$. Calculer le quotient.

b) Calculer le pgcd de $F_{m, \theta}$ et $F_{n, \theta}$ pour $m \in \mathbb{N}^*$, $n \in \mathbb{N}^*$.

Exercice 9 : Soit $\theta \in \mathbb{R}$ et $n \in \mathbb{N}^*$. Vérifier que le polynôme

$$X^n \sin \theta - X \sin n\theta + \sin(n-1)\theta$$

de $\mathbb{C}[X]$ est divisible par $X^2 - 2X \cos \theta + 1$ et calculer le quotient.

Exercice 10 : Soit r_1, r_2, \dots, r_k des naturels tels que $0 \leq r_1 < r_2 < \dots < r_k \leq n-1$ où n est un entier ≥ 2 . On donne dans \mathbb{N} des nombres m_1, m_2, \dots, m_k tels que $m_i \equiv r_i \pmod{n}$ pour $1 \leq i \leq k$. Montrer que $X^{m_1} + X^{m_2} + \dots + X^{m_k}$ est divisible par $X^{r_1} + X^{r_2} + \dots + X^{r_k}$ dans $K[X]$ et calculer le quotient.

Exercice 11 : Soit q et m dans \mathbb{N}^* . Trouver une condition nécessaire et suffisante pour que $1 + X^m + X^{2m} + \dots + X^{qm}$ soit divisible par $1 + X + X^2 + \dots + X^q$ dans $K[X]$.

Exercice 12 : Trouver un polynôme de degré aussi petit que possible dans $\mathbb{R}[X]$ dont le reste de la division par $X^4 - 2X^3 - 2X^2 + 10X - 7$ soit égal à $X^2 + X + 1$ et dont le reste de la division par $X^4 - 2X^3 - 3X^2 + 13X - 10$ soit égal à $2X^2 - 3$.

Exercice 13 : Soit n un entier ≥ 2 et K un corps dont la caractéristique ne divise pas n . Trouver des polynômes U et V dans $K[X]$ tels que

$$U \times (1 + X + X^n) + V \times (1 + nX^{n-1}) = 1.$$

Exercice 14 : Dans $\mathbb{Q}[X]$ calculer le quotient de $nX^{n+1} - (n+1)X^n + 1$ par $(X-1)^2$, avec $n \in \mathbb{N}^*$.

Exercice 15 : Soit K un corps de caractéristique p , et m et n deux entiers naturels. Calculer selon les valeurs de p le pgcd et le ppcm de $A = X(X-1)\dots(X-m)$ et de $B = X(X-1)\dots(X-n)$.

Exercice 16 : Dans $\mathbb{Q}[X]$ on donne

$$A = X^5 + 5X^4 + 9X^3 + 7X^2 + 5X + 3 \quad \text{et} \quad B = X^4 + 2X^3 + 2X^2 + X + 1.$$

Chercher le pgcd D de A et B et l'écrire sous la forme $AU + BV$.

Exercice 17 : Si A et B sont premiers entre eux dans $K[X]$ montrer qu'il en est de même de $A+B$ et de AB .

Exercice 18 : Soit A et B dans $K[X] \setminus \{0\}$. Montrer l'équivalence des deux conditions suivantes :

- A et B ne sont pas premiers entre eux.
- Il existe des polynômes U, V non constants tels que $\deg(U) < \deg(B)$, $\deg(V) < \deg(A)$ et $UA + VB = 0$.

Exercice 19 : Soit m un entier rationnel, non nul, qui n'est pas le carré d'un entier. On pose $\mu = \sqrt{m}$ si $m \geq 1$ et $\mu = i\sqrt{|m|}$ si $m \leq -1$ et l'on désigne par $\mathbb{Q}[\sqrt{m}]$ l'ensemble des nombres complexes de la forme $a + b\mu$, avec $a \in \mathbb{Q}$, $b \in \mathbb{Q}$.

a) Montrer que $\mathbb{Q}[\sqrt{m}]$ est un sous-corps de \mathbb{C} et qu'on ne nuit pas à supposer que m ne contient aucun facteur premier au carré.

b) On dit que x est un entier algébrique quadratique s'il existe des éléments u, v dans \mathbb{Z} tels que $x^2 + ux + v = 0$. Montrer que les éléments de $\mathbb{Q}[\sqrt{m}]$ qui sont « entiers » forment un sous-anneau A de \mathbb{Q} , et que A est égal à $\{a + b\mu\}_{(a,b) \in \mathbb{Z}^2}$ si $m \equiv 2$ ou $m \equiv 3 \pmod{4}$ et à $\left\{c + d \frac{1+\mu}{2}\right\}_{(c,d) \in \mathbb{Z}^2}$ si $m \equiv 1 \pmod{4}$.

c) On appelle « norme » d'un élément $\alpha + \mu\beta$ de A le nombre $\alpha^2 - m\beta^2$ dont on vérifiera qu'il est toujours dans \mathbb{Z} . Utiliser cette norme pour trouver tous les éléments inversibles de l'anneau A dans le cas où $m < 0$ (attention aux cas particuliers -1 et -3). Montrer que la restriction de cette norme à $A \setminus \{0\}$ permet de définir sur A une division euclidienne pour les valeurs $m = -1$, $m = -2$, $m = -3$, $m = -7$, $m = -11$ et que la condition (I) du théorème VII.3.4 est satisfaite par cette division.

d) Montrer qu'en revanche si $m = -5$, l'anneau A n'est pas principal.

e) Définir de même une division euclidienne dans l'anneau A pour $m = 2$, $m = 3$, $m = 5$, $m = 13$.

f) Pour $m = 2$ (resp. $m = 3$) étudier le groupe des éléments inversibles de l'anneau A (on commencera par chercher celui de ces éléments inversibles $a + b\sqrt{m}$ qui est tel que $|a|$ et $|b|$ soient minimaux dans \mathbb{N}^*).

Exercice 20 : Soit $\mathbb{Z}[X]$ l'anneau des polynômes à coefficients entiers rationnels (cf. théorème VII.1.6). Montrer que cet anneau n'est pas principal.

Indication : Considérer l'idéal formé par les polynômes de $\mathbb{Z}[X]$ dont le terme constant est un entier pair.

§ VII.3 L'ANNEAU FACTORIEL $K[X]$

Dans ce qui suit, K désigne toujours un corps commutatif.

D'après la définition IV.4.2, un polynôme $P \in K[X]$ est dit **irréductible** ssi il est *non constant*, et ses seuls diviseurs dans $K[X]$ sont les $\lambda \in K^*$, et les λP ($\lambda \in K^*$).

Les propriétés suivantes sont immédiates :

- Si P est irréductible, λP l'est également pour $\lambda \in K^*$.
- Soit $F \in K[X] \setminus \{0\}$ et $P \in K[X]$, P irréductible ; alors ou bien P divise F (ce que nous notons $P|F$), ou bien P et F sont premiers entre eux.
- Deux polynômes irréductibles sont soit associés, soit premiers entre eux.
- Deux polynômes *irréductibles* et *normalisés distincts* sont premiers entre eux.

Il est commode d'introduire l'ensemble \mathcal{J} des polynômes **irréductibles** et **normalisés** dans $K[X]$, ce qui permet d'énoncer :

PROPOSITION VII.3.1

|| Tout polynôme irréductible dans $K[X]$ est associé à un et un seul polynôme P appartenant à \mathcal{J} .

Exemple 1 : Tout polynôme P de degré 1 est irréductible

$P = AB$ ($A \in K[X], B \in K[X]$), on a :

$$1 = \deg(P) = \deg(A) + \deg(B),$$

donc l'un des polynômes A ou B est constant et l'autre est associé à P .

Exemple 2 : Si $K = \mathbb{Q}$, le polynôme $X^2 - 2$ est irréductible dans $K[X]$; mais le « même » polynôme n'est plus irréductible dans $\mathbb{R}[X]$ car $X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$. De même, le polynôme $X^2 + 1$, irréductible dans $\mathbb{R}[X]$ ne l'est plus dans $\mathbb{C}[X]$.

Soit $P \in \mathcal{J}$ et $F \in K[X] \setminus \{0\}$. L'ensemble des $m \in \mathbb{N}$ tels que $P^m \mid F$ est non vide car $P^0 \mid F$, et il est majoré par $\frac{\deg(F)}{\deg(P)}$ car $m \deg(P) \leq \deg(F)$ et $\deg(P) \geq 1$ puisque P est non constant. Donc cet ensemble admet un plus grand élément.

DÉFINITION VII.3.1

Soit $F \in K[X] \setminus \{0\}$ et $P \in \mathcal{J}$. On appelle **P-valuation** de F , et on note $\text{val}_P(F)$, le plus grand entier $m \in \mathbb{N}$ tel que $P^m \mid F$. On convient que $\text{val}_P(0) = +\infty$.

On retrouve bien pour $P = X$ (l'indéterminée) la valuation ordinaire déjà définie. Ecrire d'un naturel m que $m = \text{val}_P(F)$ signifie qu'il existe $F_1 \in K[X]$ tel que $F = P^m F_1$ et que P ne divise pas F_1 (donc que P est premier avec F_1).

PROPOSITION VII.3.2

Soit $P \in \mathcal{J}$. Pour tous F et G dans $K[X]$, on a :

(I) $\text{val}_P(F + G) \geq \min(\text{val}_P(F), \text{val}_P(G))$

et si $\text{val}_P(F) \neq \text{val}_P(G)$ il y a égalité

(II) $\text{val}_P(FG) = \text{val}_P(F) + \text{val}_P(G)$.

Démonstration :

Si $F = 0$, $\text{val}_P(F) = +\infty$ et les résultats (I) et (II) sont évidents. Supposons donc $F \neq 0$, $G \neq 0$ et posons $\alpha = \text{val}_P(F)$ et $\beta = \text{val}_P(G)$, d'où $F = P^\alpha F_1$, $G = P^\beta G_1$ avec $P \nmid F_1$ et $P \nmid G_1$. Pour fixer les idées prenons $\alpha \leq \beta$. On a : $FG = P^{\alpha+\beta} F_1 G_1$; mais P est premier avec F_1 et avec G_1 , donc avec leur produit (théorème VII.2.6). Donc $P \nmid F_1 G_1$, d'où $\text{val}_P(FG) = \alpha + \beta$, c'est-à-dire (II). Quant à $F + G$, il vaut $P^\alpha(F_1 + P^{\beta-\alpha} G_1)$, d'où $\text{val}_P(F + G) \geq \alpha$, c'est-à-dire l'inégalité (I). Il reste à vérifier que pour $\alpha < \beta$ il y a égalité dans (I). Cela résulte du fait que $P \mid P^{\beta-\alpha} G_1$ puisque $\beta - \alpha \geq 1$, mais $P \nmid F_1$, donc $P \nmid F_1 + P^{\beta-\alpha} G_1$. ■

Si $F \in K[X] \setminus \{0\}$, nous appellerons **support premier** de F l'ensemble des $P \in \mathcal{J}$ tels que $\text{val}_P(F) \geq 1$.

PROPOSITION VII.3.3

|| Le support premier d'un polynôme $F \in K[X] \setminus \{0\}$ est un ensemble fini. Il est non vide ssi F est non constant.

Démonstration :

Soit $P \in \mathcal{J}$, $Q \in \mathcal{J}$ et deux entiers naturels m et n . Par application réitérée du théorème VII.2.6, on montre que P^m et Q^n sont premiers entre eux si $P \neq Q$. Supposons le support premier \mathcal{S}_F de F non vide. Posons $\nu_P = \text{val}_P(F)$ pour $P \in \mathcal{J}$. Par le corollaire du théorème VII.2.7, pour toute partie finie J de \mathcal{S}_F , le polynôme $G_J = \prod_{P \in J} P^{\nu_P}$ divise F , d'où

$$\deg(F) \geq \deg(G_J) \geq \sum_{P \in J} \nu_P \deg(P) \geq \text{card}(J) \quad \text{car} \quad \nu_P \geq 1$$

et $\deg(P) \geq 1$ si $P \in \mathcal{S}_F$. Donc \mathcal{S}_F est nécessairement fini, de cardinal $\leq \deg(F)$. Si $\mathcal{S}_F \neq \emptyset$ il en résulte que $\deg(F) \geq 1$, c'est-à-dire que F est non constant.

Réciproquement, si F est non constant, il admet des diviseurs de degré ≥ 1 , ne serait-ce que F lui-même. Soit alors Q un diviseur non constant de F de degré minimum. Il est clair qu'un tel diviseur Q est irréductible, donc associé à un polynôme $P \in \mathcal{J}$ et par conséquent $P \in \mathcal{S}_F$ qui est non vide. ■

Au cours de la démonstration nous avons vu que F est divisible par $\prod_{P \in \mathcal{S}_F} P^{\text{val}_P(F)}$. Conformément aux conventions du chapitre III, nous pouvons donner un sens à l'expression $\prod_{P \in \mathcal{J}} P^{\text{val}_P(F)}$: c'est 1 si $\mathcal{S}_F = \emptyset$, sinon c'est $\prod_{P \in \mathcal{S}_F} P^{\text{val}_P(F)}$.

THÉORÈME VII.3.1

|| Soit $F \in K[X] \setminus \{0\}$.

(I) On a : $F = u_F \prod_{P \in \mathcal{J}} P^{\text{val}_P(F)}$, où u_F est le coefficient dominant de

F : c'est l'existence de la décomposition de F en facteurs irréductibles.

(II) Soit $u \in K^*$ et $(e_P)_{P \in \mathcal{J}}$ une famille à support fini de naturels tels que $F = u \sum_{P \in \mathcal{J}} P^{e_P}$. Alors $u = u_F$ et $e_P = \text{val}_P(F)$ pour tout

$P \in \mathcal{J}$: c'est l'unicité de la décomposition de F en facteurs irréductibles.

Démonstration :

Soit G le polynôme $\prod_{P \in \mathcal{J}} P^{\nu_P}$, où ν_P désigne $\text{val}_P(F)$.

tout $P \in \mathcal{J}$. On sait que G divise F . D'autre part le polynôme quotient H ne peut admettre aucun diviseur $P \in \mathcal{J}$, donc il est constant, par la proposition VII.3.3. Le coefficient dominant de G étant 1, on a donc $H = u_F$, d'où (I).

Prouvons (II). Le coefficient dominant de $\Phi = \prod_{P \in \mathcal{J}} P^{e_P}$ est 1, donc déjà $u = u_F$. Pour chaque $P \in \mathcal{J}$ on a :

$$\text{val}_P(F) = \text{val}_P(u) + \text{val}_P(\Phi) = \text{val}_P(\Phi).$$

Si $Q \in \mathcal{J}$ et $Q \neq P$, on a aussi $\text{val}_P(Q) = 0$ car $P \nmid Q$; d'où

$$\text{val}_P(\Phi) = \left(\sum_{Q \in \mathcal{J}, Q \neq P} e_Q \text{val}_P(Q) \right)' + e_P = e_P.$$

Finalement, $\text{val}_P(F) = e_P$ pour tout $P \in \mathcal{J}$. ■

La décomposition de $F \in K[X]$ donnée par la formule (I) du théorème VII.3.1 s'appelle *décomposition de F en facteurs irréductibles dans $K[X]$* (ou : **sur le corps K**). Bien entendu le corps K doit toujours être précisé. Par exemple $X^4 + 1$, irréductible sur le corps \mathbb{Q} se décompose en 2 facteurs sur le corps \mathbb{R} et en 4 facteurs sur le corps \mathbb{C} . Par définition les **facteurs irréductibles** de F sont les $P \in \mathcal{J}$ tels que $\text{val}_P(F) \geq 1$: F en possède ssi F est non constant. Dans la pratique on utilise aussi des décompositions où les facteurs ne sont pas toujours normalisés, tout en étant irréductibles et premiers entre eux deux à deux. Il n'y a plus alors de véritable unicité mais seulement « unicité aux facteurs inversibles près ». En tout cas, les exposants des facteurs irréductibles, eux, ne changent pas : si $P \in \mathcal{J}$, la P -valuation val_P est une fonction qui ne dépend en réalité que de l'idéal \mathfrak{p} engendré par P dans $K[X]$.

La proposition VII.3.2 et le théorème VII.3.1 entraînent immédiatement :

THÉORÈME VII.3.2

$$\left\{ \begin{array}{l} \text{(I) Soit } F, G \text{ dans } K[X] \setminus \{0\}. \text{ Pour que } F \mid G, \text{ il faut et il suffit que} \\ \qquad \qquad \qquad \forall P \in \mathcal{J} \quad \text{val}_P(F) \leq \text{val}_P(G). \\ \text{(II) Soit } F_1, F_2, \dots, F_n \text{ dans } K[X] \setminus \{0\}. \text{ On a :} \\ \qquad \text{pgcd}(F_1, F_2, \dots, F_n) = \prod_{P \in \mathcal{J}} P^{\alpha_P}, \text{ avec } \alpha_P = \min_{i=1}^n (\text{val}_P(F_i)) \\ \text{pour tout } P. \\ \qquad \text{ppcm}(F_1, F_2, \dots, F_n) = \prod_{P \in \mathcal{J}} P^{\beta_P}, \text{ avec } \beta_P = \max_{i=1}^n (\text{val}_P(F_i)) \\ \text{pour tout } P. \end{array} \right.$$

Enfin, on a :

THÉOREME VII.3.3

|| L'ensemble \mathcal{J} des polynômes irréductibles normalisés est **infini**.

Démonstration :

On sait déjà que \mathcal{J} est non vide car il contient tous les polynômes $X - a$ où $a \in K$. Soit J une partie finie non vide de \mathcal{J} ; notons F le polynôme $1 + \prod_{P \in J} P$. Alors $\deg(F) = \sum_{P \in J} \deg(P) \geq 1$. Donc F admet un diviseur $Q \in \mathcal{J}$. Or il est clair que $Q \notin J$ car le reste dans la division de F par tout $P \in J$ est 1. Donc $J \neq \mathcal{J}$, ce qui prouve que \mathcal{J} est infini. ■

Cependant les *degrés* des $P \in \mathcal{J}$ ne sont pas arbitraires. Par exemple pour $K = \mathbb{C}$, \mathcal{J} est l'ensemble des $(X - a)_{a \in \mathbb{C}}$. Pour $K = \mathbb{R}$ on ne trouve dans \mathcal{J} que des polynômes de degrés 1 et 2 (cf. § VII.6). Pour $K = \mathbb{Q}$ il y a des polynômes irréductibles de tout degré.

Anneaux factoriels

Un anneau intègre dans lequel tout élément non nul admet une décomposition en facteurs irréductibles, cette décomposition étant unique à des facteurs inversibles près, s'appelle un **anneau factoriel**. C'est le cas de l'anneau \mathbb{Z} , c'est aussi le cas de l'anneau $K[X]$ comme le montre le théorème VII.3.1 ; c'est plus généralement le cas de tout anneau principal. L'intérêt de tels anneaux est que, même en l'absence de division euclidienne et malgré l'existence éventuelle d'idéaux non principaux, on peut y définir les notions de pgcd et de ppcm, que le théorème de Gauss y reste valable avec toutes ses conséquences, bien qu'on ne puisse utiliser en général le théorème de Bezout. Parmi les propriétés remarquables des anneaux factoriels, citons le fait que si A est un anneau factoriel, l'anneau des polynômes (à une ou plusieurs indéterminées) à coefficients dans A est lui aussi un anneau factoriel. C'est le cas par exemple de $\mathbb{Z}[X]$ (cf. exercice 3).

Nous ne prouverons pas ici qu'un anneau principal est factoriel : ce n'est pas bien difficile mais cela ne présenterait que très peu d'intérêt dans le cadre de ce livre. Signalons que l'*unicité* de la factorisation en irréductibles dans un anneau principal est conséquence immédiate du théorème de Gauss, valable dans un tel anneau (voir fin du § VII. 2). L'*existence* de la factorisation est moins évidente et ne peut, en général, être prouvée de façon *constructive*, i.e. on n'a pas d'algorithme conduisant à cette factorisation. Toutefois pour une large catégorie d'anneaux principaux, un tel algorithme existe :

THÉOREME VII.3.4

|| Soit A un anneau intègre qui n'est pas un corps et muni d'un stathme euclidien φ vérifiant la condition suivante ⁽¹⁾ :

(I) Si x divise y dans $A \setminus \{0\}$ et si x et y sont non associés, on a : $\varphi(x) < \varphi(y)$.

|| Alors l'anneau A est factoriel et la factorisation en éléments irréductibles des éléments de $A \setminus \{0\}$ peut s'obtenir par algorithme.

⁽¹⁾ Condition dont on voit aisément qu'elle est satisfaite si l'axiome (I) de la définition VII.2.1 l'est.

Démonstration :

Soit \mathcal{U}_A le groupe des éléments inversibles de A , et soit m la valeur minimum de φ . La condition (I) montre d'abord que :

$$\mathcal{U}_A = \{u \in A \setminus \{0\} \mid \varphi(u) = m\}.$$

L'étude précédant l'énoncé VII.3.4 a montré qu'il suffit de prouver l'existence de la factorisation puisque A est déjà principal (cf. théorème VII.2.3). Soit donc $x \in A \setminus \{0\}$ un élément non inversible.

Il existe des suites (x_0, x_1, \dots, x_k) dans $A \setminus \{0\}$, avec $k \geq 1$, telles que : $x_0 = x$, $x_k \in \mathcal{U}_A$ et, pour $0 \leq i \leq k-1$, x_{i+1} divise x_i mais x_{i+1} et x_i sont non associés (par exemple, $(x_0 = x, x_1 = 1)$ est une telle suite). Pour une de ces suites, on a :

$$m = \varphi(x_k) < \varphi(x_{k-1}) < \dots < \varphi(x_0) = \varphi(x),$$

d'où $k \leq \varphi(x) - m$; on peut donc trouver une telle suite pour laquelle la valeur de k soit maximum. L'ayant ainsi choisie, posons :

$$p_1 = x_0/x_1, p_2 = x_1/x_2, \dots, p_k = x_{k-1}/x_k,$$

d'où $x = p_1 p_2 \dots p_k x_k$. Du fait que k est maximum, chaque p_i est irréductible, et comme $x_k \in \mathcal{U}_A$, on a donc bien décomposé x en produit d'irréductibles. Un algorithme possible pour obtenir cette factorisation consiste à choisir d'abord un diviseur q_1 non inversible de x sur lequel φ prenne la plus petite valeur possible, puis, si q_1 n'est pas associé à x , à recommencer avec x/q_1 à la place de x , et ainsi de suite : par un argument analogue à celui qui précède, l'algorithme s'arrête. ■

Exercice 1 : Montrer que les idéaux premiers de $K[X]$ (revoir la définition d'un idéal premier d'un anneau commutatif au § III.7, exercice 8) sont les idéaux $\{0\}$ et $PK[X]$, où P est irréductible.

Exercice 2 : Soit $v : K[X] \rightarrow \overline{\mathbb{N}}$ une application telle que : $v(0) = +\infty$; $v(F) \in \mathbb{N}$ pour tout $F \neq 0$; $v(F) = 0$ si $F = \lambda \in K^*$; il existe F tel que $v(F) \geq 1$;

$$v(F + G) \geq \min(v(F), v(G))$$

et enfin

$$v(FG) = v(F) + v(G)$$

pour tous F et G dans $K[X]$. Montrer que l'ensemble \mathfrak{p} des $F \in K[X]$ tels que $v(F) \geq 1$ est un idéal premier non nul, et en désignant par P le générateur normalisé de \mathfrak{p} (cf. exercice 1), montrer que $v = m \text{ val}_P$ pour un $m \in \mathbb{N}^*$ convenable.

Exercice 3 : On considère le sous-anneau $\mathbb{Z}[X]$ de l'anneau $\mathbb{Q}[X]$.

a) Si p est un nombre premier dans \mathbb{Z} , vérifier que p est élément irréductible de $\mathbb{Z}[X]$ mais non de $\mathbb{Q}[X]$, et que pX est irréductible dans $\mathbb{Q}[X]$ mais non dans $\mathbb{Z}[X]$.

b) Soit $F \in \mathbb{Z}[X] \setminus \{0\}$ et $\mathcal{C}(F)$ le pgcd dans \mathbb{Z} de ses coefficients (appelé le *contenu* de F). F est dit *primitif* ssi $\mathcal{C}(F) = 1$ (par exemple tout polynôme normalisé est primitif). Montrer que si F et G sont primitifs, alors FG l'est aussi (théorème dû à Gauss). *Indication :* soit

$$F = a_0 X^d + a_1 X^{d-1} + \dots + a_d, \quad G = b_0 X^{d'} + b_1 X^{d'-1} + \dots + b_{d'}, \quad FG = \sum_{i=0}^{d+d'} c_i X^{d+d'-i}.$$

Si le nombre p , premier dans \mathbb{Z} , divisait tous les c_i , il diviserait $a_0 b_0$, donc par exemple a_0 mais pas tous les a_i . Soit a_j le premier coefficient non divisible par p , b_k le premier coefficient de G non divisible par p . Que dire de c_{j+k} ?

En déduire que si F et G sont dans $\mathbb{Z}[X] \setminus \{0\}$, $\mathcal{C}(FG) = \mathcal{C}(F) \mathcal{C}(G)$.

c) Soit $F \in \mathbb{Z}[X] \setminus \{0\}$ non constant. Montrer l'équivalence : F est irréductible dans $\mathbb{Z}[X]$ ssi F est primitif dans $\mathbb{Z}[X]$ et irréductible dans $\mathbb{Q}[X]$.

d) Quels sont les éléments inversibles de l'anneau $\mathbb{Z}[X]$? Montrer que $\mathbb{Z}[X]$ est factoriel.

Exercice 4 : Soit p un nombre premier dans \mathbb{Z} et C un polynôme de $\mathbb{Z}[X]$ de degré ≥ 2 , $X^n + c_1 X^{n-1} + \dots + c_n$ tel que

$$c_i \equiv 0 \pmod{p} \text{ pour } 1 \leq i \leq n \text{ et } c_n \not\equiv 0 \pmod{p^2}.$$

a) Démontrer que C est irréductible dans $\mathbb{Z}[X]$.

Indication : si $C = (X^d + a_1 X^{d-1} + \dots + a_d)(X^{d'} + b_1 X^{d'-1} + \dots + b_{d'})$, produit de deux polynômes A et B non constants et normalisés de $\mathbb{Z}[X]$, montrer d'abord que $p \mid a_d$ ou $p \mid b_{d'}$; si par exemple p divise tous les a_i à partir de a_k , prouver que $p \nmid b_{d'}$ et aboutir à une contradiction en considérant le coefficient $c_{i-1+d'}$ (c'est le critère d'Eisenstein).

b) En déduire que C est irréductible dans $\mathbb{Q}[X]$ en se servant de l'exercice 3.

c) Soit p un nombre premier dans \mathbb{Z} . Montrer que le polynôme Φ défini par $\Phi = X^{p-1} + X^{p-2} + \dots + 1$ est irréductible dans $\mathbb{Q}[X]$. *Indication :* poser $X = Y + 1$.

d) Si $\alpha \in \mathbb{N}^*$ et p premier dans \mathbb{Z} , prouver de la même façon que le polynôme

$$\Psi = X^{p^\alpha - 1} + X^{p^\alpha - 1 - (p-1)} + X^{p^\alpha - 1 - (p-2)} + \dots + 1$$

est irréductible dans $\mathbb{Q}[X]$.

e) Soit $n \in \mathbb{N}^*$. Trouver des nombres entiers $a \in \mathbb{Z}^*$ tels que $X^n - a$ soit irréductible dans $\mathbb{Q}[X]$.

Exercice 5 : Soit $F \in \mathbb{Z}[X]$ non constant et normalisé. Si $G \in \mathbb{Q}[X]$ divise F , est non constant et normalisé, alors $G \in \mathbb{Z}[X]$. En déduire les diviseurs de F dans $\mathbb{Q}[X]$ en fonction des diviseurs de F dans $\mathbb{Z}[X]$.

Exercice 6 : a) Décrire une méthode systématique permettant de trouver tous les facteurs du premier degré $aX + b$ d'un polynôme donné F de $\mathbb{Z}[X]$ avec $a \in \mathbb{Z}^*$, $b \in \mathbb{Z}$.

b) Pour quelles valeurs de l'entier rationnel n le polynôme $2X^2 + nX - 7$ n'est-il pas irréductible dans $\mathbb{Q}[X]$? (cf. exercice 3).

c) Montrer que $X^3 + 3X - 1$ est irréductible dans $\mathbb{Q}[X]$.

Exercice 7 : a) Utiliser le critère d'Eisenstein (cf. exercice 4) pour prouver que les polynômes suivants sont irréductibles sur le corps \mathbb{Q} : $X^3 + 2X^2 + 4X + 2$; $X^5 + 3X + 2$.

b) Généraliser ce critère à un polynôme C de $\mathbb{Z}[X]$ de degré impair $2n + 1$, normalisé, vérifiant les conditions $c_i \equiv 0 \pmod{p}$ pour $1 \leq i \leq n$, $c_i \equiv 0 \pmod{p^2}$ pour

$$n + 1 \leq i \leq 2n + 1 \text{ et } c_{2n+1} \not\equiv 0 \pmod{p^3}.$$

Exercice 8 : Soit A l'anneau des entiers du corps $\mathbb{Q}[\sqrt{-5}]$ (cf. exercice 19 du § VII.2), c'est-à-dire l'ensemble des nombres complexes $a + b i \sqrt{5}$, avec $a \in \mathbb{Z}$ et $b \in \mathbb{Z}$.

a) Montrer, en utilisant $N(a + ib \sqrt{5}) = a^2 + 5b^2$ que cet anneau n'est pas factoriel.

Indication : on pourra utiliser par exemple le fait que

$$9 = 3 \times 3 = (2 + i \sqrt{5})(2 - i \sqrt{5})$$

alors que 3 et $2 + i \sqrt{5}$ ne sont pas associés bien qu'irréductibles.

b) Trouver dans A toutes les factorisations possibles de 21.

Exercice 9 : a) Factoriser le polynôme $X^4 - X^2 + 1$ sur le corps $\mathbb{Z}/11\mathbb{Z}$.

b) Factoriser le polynôme $X^5 + X + 1$ sur le corps des rationnels.

c) Montrer que les polynômes $X^4 + X + 1$ et $X^6 + X^2 + 1$ sont irréductibles dans $\mathbb{Q}[X]$.

Exercice 10 : Soit c fixé et n arbitraire tous deux dans \mathbb{N}^* . Montrer que l'ensemble des polynômes de $\mathbb{Q}[X]$: $F = X^n + r_1 X^{n-1} + \dots + r_n$, qui sont irréductibles dans $\mathbb{Q}[X]$ et tels que $|r_i| \leq c$ pour tout i est infini (on pourra utiliser l'exercice 4).

§ VII.4 FONCTIONS POLYNÔMES, RACINES

Dans ce paragraphe, le corps de base K est fixé. Nous avons défini au § VI.4 ce qu'est une K -algèbre. Rappelons que pour nous, sauf mention expresse du contraire, une K -algèbre est obligatoirement *associative* et *unifère*.

Substitution de x à l'indéterminée X

THÉOREME VII.4.1

|| Soit \mathcal{A} une K -algèbre, d'élément unité e ; pour tout $x \in \mathcal{A}$ il existe un, et un seul, homomorphisme de K -algèbres $\varphi : K[X] \rightarrow \mathcal{A}$ tel que $\varphi(X) = x$.

Démonstration :

Si φ existe, pour tout $F \in K[X]$, écrivons

$$F = a_0 + a_1 X + \cdots + a_n X^n ; \quad \text{alors :}$$

$$\begin{aligned} \varphi(F) &= \varphi\left(\sum_{k=0}^n a_k X^k\right) = \sum_{k=0}^n \varphi(a_k X^k) = \\ &= \sum_{k=0}^n a_k [\varphi(X)]^k = \sum_{k=0}^n a_k x^k \quad (\text{rappel : } x^0 = e). \end{aligned}$$

Réciproquement, à chaque $F = \sum_{k=0}^n a_k X^k \in K[X]$,

associons précisément $\varphi(F) = \sum_{k=0}^n a_k x^k$.

L'application $\varphi : K[X] \rightarrow \mathcal{A}$ ainsi définie est un homomorphisme de K -algèbres à cause de la définition même des opérations (somme, produit et produit par un scalaire de K) dans $K[X]$, et d'autre part on a bien $\varphi(X) = x$, donc φ existe et est unique. ■

Si $x \in \mathcal{A}$, nous noterons φ_x l'homomorphisme de K -algèbres défini dans la démonstration du théorème VII.4.1.

DÉFINITION VII.4.1

» Avec les hypothèses et notations ci-dessus, pour $x \in \mathcal{A}$, l'homomorphisme de K -algèbres $\varphi_x : K[X] \rightarrow \mathcal{A}$ tel que $\varphi_x(X) = x$ s'appelle **homomorphisme de substitution défini par x** . Pour $F \in K[X]$, l'élément $\varphi_x(F)$ s'appelle **la valeur de F en x** , dite obtenue par **substitution de x à X dans F** et notée $\tilde{F}(x)$ (ou le plus souvent $F(x)$). L'idéal $\text{Ker}(\varphi_x)$ de $K[X]$, c'est-à-dire l'idéal

$$\{F \in K[X] \mid \tilde{F}(x) = 0\}$$

» s'appelle **idéal des relations algébriques liant x** .

Il est immédiat que l'image de l'homomorphisme de substitution $\varphi_x: K[X] \rightarrow \mathcal{A}$, $F \mapsto \tilde{F}(x)$ (x fixé, $x \in \mathcal{A}$) n'est autre que le sous K -ev de \mathcal{A} engendré par la famille $(x^n)_{n \in \mathbb{N}}$. Cette image est donc tout simplement la sous- K -algèbre engendrée par x dans \mathcal{A} (cf. § VI.4, exemple 9) que nous avons convenu de noter $K[x]$.

Si $\text{Ker}(\varphi_x) = \{0\}$, φ_x est injectif et définit, par corestriction à $K[x]$, un isomorphisme de K -algèbres $K[X] \rightarrow K[x]$, $F \mapsto F(x)$.

Si $\text{Ker}(\varphi_x) \neq \{0\}$, il existe un unique générateur normalisé μ_x de cet idéal de $K[X]$, et alors les polynômes $G \in K[X]$ tels que $\tilde{G}(x) = 0$ sont exactement les multiples de μ_x .

DÉFINITION VII.4.2

Soit \mathcal{A} une K -algèbre et $x \in \mathcal{A}$. L'élément x est dit **algébriquement libre sur K** ssi l'homomorphisme de substitution $\varphi_x: K[X] \rightarrow \mathcal{A}$ est injectif.
L'élément x est dit **algébriquement lié sur K** ssi l'homomorphisme φ_x est **non injectif**, et dans ce cas l'unique polynôme normalisé $\mu_x \in K[X]$ tel que $\text{Ker}(\varphi_x) = \mu_x K[X]$ s'appelle le **polynôme minimal de x sur K** .

Pour fixer les idées, imaginons qu'on substitue à X un polynôme $P \in K[X]$. L'image du polynôme P par F (ou valeur de F en P) est un polynôme de $K[X]$ noté $F(P)$. Si P est non constant, l'homomorphisme φ_P est injectif (cf. exemple 2).

Si maintenant on substitue à X une matrice $M \in \mathfrak{M}_n(K)$, l'image de la matrice M par F (ou valeur de F en M) est une matrice de $\mathfrak{M}_n(K)$ notée $F(M)$. Nous verrons que l'homomorphisme φ_M est non injectif, ce qui permet de définir le polynôme minimal de la matrice M sur K .

Mais un cas particulier très important est celui où \mathcal{A} est un *corps commutatif extension de K* (cf. définition II.6.4). Dans ce cas on utilise un langage approprié :

DÉFINITION VII.4.3

Soit L un corps commutatif extension du corps K . Les éléments $x \in L$ algébriquement libres sur K sont dits **transcendants sur K** ; les éléments algébriquement liés sur K sont dits **algébriques sur K** . Le degré du polynôme minimal μ_x pour x algébrique sur K est appelé **degré sur K de l'élément x** .
En particulier si $K = \mathbb{Q}$ et $L = \mathbb{C}$, les nombres complexes transcendants sur \mathbb{Q} sont appelés tout simplement **nombres transcendants** ; les nombres complexes algébriques sur \mathbb{Q} sont appelés **nombres algébriques**.

Le degré sur \mathbb{Q} d'un nombre algébrique est simplement appelé son *degré*. Un élément $x \in L$ appartient à K ssi c'est un élément algébrique de degré 1 sur K ; s'il en est ainsi, son polynôme minimal sur K est X

Exemple 1 : La K -algèbre \mathcal{A} étant quelconque, prenons $x = 0_{\mathcal{A}}$. L'homomorphisme de substitution obtenu φ_0 associe, à $F \in K[X]$, l'élément $c_0(F)e$, $c_0(F)$ désignant le terme constant de F ; φ_0 n'est pas injectif, l'idéal $\text{Ker}(\varphi_0)$ est l'ensemble des multiples de X .

Exemple 2 : Prenons $\mathcal{A} = K[X]$; si $P \in \mathcal{A}$, l'homomorphisme de substitution φ_P associe à $F = \sum_{k=0}^n a_k X^k$ le polynôme $\sum_{k=0}^n a_k P^k \in \mathcal{A}$.

Au lieu de le noter $F(P)$, on l'écrit souvent $F \circ P$ et on l'appelle le *composé* de F et de P , ce qui se justifie par le fait que la valeur de $F \circ P$ en $x \in L$ est bien $F(P(x))$. On voit d'après le degré que φ_P est injectif ssi $\deg(P) \geq 1$. En particulier, si $P = X$, φ_X est l'application identique de $K[X]$, ce qui laisse le choix entre trois notations pour le même polynôme de $K[X]$: F , $F \circ X$ ou $F(X)$.

Exemple 3 : Le nombre π (rappelons que $\frac{\pi}{2}$ est la plus petite racine positive de l'équation $\cos x = 0$, où $x \in \mathbb{R}$), le nombre e (base des logarithmes naturels, défini comme somme de la série $\sum_{n=0}^{+\infty} \frac{1}{n!}$), le nombre e^π sont transcendants. Ces résultats ont été assez difficiles à obtenir : la transcendance de e n'a été démontrée qu'en 1873 par Hermite ⁽¹⁾ celle de π par Lindemann ⁽²⁾ en 1882, ce qui a mis un point final à l'antique problème de la « quadrature du cercle », celle de e^π par Gelfond ⁽³⁾ en 1929, et de façon indépendante par Schneider ⁽⁴⁾ en 1934.

Le nombre i est algébrique de degré 2 ; le nombre $\sqrt{2} + \sqrt{3} + \sqrt{5}$ est algébrique de degré 8 (le lecteur pourra chercher son polynôme minimal).

Remarque 1 : La notion de transcendance (ou d'élément algébrique) est *relative au corps de base* K qui doit toujours être précisé (sauf si c'est \mathbb{Q}) ; par exemple $\sqrt{\pi}$ est un nombre transcendant (sous-entendu sur \mathbb{Q}) ; mais $\sqrt{\pi}$ est algébrique de degré 1 sur \mathbb{R} ; et algébrique de degré 2 sur le sous-corps $\mathbb{Q}(\pi)$ de \mathbb{R} .

Exemple 4 : *Extension du corps de base.*

Soit L un corps commutatif extension de K ; soit X (resp. Y) l'indéterminée canonique sur K (resp. sur L). L'homomorphisme de substitution

⁽¹⁾ Charles *Hermite*, mathématicien français (1822-1901), qui a eu l'honneur de voir son nom adjectivé (endomorphisme hermitien, forme, matrice, norme hermitiennes).

⁽²⁾ Ferdinand *Lindemann* (1852-1939), mathématicien allemand.

⁽³⁾ A. *Gelfond*, mathématicien russe contemporain, a prouvé que si $\alpha \neq 0$, α et e^α ne peuvent être tous deux algébriques et que si $\alpha \notin \{0, 1\}$ et β algébrique irrationnel, α^β est transcendant (par exemple $(-1)^{-i}$).

⁽⁴⁾ T. *Schneider*, mathématicien allemand contemporain, a écrit en 1957 une « Introduction aux nombres transcendants ».

$K[X] \rightarrow L[Y], F \mapsto \tilde{F}_{\mathcal{A}}(Y)$ (où \mathcal{A} désigne $L[Y]$) associe, à $F = \sum_{k=0}^n a_k X^k$,

le polynôme $\sum_{k=0}^n a_k Y^k$ de $L[Y]$. On voit qu'il est injectif. A l'aide de cet

homomorphisme on peut identifier $K[X]$ à la sous- K -algèbre $K[Y]$ de $L[Y]$ définie dans le théorème VII.1.6. Procéder à cette identification, c'est, par définition, « considérer un polynôme F de $K[X]$ comme un polynôme à coefficients dans L ».

Le changement de lettre pour désigner l'indéterminée sur K et sur L ne traduit pas un simple souci de formalisme algébrique (en fait on peut l'éviter dans le cas où aucune confusion n'est possible). Il se révèle indispensable si, par exemple, on prend pour corps L le corps $K(X)$ des fractions de l'anneau intègre $K[X]$ appelé *corps des fractions rationnelles de l'indéterminée X sur K* : dans ce cas, si on ne changeait pas de lettre, l'homomorphisme $F \mapsto \tilde{F}_{\mathcal{A}}(X)$ enverrait $K[X]$ dans les constantes de $L[Y]$, ce qui n'est évidemment pas le but souhaité ! (N.B. : Cette question ne s'éclaire pleinement qu'à la lumière de la théorie du *produit tensoriel*, cf. [2] et [7]).

Pour éviter toute confusion, si l'on veut identifier $K[X]$ à la sous- K -algèbre $K[Y]$ de $L[Y]$, le plus simple est de se donner directement une indéterminée X sur L .

Fonctions polynômes

DÉFINITION VII.4.4

$\left\{ \begin{array}{l} \text{Soit } \mathcal{A} \text{ une } K\text{-algèbre et } F \in K[X] ; \text{ l'application } \mathcal{A} \rightarrow \mathcal{A}, \\ x \mapsto \tilde{F}(x) \text{ est appelée } \textbf{fonction polynôme} \text{ (ou fonction polynomiale)} \\ \text{définie par } F \text{ dans } \mathcal{A}. \end{array} \right.$

Cette fonction polynomiale sera notée $\tilde{F}_{\mathcal{A}}$ et dépend évidemment du choix de la K -algèbre \mathcal{A} , ce qui entraîne en particulier qu'au même polynôme F peut correspondre une multitude de fonctions polynomiales différentes. Si on prend pour \mathcal{A} le corps K lui-même, on notera l'application $K \rightarrow K, x \mapsto \tilde{F}(x)$ correspondante \tilde{F} , ou même F , si cela ne risque pas d'entraîner de confusion.

Soit alors $\mathcal{F}(\mathcal{A})$ l'ensemble des fonctions de \mathcal{A} dans \mathcal{A} , muni de sa structure naturelle de K -algèbre (c'est-à-dire que la somme et le produit sont définis « point par point ») ; on remarque que si $F \in K[X]$, la fonction polynôme $\tilde{F}_{\mathcal{A}}$ n'est autre que la valeur de F en l'élément $\text{Id}_{\mathcal{A}}$ (application identique de \mathcal{A} dans \mathcal{A}) de $\mathcal{F}(\mathcal{A})$: ainsi la substitution nous permet de récupérer globalement la fonction polynomiale $\tilde{F}_{\mathcal{A}}$. Sans nouvelle démonstration, on en déduit que l'application

$$K[X] \rightarrow \mathcal{F}(\mathcal{A}), \quad F \mapsto \tilde{F}_{\mathcal{A}},$$

est un *homomorphisme de K -algèbres*, qui n'est autre que l'homomorphisme de substitution $F \mapsto \tilde{F}(\text{Id}_{\mathcal{A}})$. Lorsque cet homomorphisme €

qui arrive, nous le verrons, sous des hypothèses très générales, il n'y a aucun inconvénient à confondre un polynôme F avec la fonction polynôme qu'il définit sur \mathcal{A} ; mais s'il n'est pas injectif, il faut évidemment distinguer très soigneusement le polynôme et la fonction qu'il définit.

Exemple 5 : Soit K un corps fini, de cardinal q . Désignons par F le polynôme $\prod_{x \in K} (X - x)$ qui est de degré q et par \tilde{F} la fonction polynôme de K dans K associée. Il est clair que \tilde{F} est la fonction nulle de K dans K alors que F est loin d'être le polynôme nul !

Zéros d'un polynôme

DÉFINITION VII.4.5

Soit $F \in K[X]$ et \mathcal{A} une K -algèbre. Un élément $x \in \mathcal{A}$ est appelé un **zéro de F dans \mathcal{A}** ssi $F_{\mathcal{A}}(x) = 0$ (on dit encore que **x annule F**). Dans le cas particulier où \mathcal{A} est un corps commutatif extension de K , un zéro de F dans \mathcal{A} est appelé une **racine de F dans \mathcal{A}** .

THÉORÈME VII.4.2

Soit $F \in K[X]$.
 (I) Un élément $a \in K$ est racine de F ssi F est divisible par $X - a$ dans $K[X]$.
 (II) Soit $n \in \mathbb{N}$; si $F \in K_n[X]$ et si F admet au moins $n + 1$ racines dans K , alors $F = 0$.

Démonstration :

(I) La division euclidienne de F par $X - a$ dans $K[X]$ donne :

$$F = (X - a)Q + R, \quad \text{avec } \deg(R) < 1,$$

donc $R \in K$. Substituant a à X dans cette relation, on obtient : $R = \tilde{F}(a)$. Cela prouve bien que $\tilde{F}(a) = 0$ ssi F est divisible par $X - a$ ⁽¹⁾.

(II) Pour $n = 0$ la propriété est évidente, car F est constant. Supposons donc la propriété vraie jusqu'à l'ordre $n - 1 \geq 0$ et montrons qu'elle est encore vraie pour n . Soit a_1, a_2, \dots, a_{n+1} des éléments distincts dans K qui soient racines de F . D'après (I) on a :

$$F = (X - a_{n+1})G,$$

avec $G \in K[X]$. Pour $i \in \llbracket 1, n \rrbracket$, on a :

$$\tilde{F}(a_i) = (a_i - a_{n+1}) \tilde{G}(a_i) = 0,$$

⁽¹⁾ La propriété (I) du théorème VII.4.2 reste vraie avec des polynômes à coefficients dans un anneau commutatif quelconque (utiliser le développement de $X^k - a^k$, relation (5) § III.2). En revanche, (II) ne s'étend pas si l'anneau des coefficients n'est pas intègre

mais comme $a_i - a_{n+1} \neq 0$, c'est que $G(a_i) = 0$. Or le degré de G , inférieur d'une unité à celui de F , est $\leq n$. L'hypothèse de récurrence montre donc que $G = 0$, d'où $F = 0$. ■

Remarque 2 : La propriété (II) ne serait plus vraie si l'on remplaçait les mots « $n + 1$ racines dans K » par « $n + 1$ zéros dans la K -algèbre \mathcal{A} ». Par exemple si $K = \mathbb{R}$, le polynôme du second degré $X^2 + 1$ admet *une infinité de zéros* si l'on prend pour \mathcal{A} la \mathbb{R} -algèbre des quaternions ou tout simplement la \mathbb{R} -algèbre $\mathfrak{M}_2(\mathbb{R})$ des matrices carrées d'ordre deux à coefficients réels.

COROLLAIRE 1

|| Soit $n \in \mathbb{N}^*$ et a_1, a_2, \dots, a_n des éléments distincts dans K ; les polynômes $F \in K[X]$ tels que $\tilde{F}(a_i) = 0$ pour $i \in \llbracket 1, n \rrbracket$ sont les multiples de

$$(X - a_1)(X - a_2) \dots (X - a_n).$$

La démonstration se fait par récurrence sur n en utilisant le Th. VII.4.2.

COROLLAIRE 2

|| Si K est un corps **infini**, l'homomorphisme naturel de K -algèbres $K[X] \rightarrow \mathcal{F}(K)$, $F \mapsto \tilde{F}$ est **injectif**.

C'est la raison pour laquelle lorsque K est infini, on peut sans inconvénient identifier un polynôme $F \in K[X]$ avec la fonction $\tilde{F} : K \rightarrow K$ qu'il définit.

Remarque 3 : Si \mathcal{D} est une partie *arbitraire* du corps K , au polynôme $F \in K[X]$, on peut associer la **fonction polynomiale** $\mathcal{D} \rightarrow K$, restriction à \mathcal{D} de $\tilde{F} : K \rightarrow K$. Lorsque \mathcal{D} est infini, il est clair que l'application $F \mapsto \tilde{F}|_{\mathcal{D}}$ permet d'identifier $K[X]$ à la K -algèbre des fonctions polynomiales sur \mathcal{D} .

Polynôme d'interpolation de Lagrange

Les fonctions polynômes étant parmi les plus simples que l'on connaisse, il est naturel de se poser le problème de la recherche de polynômes $F \in K[X]$ prenant des valeurs données dans $K : b_1, b_2, \dots, b_n$ ($n \in \mathbb{N}^*$) pour des valeurs données et *distinctes* a_1, a_2, \dots, a_n de la « variable » dans K . Si F est un tel polynôme, il est clair qu'en posant

$$P(X) = (X - a_1)(X - a_2) \dots (X - a_n),$$

$F + GP$ en est un autre. Le problème proposé n'admet donc pas une solution unique. Plus précisément :

THÉORÈME VII.4.3

Soit $n \in \mathbb{N}$; a_1, a_2, \dots, a_n des éléments distincts dans K , et b_1, b_2, \dots, b_n dans K .

(I) Il existe un et un seul polynôme $\Lambda \in K_{n-1}[X]$ tel que

$$\tilde{\Lambda}(a_i) = b_i \quad \text{pour } 1 \leq i \leq n.$$

(II) L'ensemble des polynômes $F \in K[X]$ tels que $\tilde{F}(a_i) = b_i$ pour $1 \leq i \leq n$ est l'ensemble des polynômes $\Lambda + GP$, où $G \in K[X]$, où $P = (X - a_1)(X - a_2) \dots (X - a_n)$, et où Λ désigne le polynôme défini en (I).

Démonstration :

(II) est une conséquence du corollaire 1 du théorème VII.4.2, car si $F_1 \in K[X]$ et $F_2 \in K[X]$ prennent les mêmes valeurs en a_1, a_2, \dots, a_n , alors la différence $F_1 - F_2$ s'annule en a_1, a_2, \dots, a_n et est donc multiple de P . Prouvons maintenant l'existence de Λ en le construisant. Pour chaque $i \in \llbracket 1, n \rrbracket$ on peut former le polynôme

$$\Lambda_i(X) = \frac{1}{\prod_{j \neq i} (a_i - a_j)} \prod_{j \neq i} (X - a_j).$$

C'est un polynôme de $K_{n-1}[X]$ qui prend la valeur 1 en a_i et la valeur 0 en a_j pour $j \neq i$ d'après sa construction même. Il en résulte que le polynôme $\Lambda(X) = \sum_{i=1}^n b_i \Lambda_i(X)$ qui est dans $K_{n-1}[X]$ vérifie $\Lambda(a_i) = b_i$ pour $i \in \llbracket 1, n \rrbracket$. L'unicité de Λ dans $K_{n-1}[X]$ résulte de (II). ■

La relation
$$\Lambda(X) = \sum_{i=1}^n b_i \Lambda_i(X) = \sum_{i=1}^n b_i \frac{1}{\prod_{j \neq i} (a_i - a_j)} \prod_{j \neq i} (X - a_j)$$

s'appelle **formule d'interpolation de Lagrange**.

Le corollaire 1 du théorème VII.4.2 que nous venons d'utiliser s'avère un outil essentiel permettant de *décomposer un polynôme en produit de facteurs*.

Exemple 6 : Soit $K = \mathbb{C}$ et n fixé dans \mathbb{N}^* . Nous avons vu (cf. théorème VI.8.2) que le polynôme $X^n - 1$ admet les n racines *distinctes* éléments du groupe μ_n . D'après le corollaire 1 du théorème VII.4.2 on a donc :

$$X^n - 1 = G \prod_{\zeta \in \mu_n} (X - \zeta) \quad \text{avec } G \in \mathbb{C}[X].$$

La comparaison des degrés et des coefficients dominants montre que $G = 1$. On en déduit l'égalité

$$(1) \quad \boxed{X^n - 1 = \prod_{\zeta \in \mu_n} (X - \zeta)}.$$

En divisant les deux membres de (1) par le facteur non nul $X - 1$, on obtient :

$$(2) \quad 1 + X + X^2 + \cdots + X^{n-1} = \prod_{\zeta \in \mu_n \setminus \{1\}} (X - \zeta).$$

Corps algébriquement clos

THÉORÈME VII.4.4

|| Pour le corps K , les propriétés suivantes sont équivalentes :

- (I) les seuls polynômes irréductibles de $K[X]$ sont ceux de degré 1
- (II) tout polynôme **non constant** $F \in K[X]$ admet au moins une racine dans K
- (III) tout polynôme $F \in K[X] \setminus \{0\}$ est un produit de polynômes de degré 1.

Démonstration :

Puisque les polynômes de degré 1 sont irréductibles, il est clair, compte tenu des résultats du § VII.3, que (I) et (III) sont équivalentes. Il est encore plus évident que (III) \Rightarrow (II). Il reste à prouver que (II) \Rightarrow (III). Supposons donc la propriété (II) vérifiée par le corps K . Une récurrence immédiate sur l'entier $\deg(F)$, en utilisant la première partie du théorème VII.4.2, montre que si $F \in K[X]$ est non constant, F est produit de facteurs de degré 1, d'où (III). ■

DÉFINITION VII.4.6

⎧ Le corps K est dit **algébriquement clos** ssi il vérifie l'une des
⎧ conditions équivalentes du théorème VII.4.3.

L'exemple le plus célèbre de corps algébriquement clos est fourni par :

THÉORÈME VII.4.5 (théorème de d'Alembert) ⁽¹⁾

|| Le corps \mathbb{C} des nombres complexes est algébriquement clos.

⁽¹⁾ Jean Le Rond d'Alembert, mathématicien et philosophe français (1717-1783). Énoncé en 1746, son théorème sera démontré plus tard par Gauss. Il porte aussi le nom de *théorème fondamental de l'Algèbre* bien qu'il soit impossible d'en donner une démonstration purement algébrique.

Ce théorème, dont toute démonstration nécessite des connaissances d'Analyse, sera prouvé dans le tome 2 de cet ouvrage, mais bien entendu nous l'utiliserons dès maintenant ainsi que ses conséquences.

Voici un autre exemple intéressant :

Exemple 7 : L'ensemble des nombres algébriques forme un sous-corps, noté $\widehat{\mathbb{Q}}$, de \mathbb{C} . Le corps $\widehat{\mathbb{Q}}$ est algébriquement clos (on pourra consulter tout traité de théorie des nombres algébriques, et par exemple [28]).

En revanche le corps \mathbb{Q} , le corps \mathbb{R} ne sont pas algébriquement clos (essayer de factoriser $X^2 + 1$). Il en est de même de tout corps fini (essayer de factoriser $1 + \prod_{x \in K} (X - x)$). Heureusement on démontre (théorème de Steinitz) que tout corps commutatif peut être plongé dans un corps algébriquement clos.

Multiplicité d'une racine

Revenons à un corps commutatif quelconque K ; l'ensemble \mathcal{J}_1 des polynômes $\{X - a\}_{a \in K}$ est une partie de l'ensemble des polynômes irréductibles et normalisés de $K[X]$. Pour chaque $a \in K$, et chaque $F \in K[X] \setminus \{0\}$, l'entier $\nu_a(F) = \text{val}_{X-a}(F)$ est donc défini. C'est, par définition, le plus grand des entiers $k \in \mathbb{N}$ tels que $(X - a)^k$ divise F . On l'appelle la **multiplicité de a dans F** (ou encore la *multiplicité de a comme racine de F*) ; a est racine de F ssi $\nu_a(F) \geq 1$. D'après la première partie du théorème VII.4.2, $\nu_a(F)$ est le seul entier $k \in \mathbb{N}$ tel que F puisse s'écrire $F = (X - a)^k G$ avec $G \in K[X]$ et $\tilde{G}(a) \neq 0$.

Polynômes dissociés sur K

Les polynômes $F \in K[X] \setminus \{0\}$ dont tous les facteurs irréductibles dans $K[X]$ sont de degré 1 sont ceux qui s'écrivent :

$$F = u \prod_{a \in K} (X - a)^{\nu_a(F)}, \quad \text{avec } u \in K^*.$$

Ces polynômes sont dits **dissociés sur K** (certains auteurs disent aussi : **scindés sur K**).

Le théorème VII.3.1 et le théorème VII.4.2(I) montrent qu'aux facteurs de K^* près, un polynôme non nul $F \in K[X] \setminus \{0\}$ s'écrit de manière unique sous la forme $F = GH$, où $G \in K[X]$ est **dissocié** (éventuellement constant), et H est un polynôme (éventuellement constant lui aussi) **sans racine dans K** . On peut appeler G la **partie dissociée** de F . Le corps K est algébriquement clos ssi tout polynôme $F \in K[X] \setminus \{0\}$ est dissocié sur K .

Exemple 8 : Si K est un corps fini de cardinal q , le polynôme $X^q - X$ est dissocié et vaut $\prod_{x \in K} (X - x)$. En effet, soit

$$A = X^q - X - \prod_{x \in K} (X - x) ;$$

c'est un polynôme de degré $\leq q - 1$ tel que $A(x) = 0$ pour tout $x \in K$ (cf. le corollaire 2 du théorème V.3.2). Donc, en utilisant le théorème VII.4.2, $A = 0$.

Exercice 1 : Soit K un corps commutatif et L une extension de K ; on considère l'algèbre de polynômes $L[X]$. Soit A, B dans $K[X]$ avec $B \neq 0$. Alors la division euclidienne de A par B dans $K[X]$ est aussi celle dans $L[X]$.

Application : B divise A dans $K[X]$ ssi B divise A dans $L[X]$.

Exercice 2 : Soit K un corps commutatif et soit $P \in K[X]$.

a) Montrer que l'application $K[X] \rightarrow K[X], F \mapsto F \circ P$ est un automorphisme de la K -algèbre $K[X]$ si, et seulement si, $\deg(P) = 1$.

b) Montrer que les automorphismes de la K -algèbre $K[X]$ sont précisément ceux définis au a). Ces automorphismes forment un groupe que l'on définira de manière simple.

Exercice 3 : Soit L un corps commutatif et K un sous-corps de L . On considère la sous- K -algèbre $K[X]$ de la L -algèbre $L[X]$. Si F_1, \dots, F_n sont des éléments de $K[X]$, tout pgcd (resp. ppcm) de F_1, \dots, F_n dans $K[X]$ est un pgcd (resp. un ppcm) de F_1, \dots, F_n dans $L[X]$.

Exercice 4 : a) Résoudre dans $\mathbb{C}[X]$ l'équation $P^2 + Q^2 = R^2$, où les polynômes inconnus P, Q, R sont non nuls.

b) Si $n \geq 3$ montrer que l'équation $P^n + Q^n = R^n$ (où P, Q et R sont dans $\mathbb{C}[X] \setminus \{0\}$) n'admet que les solutions de la forme $P = \lambda D, Q = \mu D, R = \nu D$, où $D \in \mathbb{C}[X] \setminus \{0\}$ et λ, μ, ν sont des éléments de \mathbb{C}^* vérifiant $\lambda^n + \mu^n = \nu^n$ (on commencera par le cas $n = 3$).

Exercice 5 : On suppose le corps K fini, de cardinal q , et on considère l'homomorphisme naturel de K -algèbres $f : K[X] \rightarrow \mathcal{F}(K), F \mapsto \tilde{F}$.

a) En utilisant le polynôme d'interpolation de Lagrange, montrer que f est surjectif, autrement dit que toute application de K dans K est une fonction polynomiale.

b) Montrer que l'idéal $\text{Ker}(f)$ est engendré par $\prod_{x \in K} (X - x) = X^q - X$.

Exercice 6 : On prend $K = \mathbb{Q}$. Soit a_1, a_2, \dots, a_n des entiers rationnels distincts. Démontrer que le polynôme $(X - a_1)(X - a_2) \dots (X - a_n) - 1$ est irréductible dans l'anneau $\mathbb{Z}[X]$, donc (cf. exercice 3 du § VII.3) irréductible aussi dans $\mathbb{Q}[X]$.

Exercice 7 : Pour les besoins de cet exercice, on admet que le nombre π est irrationnel.

a) Trouver tous les réels λ tels qu'il existe $F \in \mathbb{C}[X]$ vérifiant :

$$\forall \theta \in \mathbb{R} \quad e^{i\lambda\theta} = F(e^{i\theta}).$$

b) Soit $n \in \mathbb{N}^*$. Montrer, sans chercher à les calculer, qu'il existe des polynômes $\mathcal{T}_n \in \mathbb{Z}[X], \mathcal{U}_n \in \mathbb{Z}[X]$ de degré n , tels que :

$$\forall \theta \in \mathbb{R} \quad \mathcal{T}_n(2 \cos \theta) = 2 \cos n\theta ; \quad \mathcal{U}_n(2 \cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}.$$

c) Trouver tous les réels λ tels qu'il existe un polynôme $F \in \mathbb{C}[X]$ vérifiant : $\forall \theta \in \mathbb{R} \quad \cos \lambda\theta = F(\cos \theta)$. Résoudre la même question avec la condition :

$$\forall \theta \in \mathbb{R} \quad \frac{\sin \lambda\theta}{\sin \theta} = F(\cos \theta).$$

Exercice 8 : Avec les notations de l'exercice 7, les polynômes

$$T_n(X) = \frac{1}{2} \mathcal{T}_n(2X) \quad \text{et} \quad U_n(X) = \mathcal{U}_n(2X)$$

s'appellent respectivement *polynômes de Tchebychev* de première et de d

Obtenir les factorisations suivantes :

$$U_n(X) = 2^{n-1} \prod_{k=1}^n \left(X - \cos \frac{k\pi}{n+1} \right); \quad T_n(X) = 2^{n-1} \prod_{k=0}^{n-1} \left(X - \cos \left(\frac{\pi}{2n} + \frac{k\pi}{n} \right) \right).$$

N.B. : Le calcul des coefficients de U_n et T_n sera proposé en exercice au § VII.5.

Exercice 9 : En désignant par G le polynôme $1 + X^2 + X^4 + \dots + X^{2n} \in \mathbb{C}[X]$ et par U_n le polynôme de Tchebychev de seconde espèce de l'exercice précédent, on a :

$$(\forall \theta \in \mathbb{R}) \quad U_n(\cos \theta) = e^{-in\theta} G(e^{i\theta}).$$

Exercice 10 : a) Pour quels entiers $n \in \mathbb{N}^*$ existe-t-il un polynôme $F \in \mathbb{C}[X]$ tel que

$$(\forall \theta \in \mathbb{R}) \quad \sin n\theta = F(\sin \theta)?$$

Lorsque n convient, factoriser le polynôme F en facteurs de degré 1.

b) Trouver tous les réels λ tels qu'il existe $F \in \mathbb{C}[X]$ vérifiant :

$$(\forall \theta \in \mathbb{R}) \quad \sin \lambda \theta = F(\sin \theta).$$

Exercice 11 : Utiliser les formules (1) et (2) de l'exemple 6 et les formules d'Euler pour réduire les expressions suivantes :

$$a) P_n(x) = \prod_{k=0}^{n-1} \sin \left(x + \frac{k\pi}{n} \right), \quad x \in \mathbb{R}, n \in \mathbb{N}^*.$$

$$b) Q_n(x) = \prod_{k=0}^{n-1} \cos \left(x + \frac{k\pi}{n} \right), \quad x \in \mathbb{R}, n \in \mathbb{N}^*.$$

$$c) P_n = \prod_{k=1}^{n-1} \sin \frac{k\pi}{n} \quad \left(\text{Réponse : } \frac{n}{2^{n-1}} \right).$$

Exercice 12 : Calculer le polynôme d'interpolation de Lagrange $\Lambda \in \mathbb{C}[X]$ tel que

$$\Lambda(\omega^k) = (k+1)^r \quad \text{pour } k \in \llbracket 0, n-1 \rrbracket,$$

où $\omega = e^{\frac{2i\pi}{n}}$, et où $r \in \mathbb{N}^*$ est donné. On commencera par étudier à fond le cas $r=1$ et le cas $r=2$ et on indiquera une méthode pour r quelconque.

Exercice 13 : Opérateur de différence. Le corps de base est \mathbb{C} . Soit τ l'endomorphisme du \mathbb{C} -ev $\mathbb{C}[X]$ qui envoie $F \in \mathbb{C}[X]$ sur $F(X+1)$. On note Id l'opérateur identique de $\mathbb{C}[X]$ et on pose $\Delta = \tau - \text{Id}$.

a) Vérifier que, pour tout $n \in \mathbb{N}^*$, Δ envoie $\mathbb{C}_n[X]$ dans $\mathbb{C}_{n-1}[X]$. Préciser la valeur de $\Delta \langle^n \rangle(F)$ pour $F \in \mathbb{C}_n[X]$. On raisonnera dans la \mathbb{C} -algèbre $\text{Hom}_{\mathbb{C}}(\mathbb{C}[X])$.

b) En déduire les identités

$$\sum_{k=0}^n (-1)^k \binom{n}{k} (X-k)^p = \begin{cases} 0 & \text{si } n > p \\ n! & \text{si } n = p. \end{cases}$$

c) A partir de l'identité

$$\sum_{k=0}^n (-1)^{k+n} \binom{n}{k} k^n = n!,$$

qui est une conséquence de b), démontrer que le théorème de Fermat IV.4.5 entraîne le théorème de Wilson IV.4.6.

Exercice 14 : En utilisant seulement l'existence du polynôme $\mathcal{F}_n \in \mathbb{Z}[X]$ de l'exercice 7, résoudre l'équation $\cos r\pi = s$, où $r \in \mathbb{Q}_+^*$ et $s \in \mathbb{Q}_+^*$.

Application : Trouver les polygones réguliers du plan d'Argand-Cauchy dont tous les sommets ont des coordonnées rationnelles.

Exercice 15 : On fixe $n \in \mathbb{N}^*$ ($n \geq 2$), et on pose

$$\zeta = e^{2i\pi/n}, \quad \Delta_n = \prod_{\substack{k \neq l \\ (k,l) \in \llbracket 0, n-1 \rrbracket^2}} (\zeta^k - \zeta^l), \quad \Delta'_n = \prod_{\substack{k < l \\ (k,l) \in \llbracket 0, n-1 \rrbracket^2}} (\zeta^k - \zeta^l),$$

$$p_n = \prod_{k=1}^{n-1} \sin \frac{k\pi}{n} \quad (\text{égal à } \frac{n}{2^{n-1}} \text{ d'après l'exercice 11}).$$

a) Démontrer que $\Delta_n = 2^{n(n-1)} \mu_n (p_n)^n$, où μ_n est un nombre de module 1 à préciser, ce qui donnera une expression simple pour Δ_n .

b) Calculer Δ'_n .

Exercice 16 : (Polynômes cyclotomiques).

Pour $n \in \mathbb{N}^*$, soit \mathcal{P}_n l'ensemble des racines n -ièmes primitives de l'unité dans \mathbb{C} . On pose

$$\Phi_1(X) = X - 1 \quad \text{et} \quad \Phi_n(X) = \prod_{\zeta \in \mathcal{P}_n} (X - \zeta).$$

Φ_n est appelé le n -ième polynôme cyclotomique (son degré est évidemment l'indicateur d'Euler $\varphi(n)$).

a) Démontrer : $(\forall n \in \mathbb{N}^*) \quad X^n - 1 = \prod_{d|n} \Phi_d(X).$

En déduire, par récurrence, que $\Phi_n(X)$ a tous ses coefficients dans \mathbb{Z} .

b) Calculer explicitement les $\Phi_n(X)$ pour $n \leq 16$.

c) Démontrer que, pour p premier et α dans \mathbb{N}^* ,

$$\Phi_{p^\alpha}(X) = \sum_{k=0}^{p-1} X^{kp^{\alpha-1}}.$$

d) Calculer le terme constant de chaque $\Phi_n(X)$.

Remarque. Une généralisation hâtive pourrait laisser penser que les coefficients de Φ_n sont égaux à 0, 1, ou -1 . Il n'en est rien : par exemple le polynôme Φ_{105} qui est égal à $\frac{\Phi_{15}(X^7)}{\Phi_{15}(X)}$ admet le coefficient (-2) pour le terme en X^7 comme on pourra s'en assurer en effectuant la division.

Exercice 17 : Factoriser le polynôme $X^{10} + X^5 + 1$ sur \mathbb{C} , puis sur \mathbb{R} , enfin sur \mathbb{Q} .

Exercice 18 : Montrer qu'un polynôme de $\mathbb{Q}[X]$ qui est de degré 3 et qui n'admet pas de racine rationnelle (cf. Chap. IX pour la recherche de telles racines) est nécessairement irréductible sur \mathbb{Q} .

Exercice 19 : a) Factoriser $X^{2n} - 2X^n \cos \alpha + 1$ sur le corps des réels ($\alpha \in \mathbb{R}$ donné). Ecrire l'égalité obtenue en donnant à X la valeur 1.

b) Le polynôme $X^3 - X + 1$ est irréductible sur \mathbb{Q} . Le factoriser dans $(\mathbb{Z}/3\mathbb{Z})[X]$.

Exercice 20 : Soit $P_n \in \mathbb{C}[X]$ défini par

$$P_n(X) = \frac{1}{2i} \left[\left(1 + \frac{iX}{n} \right)^n - \left(1 - \frac{iX}{n} \right)^n \right].$$

Quel est son degré ? Montrer que ses coefficients sont dans \mathbb{R} , ses racines aussi. Le factoriser.

Exercice 21 : Le corps de base est \mathbb{C} . Pour quels $m \in \mathbb{N}^*$ le polynôme

$$F_m = 1 + X^m + X^{2m} + \dots + X^{(k-1)m}$$

est-il divisible par F_1 ?

Exercice 22 : Le corps de base est \mathbb{C} . On donne des entiers

$$m_1, m_2, \dots, m_k \in \mathbb{N}^* \quad (k \geq 1).$$

Utiliser les polynômes cyclotomiques, dont on admettra qu'ils sont *irréductibles sur \mathbb{Q}* pour déterminer le ppcm des $(X^{m_i} - 1)_{1 \leq i \leq k}$.

Exercice 23 : Pour tout $n \in \mathbb{N}$ on appelle *n -ième nombre de Fermat* et on note F_n l'entier $2^{2^n} + 1$. Pour tout $m \in \mathbb{N}^*$ on appelle *système complet mod (m)* toute partie \mathcal{R} de \mathbb{Z} telle que la restriction à \mathcal{R} de l'application canonique $\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}$ soit bijective de \mathcal{R} sur $\mathbb{Z}/m\mathbb{Z}$.

1^{re} Partie : Soit $p = F_n$ un nombre de Fermat qui soit **premier**, avec $n \geq 2$ (c'est sûrement le cas pour $n = 2, 3$ ou 4). On note g un générateur du groupe multiplicatif $\mathbb{Z}/p\mathbb{Z} \setminus \{0\}$, fixé une fois pour toutes, et appelé *entier primitif mod (p)* . On pose $\zeta = e^{2i\pi/p}$.

1) Soit s un entier, $s \in \llbracket 0, 2^n - 1 \rrbracket$, et soit $k \in \mathbb{Z}$.

a) Soit \mathcal{R} un système complet mod $(2^{2^n - s})$. Montrer que $\sum_{\lambda \in \mathcal{R}} \zeta^{(g^k + \lambda 2^s)}$ est indépendant du choix de \mathcal{R} . On notera cet élément $\Psi_{s,k}$ et on écrira :

$$\Psi_{s,k} = \sum_{\lambda \bmod (2^{2^n - s})} \zeta^{(g^k + \lambda 2^s)}.$$

b) Pour $k \in \mathbb{Z}$ calculer $\Psi_{0,k}$.

c) Pour $s \in \llbracket 1, 2^n - 1 \rrbracket$ et $k \in \mathbb{Z}$, prouver :

$$\Psi_{s,k} + \Psi_{s,k+2^s-1} = \Psi_{s-1,k}.$$

2) On donne des entiers $s \in \llbracket 0, 2^n - 1 \rrbracket$ et $k \in \mathbb{Z}$.

a) Démontrer :

$$\Psi_{s,k} + \Psi_{s,k+2^s-1} = \sum_{\mu \bmod (2^{2^n - s})} \left(\sum_{\lambda \bmod (2^{2^n - s})} \zeta^{K_\mu \times g^{\lambda 2^s}} \right)$$

où, pour chaque $\mu \in \mathbb{Z}$, on a posé $K_\mu = g^k + g^{k+(2\mu+1)2^s-1}$.

Indication : vérifier d'abord que

$$(\forall \alpha \in \mathbb{Z}, \forall l \in \mathbb{Z}) \quad \Psi_{s,l} = \sum_{\mu \bmod (2^{2^n - s})} \zeta^{(g^l + (\mu + \alpha)2^s)}.$$

b) Montrer : pour tout $\mu \in \mathbb{Z}$, $K_\mu \not\equiv 0 \bmod (p)$. En déduire qu'il existe un unique entier $t \in \llbracket 0, 2^{2^n} - 1 \rrbracket$ tel que $K_\mu \equiv g^t \bmod (p)$, qu'on notera $t = k_\mu$.

c) Exprimer simplement $\Psi_{s,k} \times \Psi_{s,k+2^s-1}$ à l'aide des nombres Ψ_{s,k_μ} .

2^e Partie : On se limite au cas où $p = F_2 = 17$, et on prend donc $\zeta = e^{2i\pi/17}$.

1) Vérifier que 3 est primitif mod (17). Dans toute la suite on utilise cette valeur de g .

2) a) Montrer que $(X - \Psi_{1,0})(X - \Psi_{1,1}) = X^2 + X - 4$. Calculer $\Psi_{1,0}$ et $\Psi_{1,1}$.

b) Montrer que $(X - \Psi_{2,0})(X - \Psi_{2,2}) = X^2 - \Psi_{1,0}X - 1$. Calculer $\Psi_{2,0}$ et $\Psi_{2,2}$.

c) Montrer que $(X - \Psi_{2,1})(X - \Psi_{2,3}) = X^2 - \Psi_{1,1}X - 1$. En déduire $\Psi_{2,1}$ et $\Psi_{2,3}$.

d) Montrer que $(X - \Psi_{3,0})(X - \Psi_{3,4}) = X^2 - \Psi_{2,0}X + \Psi_{2,1}$. Préciser $\Psi_{3,0}$ et $\Psi_{3,4}$.

3) On a donc $\Psi_{3,0} = \zeta + \zeta^{-1} = 2 \cos \frac{2\pi}{17}$ (le vérifier). Donner alors l'expression exacte de $2 \cos \frac{2\pi}{17}$ en n'utilisant qu'un nombre fini de racines carrées, portant au départ sur des nombres rationnels. *Réponse :*

$$2 \cos \frac{2\pi}{17} = \frac{1}{8}(-1 + \sqrt{17}) + \frac{1}{4} \sqrt{\frac{17 - \sqrt{17}}{2}} + \frac{1}{2} \sqrt{\frac{17 + 3\sqrt{17}}{4} - 2} \sqrt{\frac{17 + \sqrt{17}}{2} + \frac{1}{4}}$$

Exercice 24 : Déterminer tous les polynômes $P \in \mathbb{C}[X]$ tels que

$$P(X^2) = P(X)P(X-1).$$

Exercice 25 : Soit $A = X^3 + X - 2 \in \mathbb{C}[X]$ de racines $1, \alpha$ et β . Déterminer un polynôme B du second degré tel que $B(1) = 1$, $B(\alpha) = \beta$ et $B(\beta) = \alpha$. Montrer qu'alors $B(B(X)) - X$ est divisible par A .

§ VII.5 RACINES D'UN POLYNÔME. FORMULE DE TAYLOR

On prend comme corps de base un corps commutatif quelconque K .

Fonctions symétriques élémentaires

Avant de poursuivre l'étude des polynômes de $K[X]$ nous allons compléter les règles de calcul dans un anneau commutatif dont les premières ont été vues au Chapitre III. Considérons donc un anneau commutatif quelconque A , et soit t, a_1, a_2, \dots, a_n ($n \in \mathbb{N}^*$) des éléments de A . Nous allons développer le produit

$$(1) \quad P_n = (t - a_1)(t - a_2) \dots (t - a_n) = \prod_{i=1}^n (t - a_i).$$

La relation (3) du § III.2 montre que P_n est la somme de tous les termes $b_1 b_2 \dots b_n$ où, pour chaque $i \in \llbracket 1, n \rrbracket$, b_i est l'un des deux éléments t et $-a_i$ du i -ième facteur de P_n . Soit k le nombre de facteurs où $b_i = -a_i$ et soit J l'ensemble des indices de ces facteurs. Si $k = 0$ (c'est-à-dire si $J = \emptyset$), on a : $b_1 b_2 \dots b_n = t^n$; si $k \geq 1$ (c'est-à-dire si $J \neq \emptyset$), on a :

$$b_1 b_2 \dots b_n = (-1)^k t^{n-k} \prod_{i \in J} a_i.$$

Ordonnons alors *par paquets* la somme de tous les $b_1 b_2 \dots b_n$, en groupant dans un même paquet, pour chaque $k \in \llbracket 0, n \rrbracket$, ceux pour lesquels $\text{card}(J) = k$ (cf. théorème III.1.3). On obtient ainsi :

$$(2) \quad P_n = t^n + \sum_{k=1}^n (-1)^k t^{n-k} \left(\sum_{J \subset \llbracket 1, n \rrbracket, \text{card}(J)=k} \left(\prod_{i \in J} a_i \right) \right).$$

La donnée d'une k -partie J de $\llbracket 1, n \rrbracket$ équivaut à celle d'une suite strictement croissante à k termes à valeurs dans $\llbracket 1, n \rrbracket$, la suite (i_1, i_2, \dots, i_k) telle que $J = \{i_1, i_2, \dots, i_k\}$: on a alors

$$\prod_{i \in J} a_i = a_{i_1} a_{i_2} \dots a_{i_k}.$$

Introduisons les fonctions $\sigma_1, \sigma_2, \dots, \sigma_n$ de A^n dans A définies par :

$$(3) \quad \begin{aligned} \sigma_k(a_1, a_2, \dots, a_n) &= \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} a_{i_1} a_{i_2} \dots a_{i_k} \\ &= \sum_{J \subset \llbracket 1, n \rrbracket, \text{card}(J) = k} \left(\prod_{i \in J} a_i \right). \end{aligned}$$

Alors (2) peut s'écrire ainsi :

$$(4) \quad \boxed{P_n = t^n + \sum_{k=1}^n (-1)^k \sigma_k(a_1, a_2, \dots, a_n) t^{n-k}}$$

Notons en particulier qu'on a :

$$\sigma_1(a_1, \dots, a_n) = a_1 + a_2 + \dots + a_n \quad \text{et} \quad \sigma_n(a_1, \dots, a_n) = a_1 a_2 \dots a_n.$$

Les fonctions $\sigma_k : A^n \rightarrow A$ mises en évidence ci-dessus présentent un grand intérêt en Algèbre. On les appelle **fonctions symétriques élémentaires** de la variable $(a_1, a_2, \dots, a_n) \in A^n$. Cette appellation se justifie par le fait que, pour toute permutation $s \in \mathfrak{S}_n$, et tout $k \in \llbracket 1, n \rrbracket$, on a :

$$\sigma_k(a_{s(1)}, a_{s(2)}, \dots, a_{s(n)}) = \sigma_k(a_1, a_2, \dots, a_n)$$

qui est une conséquence de la définition des σ_k et que l'écriture

$$\sigma_k(a_1, a_2, \dots, a_n) = \sum_{J \subset \llbracket 1, n \rrbracket, \text{card}(J) = k} \left(\prod_{i \in J} a_i \right)$$

a l'avantage de rendre pratiquement évident, compte tenu des résultats du § III.1.

De manière générale nous dirons qu'une fonction $f : A^n \rightarrow A$ est *symétrique* ssi elle vérifie

$$f(a_{s(1)}, a_{s(2)}, \dots, a_{s(n)}) = f(a_1, a_2, \dots, a_n)$$

pour tout $s \in \mathfrak{S}_n$ et tout $(a_1, a_2, \dots, a_n) \in A^n$.

Relations entre coefficients et racines

Appliquons ce qui précède à l'anneau $K[X]$, en prenant a_1, a_2, \dots, a_n dans K , et en prenant pour t l'indéterminée X . Nous obtenons :

THÉORÈME VII.5.1

|| *Les coefficients du polynôme*

$$(X - a_1)(X - a_2) \dots (X - a_n) \in K[X]$$

sont donnés par

$$(X - a_1) \dots (X - a_n) = X^n + \sum_{k=1}^n (-1)^k \sigma_k(a_1, \dots, a_n) X^{n-k},$$

où les σ_k sont les fonctions symétriques élémentaires des a_i .

Si maintenant $F \in K[X]$ est un polynôme **dissocié** sur K , de degré $n \geq 1$, de coefficients c_i , noté

$$F = c_0 X^n + c_1 X^{n-1} + \dots + c_n,$$

appelons *numérotation* (ou *listage*) des racines de F , toute application $i \mapsto a_i$ de $\llbracket 1, n \rrbracket$ dans l'ensemble \mathcal{R} des racines de F , telle que

$$F = c_0(X - a_1)(X - a_2) \dots (X - a_n).$$

Cela exige d'une numérotation d'être une application $i \mapsto a_i$ telle que, pour tout $x \in \mathcal{R}$, le nombre des i tels que $a_i = x$ soit la multiplicité $\nu_x(F)$ de la racine x de F (autrement dit « chaque racine doit être comptée autant de fois que sa multiplicité »).

Le théorème VII.5.1 montre alors que, pour n'importe quelle numérotation $i \mapsto a_i$ des racines de F , on a :

$$(5) \quad \sigma_1 = \frac{-c_1}{c_0}, \dots, \sigma_k = (-1)^k \frac{c_k}{c_0}, \dots, \sigma_n = (-1)^n \frac{c_n}{c_0}$$

où l'on a noté en abrégé σ_k à la place de $\sigma_k(a_1, a_2, \dots, a_n)$ pour $1 \leq k \leq n$. Les relations (5) sont appelées **relations entre les coefficients et les racines** du polynôme dissocié F . Bien sûr, si K est algébriquement clos, elles pourront s'appliquer à tout polynôme F de degré ≥ 1 .

Ces relations confirment que les fonctions $\sigma_k : K^n \rightarrow K$ sont bien symétriques, puisque

$$F = c_0(X - a_{s(1)})(X - a_{s(2)}) \dots (X - a_{s(n)})$$

pour toute permutation $s \in \mathfrak{S}_n$. Mais elles constituent surtout un puissant outil de calcul : on s'en rendra compte en revenant à l'exemple 5 du § VI.7. Voici deux autres exemples simples :

Exemple 1 : Si $\theta \in \mathbb{R}$ et $n \in \mathbb{N}^*$, calculer le produit des racines du polynôme $F \in \mathbb{C}[X]$, $F = (X + 1)^n - e^{2in\theta}$, et en déduire

$$P_n(\theta) = \prod_{k=0}^{n-1} \sin \left(\theta + \frac{k\pi}{n} \right).$$

Le polynôme F admet n zéros distincts dans \mathbb{C} , ce qui permet de le factoriser en

$$F = \prod_{\zeta \in \mu_n} (X + 1 - \zeta e^{2i\theta}) = \prod_{k \in \llbracket 0, n-1 \rrbracket} \left(X - 2i e^{i\left(\theta + \frac{k\pi}{n}\right)} \sin\left(\theta + \frac{k\pi}{n}\right) \right).$$

Le produit des racines donne :

$$(-1)^n (1 - e^{2in\theta}) = 2^n i^n e^{in\theta} e^{i\pi\left(\frac{n-1}{2}\right)} P_n(\theta),$$

d'où
$$P_n(\theta) = \frac{1}{2^{n-1}} \sin n\theta.$$

En prenant $\lim_{\theta \rightarrow 0} \frac{P_n(\theta)}{\sin \theta}$ on retrouve $\prod_{k=1}^{n-1} \sin \frac{k\pi}{n} = \frac{n}{2^{n-1}}$ (cf. l'exercice 15 du § VII.4).

Exemple 2 : Soit p un nombre premier impair et K le corps $\mathbb{Z}/p\mathbb{Z}$. Nous avons vu (cf. l'exemple 8 du § VII.4) que dans $K[X]$, on a :

$$X^p - X = \prod_{a \in K} (X - a).$$

Après division par X , cela donne

$$X^{p-1} - \bar{1} = \prod_{a \in K^*} (X - a).$$

Le produit des racines donne alors $(-\bar{1})^{p-1} \prod_{a \in K^*} a = -\bar{1}$, et puisque $p-1$ est pair, $\prod_{k \in \llbracket 1, p-1 \rrbracket} \bar{k} = -\bar{1}$ dans $\mathbb{Z}/p\mathbb{Z}$, ou encore en revenant dans

$$\mathbb{Z} : (p-1)! \equiv -1 \pmod{p}$$

ce qui redémontre le théorème de Wilson déjà vu en Arithmétique (théorème IV.4.6).

Voici quelques autres applications un peu plus subtiles de la théorie des racines.

THÉORÈME VII.5.2

|| Soit G un sous-groupe **fini** du groupe multiplicatif K^* . Alors G est **cyclique**.

Démonstration :

Soit $n = \text{card}(G)$ et soit d un diviseur de n ; les éléments $x \in G$ dont l'ordre divise d sont ceux tels que $x \in G$ et $x^d = 1_K$, c'est-à-dire $x^d - 1 = 0$. Ce sont donc les racines de $X^d - 1_K$ qui appartiennent à G . Or $X^d - 1_K$ admet au plus d racines dans K . Donc G possède au plus d

l'ordre divise d . Il suffit alors de se reporter au corollaire de la proposition V.3.3 pour conclure que G est cyclique. ■

COROLLAIRE

|| Si K est un **corps fini**, le groupe multiplicatif K^* est cyclique. En particulier, si p est un nombre premier, le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$ est cyclique.

Dans le cas de $(\mathbb{Z}/p\mathbb{Z})^*$, on en déduit qu'il admet $\varphi(p-1)$ générateurs, φ désignant l'indicateur d'Euler. Chacun de ces générateurs est appelé un **nombre primitif modulo p** . Par exemple, 3 est primitif modulo 17 (cf. exercice 24 du § VII.4).

Exemple 3 : Critère d'Euler. Soit p un nombre **premier impair**. Etudions les éléments du groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$ qui sont des **carrés**. Bien évidemment ces carrés forment un sous-groupe Γ_2 . L'application $f : (\mathbb{Z}/p\mathbb{Z})^* \rightarrow (\mathbb{Z}/p\mathbb{Z})^*$, $x \mapsto x^2$ est un homomorphisme de groupes, dont le noyau $\text{Ker}(f)$ est $\{-1, +1\}$ car

$$x^2 = 1 \Leftrightarrow (x-1)(x+1) = 0.$$

Le corollaire 3 du théorème V.3.1 permet de conclure que $\text{card}(\Gamma_2) = \frac{p-1}{2}$. Les entiers $n \in \mathbb{Z}$ tels que $\bar{n} \in \Gamma_2$ (resp. $\bar{n} \notin \Gamma_2$) s'appellent les **résidus quadratiques mod (p)** (resp. les **non-résidus mod (p)**). Nous allons montrer que Γ_2 est l'ensemble des racines du polynôme $X^{\frac{p-1}{2}} - 1 \in \mathbb{Z}/p\mathbb{Z}[X]$ dans $\mathbb{Z}/p\mathbb{Z}$ (critère d'Euler). En effet si $x \in \Gamma_2$, il existe $y \in (\mathbb{Z}/p\mathbb{Z})^*$ tel que $x = y^2$ d'où $x^{\frac{p-1}{2}} = y^{p-1} = 1$ (on sait que l'ordre d'un élément divise le cardinal du groupe : théorème de Lagrange, corollaire 2 du théorème V.3.1). Mais d'après le théorème VII.4.2, le polynôme $X^{\frac{p-1}{2}} - 1$ ne peut pas admettre plus de $\frac{p-1}{2}$ racines dans $\mathbb{Z}/p\mathbb{Z}$. Or, on vient de lui en trouver $\frac{p-1}{2}$, donc ce sont les seules, et $X^{\frac{p-1}{2}} - 1$ est dissocié sur le corps $\mathbb{Z}/p\mathbb{Z}$, d'où le résultat annoncé. Cela permet d'écrire

$$X^{\frac{p-1}{2}} - 1 = \prod_{x \in \Gamma_2} (X - x).$$

Puisque

$$(X^{\frac{p-1}{2}} - 1)(X^{\frac{p-1}{2}} + 1) = X^{p-1} - 1 = \prod_{x \in (\mathbb{Z}/p\mathbb{Z})^*} (X - x),$$

comme on l'a vu dans l'exemple 2, il s'ensuit aussi que

$$X^{\frac{p-1}{2}} + 1 = \prod_{x \in (\mathbb{Z}/p\mathbb{Z})^* \setminus \Gamma_2} (X - x).$$

En résumé pour reconnaître qu'un entier x de \mathbb{Z}^* est un **résidu quadratique mod (p)** (resp. un non résidu), il suffit de vérifier que $x^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ (resp. $x^{\frac{p-1}{2}} \equiv -1 \pmod{p}$). Par exemple -1 est un résidu quadrati

condition que $(-1)^{\frac{p-1}{2}}$ soit $\equiv 1 \pmod{p}$, ce qui impose que $\frac{p-1}{2}$ soit pair, ou encore que $p \equiv 1 \pmod{4}$. Au contraire si $p \equiv 3 \pmod{4}$, (-1) est un non-résidu.

Dérivation d'un polynôme

Soit $F \in K[X]$, $F = \sum_{n \in \mathbb{N}} a_n X^n$, où la suite (a_n) est à support fini. On appelle polynôme dérivé de F (ou plus brièvement **dérivée** de F) le polynôme $\sum_{n \in \mathbb{N}} (n+1) a_{n+1} X^n$, et on le note F' , ou $F'(X)$, ou $\frac{dF}{dX}$. Le lecteur n'aura aucun mal à vérifier par des calculs élémentaires :

THÉOREME VII.5.3

|| L'application $K[X] \rightarrow K[X]$, $F \mapsto F'$ est **K-linéaire** et vérifie :

(I) $\forall (F, G) \in K[X] \times K[X] \quad (FG)' = F'G + FG'$.

(II) $\forall (F, G) \in K[X] \times K[X] \quad (F \circ G)' = (F' \circ G)G'$.

En fait, on déduit de (I) par récurrence $(F^m)' = mF^{m-1}F'$ ($m \in \mathbb{N}^*$), d'où (II) en utilisant la K -linéarité de $F \mapsto F \circ G$ lorsque G est fixé.

Si $K = \mathbb{R}$ (resp. $K = \mathbb{C}$), la dérivée de $F \in \mathbb{R}[X]$ définit la *fonction dérivée* (resp. la \mathbb{C} -dérivée), au sens de l'Analyse, de la fonction polynomiale définie par F (se reporter au tome 2).

L'itération de la dérivation se note ainsi : $F^{(0)} = F$, $F^{(1)} = F'$ et

$$F^{(n+1)} = (F^{(n)})' \quad \text{pour } n \geq 0.$$

On vérifie sans peine, par récurrence sur n , la *formule de Leibniz* qui donne la dérivée n -ième d'un produit :

$$\forall F \in K[X], \forall G \in K[X] \quad (FG)^{(n)} = \sum_{k=0}^n \binom{n}{k} F^{(n-k)} G^{(k)}.$$

Formule de Taylor

Dans ce qui suit, nous supposons le corps K de *caractéristique nulle*, de sorte que c'est une extension de \mathbb{Q} .

THÉOREME VII.5.4 (formule de Taylor ⁽¹⁾)

|| Soit $F \in K_n[X]$ et $h \in K$. On a :

(I) $F(X+h) = F(X) + hF'(X) + \dots + \frac{h^n}{n!} F^{(n)}(X) = \sum_{k=0}^n \frac{h^k}{k!} F^{(k)}(X).$

(II) $F(X+h) = F(h) + XF'(h) + \dots + \frac{X^n}{n!} F^{(n)}(h) = \sum_{k=0}^n \frac{X^k}{k!} F^{(k)}(h).$

⁽¹⁾ Brook Taylor (1685-1731), mathématicien anglais, dont le nom est développé en série entière de $f(x+h)$ étudié dans tout cours d'Anal

Démonstration :

Par K -linéarité, il suffit de prouver ces relations lorsque F est un monôme : $F = X^m$ ($m \in \mathbb{N}$). Or dans ce cas, en développant $(X + h)^m$ par la formule du binôme :

$$(X + h)^m = \sum_{k=0}^m \frac{m!}{k!(m-k)!} X^{m-k} h^k = \sum_{k=0}^n \frac{m!}{k!(m-k)!} X^k h^{m-k}$$

on remarque que

$$F^{(k)}(X) = \frac{m!}{(m-k)!} X^{m-k} \quad \text{pour tout } k \in \llbracket 0, m \rrbracket ;$$

mais du fait que K est de caractéristique nulle $\frac{m!}{k!(m-k)!}$ est le produit dans K des deux rationnels $\frac{1}{k!}$ et $\frac{m!}{(m-k)!}$, et les deux expressions de $(X + h)^m$ développées ci-dessus fournissent donc (I) et (II). ■

Dans la formule (I), si $h \neq 0$, les dérivées $F^{(k)}(X)$ sont non nulles tant que k reste $\leq \deg(F)$. Si d est ce degré on remarque que $F^{(d)}$ est une constante non nulle et que les dérivées suivantes sont toutes nulles. En fait, en caractéristique nulle, seuls les polynômes constants ont une dérivée nulle. Mais ce ne serait pas vrai en caractéristique $p > 0$ (par exemple sur $\mathbb{Z}/p\mathbb{Z}$ le polynôme X^p a une dérivée nulle et pourtant il prend les mêmes valeurs que X qui n'est pas constant !)

COROLLAIRE

$$\left\| \begin{array}{l} \text{Soit } F \in K[X] \setminus \{0\} \text{ et } a \in K ; \text{ la multiplicité } \nu_a(F) \text{ de } a \text{ dans } F \text{ est} \\ \text{l'entier } \nu \text{ tel que} \\ F^{(\nu)}(a) \neq 0 \text{ et } F^{(j)}(a) = 0 \text{ pour } j \leq \nu - 1. \end{array} \right.$$

Démonstration :

D'après la forme (II) de la formule de Taylor, dire que ν est la multiplicité de a dans F (i.e. $\text{val}_{X-a} F(X)$), mais dans l'automorphisme de la K -algèbre $K[X]$ défini par $F(X) \mapsto F(X - a)$, cela équivaut à $\text{val}_X (F(X + a)) = \nu$ est bien équivalent aux conditions énoncées dans le corollaire. ■

Si $\nu = 0$, a n'est pas racine de F ; si $\nu = 1$, a est racine simple ; si $\nu = 2$, a est racine double, etc...

Exercice 1 : Vérifier dans le détail que si A est un anneau commutatif, pour $n \in \mathbb{N}^*$ et $k \in \llbracket 1, n \rrbracket$, la fonction

$$\sigma_k : A^n \rightarrow A, (a_1, a_2, \dots, a_n) \mapsto \sum_{\substack{J \subset \llbracket 1, n \rrbracket \\ \text{card}(J) = k}} \left(\prod_{i \in J} a_i \right)$$

est symétrique.

Indication : si $s \in \mathfrak{S}_n$, s induit une bijection \hat{s} de l'ensemble des k -parties de $\llbracket 1, n \rrbracket$ sur lui-même. On utilise alors \hat{s} et les résultats des § III.1 et III.2.

Exercice 2 : Soit K un corps commutatif et $F \in K[X]$ un polynôme dissocié normalisé de degré $n \geq 1$, \mathcal{R} l'ensemble des racines de F , de sorte que $F = \prod_{a \in \mathcal{R}} (X - a)^{\nu_a}$, où $\nu_a = \nu_a(F)$ pour tout a .

Calculer le nombre de listages des racines de F . **Réponse :** $\frac{n!}{\prod_{a \in \mathcal{R}} (\nu_a!)}$.

Exercice 3 : Soit K un corps commutatif et a_1, a_2, \dots, a_n des éléments *distincts* dans K , et b_1, b_2, \dots, b_n des éléments de K . Le polynôme d'interpolation de Lagrange Λ tel que

$$\Lambda(a_i) = b_i \quad (1 \leq i \leq n)$$

peut s'écrire

$$\Lambda = \sum_{1 \leq i \leq n} b_i \frac{F(X)}{(X - a_i) F'(a_i)},$$

en désignant par F le polynôme $F(X) = \prod_{i \in \llbracket 1, n \rrbracket} (X - a_i)$.

Exercice 4 : Avec les notations de l'exercice 3, démontrer en choisissant des polynômes convenables :

$$\sum_{i=1}^n \frac{a_i^k}{F'(a_i)} = \begin{cases} 0 & \text{si } k < n-1 \\ 1 & \text{si } k = n-1 \end{cases} \quad (\text{Euler ; et aussi Jacobi}).$$

Exercice 5 : Le corps de base est \mathbb{C} . Tous les polynômes considérés ci-après sont *normalisés*. Soit $F \in \mathbb{C}[X]$ de degré $n \geq 2$. On note

$$G_k = \prod_{a \in \mathbb{C}, \nu_a(F)=k} (X - a).$$

On se propose de construire un algorithme donnant les G_k sans avoir besoin d'explicitier les racines de F .

a) Calculer d'abord $F_1 = \text{pgcd}(F, F')$ et $Q_1 = \frac{F}{F_1}$. Toutes les racines multiples

de F sont racines simples de Q_1 , mais il peut encore rester dans F_1 des racines multiples. Alors :

b) Définir par récurrence F_1, F_2, \dots et Q_1, Q_2, \dots par $F_{k+1} = \text{pgcd}(F_k, F'_k)$ et $Q_{k+1} = \frac{F_k}{F_{k+1}}$. Comparer alors G_k et Q_k/Q_{k+1} . Quand le processus s'arrête-t-il ?

Application : factoriser $X^7 - 2X^5 - X^4 + X^3 + 2X^2 - 1$.

Exercice 6 : Le corps de base est de caractéristique nulle. On donne des entiers $n \geq 2$, $p \geq 2$ et $(n_i)_{1 \leq i \leq p}$ tels que $\sum_{i=1}^p n_i = n$, ainsi que des éléments a_1, a_2, \dots, a_p de K distincts. Pour chaque $i \in \llbracket 1, p \rrbracket$, soit

$$\Phi_i(X) = \prod_{j \neq i} (X - a_j)^{n_j}.$$

a) Montrer que les $(X - a_i)^k \Phi_i$ ($0 \leq k \leq n_i - 1$) forment une base du K -ev U_i des $F \in K[X]$ tels que $F^{(\nu)}(a_j) = 0$ pour $j \neq i$ et $\nu \leq n_j - 1$.

b) On donne une famille $(b_{i,k})_{i \in \llbracket 1, p \rrbracket, k \in \llbracket 1, n_i - 1 \rrbracket}$ d'éléments de K .

Montrer qu'il existe un, et un seul, polynôme $F \in K[X]$ tel que $F^{(\nu)}(a_j) = b_{j,\nu}$ pour $j \in \llbracket 1, p \rrbracket$ et $\nu \leq n_j - 1$.

Que retrouve-t-on pour tous les n_i égaux à 1 ? Expliciter le résultat pour tous les n_i égaux à 2 (formule de Hermite).

Exercice 7 : On donne $n \in \mathbb{N}$. Trouver l'ordre de multiplicité de la racine

suivants de $\mathbb{R}[X]$:

- a) $X^{2n+1} - (2n+1)X^{n+1} + (2n+1)X^n - 1$.
 b) $X^{2n} - n^2 X^{n+1} + 2(n^2-1)X^n - n^2 X^{n-1} + 1$.

Exercice 8 : Déterminer un polynôme $P \in K[X]$ de degré n (le corps K est de caractéristique 0) sachant que $P(X) + a$ est divisible par $(X-b)^p$ et que $P(X) - a$ est divisible par $(X+b)^q$ où p et q sont des naturels de somme égale à $n+1$ et a, b des éléments de K .

Exemple : $n = 7$, 1 est racine d'ordre 4 de $P+1$ et -1 est racine d'ordre 4 de $P-1$.

Exercice 9 : Déterminer les polynômes P solutions des équations différentielles suivantes :

- a) $(1-X)P'(X) - P(X) = X^n$.
 b) $X(X-1)P' + P^2 - (2X+1)P + 2X = 0$.
 c) $XP'' - (X+m)P' + nP = 0$ (avec $m \in \mathbb{N}^*$, $n \in \mathbb{N}^*$).
 d) $X(X+1)P'' + (X+2)P' - P = 0$.
 e) $(1+X^2)P'' - (2X+1)P' + 2P = 0$.

Exercice 10 : Un calcul des polynômes de Tchebychev

On reprend les notations de l'exercice 8 du § VII.4.

a) Démontrer que, pour n fixé ≥ 2 , le polynôme T_n satisfait l'équation différentielle

$$(1-X^2)T_n''(X) - XT_n'(X) + n^2 T_n(X) = 0$$

(on suppose connues les dérivées des fonctions élémentaires utilisées en Analyse).

Déduire de là des relations de récurrence entre les coefficients de T_n .

Calculer $T_n(0)$ et $T_n'(0)$ et en déduire l'expression exacte de chaque coefficient.

b) Démontrer que le polynôme U_n vérifie l'équation différentielle

$$(1-X^2)U_n''(X) - 3XU_n'(X) + n(n+2)U_n(X) = 0.$$

Par une méthode analogue trouver l'expression complète de U_n .

Exercice 11 : Soit $n \in \mathbb{N}^*$ et $T_n \in \mathbb{C}[X]$ le n -ième polynôme de Tchebychev de 1^{re} espèce défini dans l'exercice 8 du § VII.4. On observera d'abord que $T_p \circ T_q = T_{pq}$ pour tous p et q dans \mathbb{N} .

a) Montrer que T_n et $-T_n$ sont les seules solutions polynômiales de l'équation différentielle

$$(1-X^2)Y'^2 - n^2(1-Y^2) = 0.$$

b) Montrer que les seuls polynômes $F \in \mathbb{C}[X]$ tels que $F \circ T_n = T_n \circ F$ sont les $(\varepsilon T_k)_{\varepsilon \in \{-1, +1\}, k \in \mathbb{N}}$.

Exercice 12 : On reprend les notations de l'exercice 15 du § VII.4 et on pose $F = X^n - 1$. Calculer directement Δ_n en remarquant que $\Delta_n = \prod_{\zeta \in \mu_n} F'(\zeta)$, et en déduire une nouvelle

preuve de

$$p_n = \prod_{k=1}^{n-1} \sin \frac{k\pi}{n} = \frac{n}{2^{n-1}}.$$

Exercice 13 : Soit $F = X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{Z}[X]$.

a) Démontrer que si G est un facteur normalisé de F dans $\mathbb{Q}[X]$, alors $G \in \mathbb{Z}[X]$.

b) Utiliser les relations entre coefficients et racines pour majorer les coefficients de G en fonction des a_i , G désignant toujours un facteur normalisé de F .

c) En déduire qu'un nombre fini d'essais permet de déterminer tous les facteurs irréductibles de F dans $\mathbb{Q}[X]$.

d) Application : Factoriser dans $\mathbb{Q}[X]$ le polynôme $X^7 + X + 1$.

Exercice 14 : Soit $F = a_0 X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{Z}[X]$. On se propose de chercher les facteurs du 1^{er} degré qui figurent dans la décomposition de F en facteurs irréductibles dans $\mathbb{Q}[X]$. D'après le théorème VII.4.2 cela revient à chercher les racines rationnelles.

de F . Montrer que l'on peut toujours supposer les a_i ($0 \leq i \leq n$) premiers entre eux dans leur ensemble.

Prouver ensuite que si $\frac{p}{q}$, avec p et q premiers entre eux, est une racine rationnelle de F , alors $p \mid a_n$ et $q \mid a_0$. En déduire qu'un nombre fini d'essais suffit pour trouver toutes les racines rationnelles de F .

Exemples :

$$F = X^6 + 3X^5 + 4X^4 + 3X^3 - 15X^2 - 16X + 20 ;$$

$$F = 6X^4 - 11X^3 - X^2 - 4 ;$$

$$F = 2X^3 + 12X^2 + 13X + 15 ; \quad F = 6X^3 - 13X^2 + 16X + 7 ;$$

$$F = 6X^5 + 11X^4 - X^3 + 5X - 6 ; \quad F = 33X^3 - 32X^2 + 62X - 35 .$$

Exercice 15 : Soit encore

$$F = a_0 X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{Z}[X]$$

qu'on se propose de factoriser dans $\mathbb{Z}[X]$. Supposons par exemple $F = F_1 F_2$ avec F_1 et F_2 non constants. Remplaçons chaque a_i par sa classe mod (p) (où p est un nombre premier qu'on choisit assez petit et tel que $p \nmid a_0$), ce qui remplace F par

$$\bar{F} \in \mathbb{Z}/p\mathbb{Z}[X], \text{ d'où } \bar{F} = \bar{F}_1 \times \bar{F}_2.$$

Or dans $\mathbb{Z}/p\mathbb{Z}[X]$ le nombre de polynômes à essayer comme diviseurs éventuels de \bar{F} est fini (on limite le degré à $\frac{1}{2} \deg(F)$ et on n'essaie que des polynômes irréductibles dans $\mathbb{Z}/p\mathbb{Z}[X]$). Le facteur \bar{F}_1 étant trouvé, on essaie pour F_1 un polynôme de même degré que \bar{F}_1 , avec des coefficients dont on connaît seulement les classes résiduelles mod (p) . On obtient des renseignements complémentaires en choisissant une nouvelle valeur de p ou en se reportant à l'exercice 13.

Exercice 16 : a) Déterminer deux polynômes P et Q de $\mathbb{R}[X]$ premiers entre eux tels que

$$P^2 + Q^2 = (X^2 + 1)^2 .$$

Réponse :

$$P = (X^2 - 1) \cos \alpha - 2X \sin \alpha \quad (\alpha \in \mathbb{R})$$

$$Q = (X^2 - 1) \sin \alpha + 2X \cos \alpha .$$

b) Calculer $P'^2 + Q'^2$. Montrer qu'on peut choisir deux « primitives » P_1 de P et Q_1 de Q vérifiant $P_1^2 + Q_1^2 = k(X^2 + 1)^3$ où $k \in \mathbb{R}$ est une constante à déterminer.

Exercice 17 : Montrer que le nombre 2 est un résidu quadratique pour tous les nombres premiers de la forme $8n \pm 1$ et un non-résidu pour tous les nombres premiers de la forme $8n \pm 3$.

Indication : Soit par exemple $p = 8n + 1$. Le polynôme $X^{8n} - 1$ se factorise en $(X^{4n} - 1)(X^{4n} + 1)$. Dans $\mathbb{Z}/p\mathbb{Z}[X]$, $X^{4n} + 1$ admet $4n$ racines. Si x est l'une d'elles $(x^{2n} + 1)^2 - 2x^{2n} = 0$, d'où $t^2 - 2u^2 = 0$, d'où l'on conclura que 2 est résidu quadratique de p . On pourra adapter la méthode aux autres cas (Legendre).

Exercice 18 : (loi de réciprocité quadratique, méthode des racines de l'unité).

Soit p et q deux nombres premiers impairs distincts, et $\zeta = e^{2i\pi/p}$.

a) Montrer que $\prod_{k=1}^{p-1} (1 - \zeta^{2k}) = p$ et en déduire

$$p = (-1)^{\frac{p-1}{2}} \prod_{k=1}^{\frac{p-1}{2}} (\zeta^k - \zeta^{-k})^2 .$$

b) On considère l'ensemble des complexes de la forme

$$c_0 + c_1 \zeta + \dots + c_{p-2} \zeta^{p-2} ,$$

avec $c_i \in \mathbb{Z}$ pour tout i . Vérifier qu'on obtient un anneau A . Soit \mathfrak{q} l'idéal

Montrer que, pour tout $k \in \mathbb{N}$,

$$(\zeta^k - \zeta^{-k})^q - (\zeta^{qk} - \zeta^{-qk}) \in \mathfrak{q}, \text{ et que } A \cap \mathbb{Q} = \mathbb{Z} \text{ (cf. e)}).$$

En déduire, à l'aide de a), que si

$$Q = \prod_{k=1}^{\frac{p-1}{2}} \frac{\zeta^{qk} - \zeta^{-qk}}{\zeta^k - \zeta^{-k}}, \text{ alors } p^{\frac{q-1}{2}} - (-1)^{\frac{p-1}{2} \frac{q-1}{2}} Q \in \mathfrak{q}.$$

c) Pour $a \in \left[1, \frac{p-1}{2}\right]$, soit ρ_a l'unique élément de $\left[1 - \frac{p-1}{2}, \frac{p-1}{2}\right]$ tel que $qa = bp + \rho_a$. Montrer que les ρ_a sont tous distincts. Soit

$$\mathcal{E}_+ = \left\{ a \in \left[1, \frac{p-1}{2}\right] \mid \rho_a \geq 1 \right\} \text{ et } \mathcal{E}_- = \left\{ a \in \left[1, \frac{p-1}{2}\right] \mid \rho_a \leq -1 \right\}.$$

Alors $\mathcal{E}_+ \cup \mathcal{E}_- = \left[1, \frac{p-1}{2}\right]$. En déduire $Q = (-1)^\mu$ où $\mu = \text{card}(\mathcal{E}_-)$.

d) On pose $\left(\frac{p}{q}\right) = +1$ (resp. -1) ssi l'image de p dans $\mathbb{Z}/q\mathbb{Z}$ appartient au groupe des carrés dans $(\mathbb{Z}/q\mathbb{Z})^*$ (resp. n'appartient pas à ce groupe) et on définit de même $\left(\frac{q}{p}\right)$ (on les appelle *symboles de Legendre*).

Démontrer d'abord que $p^{\frac{q-1}{2}} - \left(\frac{p}{q}\right) \in q\mathbb{Z}$ en utilisant l'exemple 3 du § VII.5.

Prouver ensuite par la même méthode que $\left(\frac{q}{p}\right) = (-1)^\mu$.

En déduire que $\left(\frac{p}{q}\right) \left(\frac{q}{p}\right) - (-1)^{\frac{p-1}{2} \frac{q-1}{2}} \in \mathfrak{q} \cap \mathbb{Z}$.

e) Le polynôme $X^{p-1} + X^{p-2} + \dots + 1$ est irréductible dans $\mathbb{Q}(X)$ (cf. exercice 4 du § VII.3). En déduire que $1, \zeta, \zeta^2, \dots, \zeta^{p-2}$ sont linéairement indépendants sur \mathbb{Q} , puis, que $\mathfrak{q} \cap \mathbb{Z} = q\mathbb{Z}$. En déduire enfin la *loi de réciprocité quadratique* (démontrée d'abord par Legendre presque rigoureusement (1798), puis entièrement par Gauss (1801)) qui s'écrit

$$\left(\frac{p}{q}\right) \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \frac{q-1}{2}}.$$

Remarque. On pourra se reporter au problème posé à l'ENS d'Ulm en 1978 pour une autre démonstration, parmi beaucoup d'autres de cette célèbre loi.

Exercice 19 : Anneaux $\mathbb{Z}/n\mathbb{Z}$ dont le groupe des éléments inversibles est cyclique.

a) Soit k un entier naturel de la forme $k = p^\beta q$, avec p premier, $\beta \geq 1$, q premier avec p , tel que $k < p^\alpha$. Démontrer :

$$(1) \quad \left(\frac{p^\alpha}{k}\right) \equiv 0 \pmod{(p^{\alpha-\beta})}.$$

b) Si $m \in \mathbb{Z}$, on note \overline{m} la classe de $m \pmod{p}$, et $\overline{\overline{m}}$ la classe de $m \pmod{p^\alpha}$. Soit $G(p)$ le groupe $(\mathbb{Z}/p\mathbb{Z})^*$ et $G(p^\alpha)$ le groupe multiplicatif des éléments inversibles de l'anneau $\mathbb{Z}/p^\alpha\mathbb{Z}$. Soit \overline{g} un générateur de $G(p)$, où $g \in \mathbb{Z}$ est choisi convenablement (cf. corollaire du théorème VII.5.2). On sait que

$$\text{card}(G(p^\alpha)) = \varphi(p^\alpha) = p^{\alpha-1}(p-1).$$

Prouver qu'il existe $k \in \mathbb{Z}$ tel que l'entier $\theta = g + kp$ vérifie (2) $\theta^{p-1} \not\equiv 1 \pmod{p^2}$; k étant ainsi choisi, démontrer que $\overline{\overline{\theta}}^{(p-1)p^{\alpha-2}} \not\equiv 1 \pmod{p^\alpha}$ en utilisant (1).

Prouver que l'ordre de $\overline{\overline{\theta}}$ dans le groupe $G(p^\alpha)$ est multiple de $p-1$, puis que $G(p^\alpha)$ est cyclique, engendré par $\overline{\overline{\theta}}$, que l'on notera ρ pour abrégé.

c) Pour tout $a \in G(p^\alpha)$ soit s_a la bijection de $\mathbb{Z}/p^\alpha\mathbb{Z}$ dans lui-même, $x \mapsto s_a(x) = ax$. Déterminer la décomposition en cycles de s_p et en déduire la signature $\varepsilon(s)$

d) Pour tout entier impair a , montrer (3) $a^2 \equiv 1 \pmod{(2^3)}$ et $a^{(2^j)} \equiv 1 \pmod{(2^4)}$. En déduire (4) $a^{2^a-2} \equiv 1 \pmod{(2^a)}$.

Si on désigne par $G(n)$ le groupe multiplicatif des éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$, pour $n \in \mathbb{N}^*$, déterminer les entiers $\beta \geq 1$ tels que $G(2^\beta)$ soit cyclique.

e) Soit G_1 et G_2 deux groupes cycliques finis, de cardinaux respectifs n_1 et n_2 . Donner une condition nécessaire et suffisante sur n_1 et n_2 pour que le groupe $G_1 \times G_2$ soit cyclique.

f) Soit $n \in \mathbb{N}^*$, et $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ sa décomposition en facteurs premiers ($\alpha_i \geq 1$ pour $1 \leq i \leq r$). Montrer que l'application naturelle $\mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \dots \times \mathbb{Z}/p_r^{\alpha_r}\mathbb{Z}$ est un isomorphisme d'anneaux qui induit un isomorphisme de groupes

$$G(n) \cong G(p_1^{\alpha_1}) \times \dots \times G(p_r^{\alpha_r}).$$

Déduire de cette étude tous les $n \in \mathbb{N}^*$ tels que le groupe $G(n)$ soit cyclique.

Exercice 20 : Sommes de deux carrés.

a) Vérifier d'abord que si a, b, α, β sont dans \mathbb{Z} , le produit $(a^2 + b^2)(\alpha^2 + \beta^2)$ est la somme de deux carrés d'entiers rationnels (de combien de façons différentes en général?).

b) On se propose alors de démontrer que *tout nombre premier impair de la forme $p = 4n + 1$ est la somme de deux carrés*. On sait déjà (cf. critère d'Euler, exemple 3) que -1 est un résidu quadratique de p , c'est-à-dire qu'il existe $t \in \mathbb{N}^*$ tel que p divise $1 + t^2$. On peut même choisir $t < p$ tel que $1 + t^2 = mp$ avec $0 < m < p$, ou mieux $t \in \left[1, \frac{p-1}{2}\right]$ tel que $1 + t^2 = mp$ de sorte que $0 < m < \frac{p}{4}$.

Démontrons plus généralement que si un nombre premier divise une somme de deux carrés premiers entre eux, il est lui-même la somme de deux carrés. Partons de l'égalité (1) $a^2 + b^2 = mp$ avec $|a| < \frac{p}{2}$, $|b| < \frac{p}{2}$, de sorte que $m < \frac{p}{2}$. En divisant a et b par m on peut choisir les restes a' et b' dans \mathbb{Z} tels que $|a'| < \frac{m}{2}$ et $|b'| < \frac{m}{2}$, ce qui conduit à (2) $a'^2 + b'^2 = mm'$ avec $m' < \frac{m}{2}$. En multipliant membre à membre (1) et (2) et en utilisant le a) on s'aperçoit que l'on arrive à une égalité de la forme (3) $A^2 + B^2 = pm'$ analogue à (1) mais avec $m' < \frac{m}{2}$. La méthode de « descente infinie » de Fermat permet alors de conclure que $p = X^2 + Y^2$.

c) Il est évident que le nombre premier 2 est égal à $1^2 + 1^2$. Montrer qu'en revanche aucun nombre premier de la forme $4n + 3$ ne peut être une somme de deux carrés.

d) Exprimer une condition nécessaire et suffisante pour qu'un élément de \mathbb{N} soit somme de deux carrés.

e) Montrer que dans un corps fini tout élément est somme de deux carrés.

Exercice 21 : Soit p un nombre premier de la forme $4n + 3$. Montrer que dans le corps $F_p = \mathbb{Z}/p\mathbb{Z}$ la relation $a^2 + b^2 = 0$ implique $a = b = 0$. En imitant la façon dont \mathbb{C} a été construit à partir de \mathbb{R} au chapitre VI, montrer qu'on peut construire un corps noté F_{p^2} tel que $F_p \subset F_{p^2}$ et dans lequel $-\bar{1}$ est un carré.

Exercice 22 : Soit $P \in \mathbb{Z}[X]$ un polynôme de degré ≥ 1 et $n \in \mathbb{Z}$. On pose $m = P(n)$.

a) Montrer que pour tout $k \in \mathbb{Z}$, $P(n + km)$ est divisible par m .

b) Montrer qu'il n'existe pas de polynôme non constant $P \in \mathbb{Z}[X]$ tel que pour tout $n \in \mathbb{Z}$, $P(n)$ soit premier.

c) Vérifier cependant que $n^2 - n + 41$ prend des valeurs premières pour $0 \leq n \leq 40$ et $n^2 - 79n + 1601$ pour $0 \leq n \leq 79$.

Exercice 23 : (Polynômes d'Euler) : K est de caractéristique nulle. Montrer qu'il existe un polynôme $P \in K[X]$ et un seul vérifiant

$$(1) \quad P(X+1) + P(X) = 2X^n \quad (n \in \mathbb{N}).$$

Soit E_n le polynôme correspondant. Trouver une relation simple entre E'_n et E_{n-1} . Développer $E_n(X+h)$ sous la forme $\sum a_p E_p(X)$ et en déduire une relation de récurrence entre E_n et les précédents. Expliciter E_n pour $n \in \llbracket 0, 5 \rrbracket$. Démontrer que

$$E_n(1-X) = (-1)^n E_n(X).$$

§ VII.6 FACTORISATION DANS $\mathbb{R}[X]$

On sait que, contrairement à \mathbb{C} , le corps \mathbb{R} n'est pas algébriquement clos. Considérons alors la sous- \mathbb{R} -algèbre $\mathbb{R}[X]$ de la \mathbb{C} -algèbre $\mathbb{C}[X]$ et proposons-nous de déterminer *tous les polynômes irréductibles* dans $\mathbb{R}[X]$. Commençons par observer que, si

$$F = X^2 - 2\alpha X + \beta \in \mathbb{R}[X] \quad \text{avec} \quad \alpha^2 - \beta < 0$$

(c'est-à-dire si le trinôme F n'a *pas de racine réelle*), alors F est *irréductible* dans $\mathbb{R}[X]$. En effet si F n'était pas irréductible dans $\mathbb{R}[X]$, il aurait un facteur de degré 1, donc une racine réelle.

A tout polynôme $F = \sum_{n \in \mathbb{N}} a_n X^n \in \mathbb{C}[X]$, associons $\bar{F} = \sum_{n \in \mathbb{N}} \bar{a}_n X^n$.

L'application $F \rightarrow \bar{F}$ est un automorphisme de la \mathbb{R} -algèbre $\mathbb{C}[X]$ (mais pas de la \mathbb{C} -algèbre !), et $F = \bar{F}$ ssi $F \in \mathbb{R}[X]$. En outre il est clair que $(\forall z \in \mathbb{C}) \bar{F}(\bar{z}) = \overline{F(z)}$ (notation qui désigne évidemment le *conjugué* du nombre complexe $F(z)$). En particulier $F(z) = 0$ ssi $\bar{F}(\bar{z}) = 0$.

PROPOSITION VII.6.1

|| Soit $F \in \mathbb{C}[X]$ un polynôme non constant ; pour tout $a \in \mathbb{C}$, on a :
|| $\text{val}_{X-a}(F) = \text{val}_{X-a}(\bar{F})$.

Démonstration :

Donnons-nous $a \in \mathbb{C}$ et soit $\nu_a = \text{val}_{X-a}(F)$ la multiplicité de a dans F . On a :

$$F(X) = (X-a)^{\nu_a} G(X) \quad \text{avec} \quad G \in \mathbb{C}[X], \quad G(a) \neq 0.$$

Puisque $P \mapsto \bar{P}$ est un automorphisme de l'anneau $\mathbb{C}[X]$, on en déduit (compte tenu de $\overline{(X-a)} = X-a$) : $\bar{F}(X) = (X-a)^{\nu_a} \bar{G}(X)$. Mais $\bar{G}(\bar{a})$ est le conjugué de $G(a)$, donc $\bar{G}(\bar{a}) \neq 0$, d'où $\nu_a = \text{val}_{X-a}(\bar{F})$. ■

THÉORÈME VII.6.1

|| Les polynômes irréductibles et normalisés dans $\mathbb{R}[X]$ sont les
|| suivants : les $X-a$ pour $a \in \mathbb{R}$, les $X^2 - 2\alpha X + \beta$ avec
|| $(\alpha, \beta) \in \mathbb{R}^2$ et $\alpha^2 - \beta < 0$.

Démonstration :

On sait déjà que ces polynômes sont irréductibles dans $\mathbb{R}[X]$. Il suffit donc de voir que tout $F \in \mathbb{R}[X]$ est produit de facteurs irréductibles de ce type, et d'une constante réelle. Soit donc $F \in \mathbb{R}[X]$ non constant. Notons H l'ensemble $\{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$; F est dissocié dans \mathbb{C} d'après le théorème de d'Alembert. La proposition VII.6.1 permet d'écrire sa décomposition sous la forme

$$(1) \quad F = \lambda \left[\prod_{a \in \mathbb{R}} (X - a)^{\nu_a} \right] \left[\prod_{b \in H} (X - b)^{\nu_b} (X - \bar{b})^{\nu_b} \right],$$

avec $\nu_z = \text{val}_{X-z}(F)$ pour tout $z \in \mathbb{C}$.

Mais si $b \in H$ on a $b = \alpha + i\beta$, $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}_+^*$, et

$$(X - b)^{\nu_b} (X - \bar{b})^{\nu_b} = [(X - b)(X - \bar{b})]^{\nu_b} = (X^2 - 2\alpha X + \alpha^2 + \beta^2)^{\nu_b},$$

et le polynôme $X^2 - 2\alpha X + \alpha^2 + \beta^2$ est irréductible dans $\mathbb{R}[X]$, car il n'a pas de racine réelle. On remarquera que $\prod_{a \in \mathbb{R}} (X - a)^{\nu_a}$ est la **partie dissociée** de F dans $\mathbb{R}[X]$. ■

On peut ajouter qu'il résulte de la démonstration précédente que dans le cas où le degré du polynôme $F \in \mathbb{R}[X]$ est **impair**, la partie dissociée de F est elle-même de degré impair et donc non constante, ce qui prouve que tout polynôme de $\mathbb{R}[X]$ de degré impair *admet au moins une racine réelle*. Mais ce résultat se démontre de façon bien plus naturelle en Analyse en se servant du fait que la fonction polynôme F est une fonction **continue** de \mathbb{R} dans \mathbb{R} , ce qui permet d'utiliser le théorème des valeurs intermédiaires. On peut même prendre ce résultat simple comme point de départ d'une démonstration du théorème de d'Alembert, en n'utilisant plus à partir de là que des considérations algébriques.

La pratique de la décomposition dans $\mathbb{R}[X]$ de $F \in \mathbb{R}[X]$ n'est pas toujours facile. Bien sûr on peut, en suivant la démonstration du théorème VII.6.1, commencer par décomposer F sur \mathbb{C} , puis regrouper les facteurs conjugués, mais ce n'est pas nécessairement la meilleure méthode.

Exemple 1 : Soit $F = X^4 + 2\alpha X^2 + \beta$ un polynôme de $\mathbb{R}[X]$.

Si $\alpha^2 - \beta \geq 0$ on pourra partir de $U^2 + 2\alpha U + \beta$ qui se factorise en $(U - a)(U - b)$ avec $(a, b) \in \mathbb{R}^2$, d'où $F = (X^2 - a)(X^2 - b)$ et on termine selon les signes de a et b . Si $\alpha^2 - \beta < 0$, on a $\beta > 0$ et $\sqrt{\beta} > |\alpha|$, ce qui incite à mettre F sous la forme

$$\begin{aligned} F &= (X^2 + \sqrt{\beta})^2 - 2(\sqrt{\beta} - \alpha)X^2 = \\ &= (X^2 + \sqrt{\gamma}X + \sqrt{\beta})(X^2 - \sqrt{\gamma}X + \sqrt{\beta}) \end{aligned}$$

avec $\gamma = \sqrt{\beta} - \alpha > 0$.

Exemple 2 : Soit à factoriser $F = X^8 + 1$, élément de $\mathbb{R}[X]$, sur le corps des réels. On a

$$F = (X^4 + 1)^2 - 2X^4 = (X^4 + \sqrt{2}X^2 + 1)(X^4 - \sqrt{2}X^2 + 1)$$

et en utilisant la méthode précédente (en prenant $\varepsilon \in \{-1, +1\}$ pour économiser des calculs) :

$$\begin{aligned} X^4 + \varepsilon \sqrt{2}X^2 + 1 &= (X^2 + 1)^2 - (2 - \varepsilon \sqrt{2})X^2 = \\ &= (X^2 + \sqrt{2 - \varepsilon \sqrt{2}}X + 1)(X^2 - \sqrt{2 - \varepsilon \sqrt{2}}X + 1) \end{aligned}$$

d'où

$$\begin{aligned} X^8 + 1 &= (X^2 + \sqrt{2 + \sqrt{2}}X + 1)(X^2 - \sqrt{2 + \sqrt{2}}X + 1) \times \\ &\times (X^2 + \sqrt{2 - \sqrt{2}}X + 1)(X^2 - \sqrt{2 - \sqrt{2}}X + 1). \end{aligned}$$

Exercice 1 : Décomposer en facteurs irréductibles dans $\mathbb{R}[X]$ les polynômes suivants :

a) $F = X^6 + 1$; b) $F = X^{12} + 1$; c) $F = X^{24} + 1$; d) $X^8 + X^4 + 1$.

Exercice 2 : Décomposer dans $\mathbb{R}[X]$ le polynôme $F = X^{2^n} + 1$, où $n \in \mathbb{N}^*$, et en déduire des expressions « par radicaux » des nombres $\cos \frac{k\pi}{2^n}$, $0 \leq k \leq 2^{n-1} - 1$.

Exercice 3 : On donne un entier $n \geq 3$ et on cherche la forme d'un polynôme normalisé du 3^e degré $X^3 + aX^2 + bX + c \in \mathbb{R}[X]$ qui divise $X^n - 1$.

a) Montrer que si n est impair les polynômes cherchés sont $X^3 - uX^2 + uX - 1$, avec

$$u = 1 + 2 \cos \frac{2k\pi}{n} \quad \text{et} \quad k \in \left[\left[1, \frac{n-1}{2} \right] \right].$$

b) Montrer que si n est pair, les polynômes cherchés sont $X^3 - vX^2 + \varepsilon vX - \varepsilon$ avec

$$\varepsilon \in \{-1, +1\} \quad \text{et} \quad v = \varepsilon + 2 \cos \frac{2k\pi}{n}, \quad k \in \left[\left[1, \frac{n}{2} - 1 \right] \right].$$

Exercice 4 : Soit $(a, b) \in \mathbb{R}^2 \setminus \{0, 0\}$ et $\lambda \in \mathbb{R}^*$. On se propose de chercher la décomposition dans $\mathbb{R}[X]$ du polynôme $F = (X^2 + a^2)^2 + \lambda^2(X + b)^2$. Soit $t \in \mathbb{R}$; on pose

$$G_t(X) = (X^2 + a^2 + t)^2 - F(X).$$

Chercher $t \neq 0$ pour que le trinôme G_t ait un discriminant nul dans $\mathbb{R}[X]$, et en déduire la factorisation cherchée (méthode de Ferrari).

Exercice 5 : Factoriser sur \mathbb{C} et sur \mathbb{R} le polynôme $X^5 - 1$ et en déduire la valeur « par radicaux » de $\cos \frac{2k\pi}{5}$ et de $\sin \frac{2k\pi}{5}$ où $k \in \mathbb{Z}$.

Exercice 6 : Factoriser dans $\mathbb{R}[X]$: $16X^5 - 20X^3 + 5X - 1$ et

$$X^5 - 13X^4 + 67X^3 - 171X^2 + 216X - 108$$

sachant qu'ils admettent des racines multiples.

Exercice 7 : Soit $F \in \mathbb{R}[X]$; on identifie F avec la fonction polynomiale $\mathbb{R} \rightarrow \mathbb{R}$ qu'il définit :

a) Pour que $F(\mathbb{R}) \subset \mathbb{R}_+$, il faut et il suffit qu'il existe A et B dans $\mathbb{R}[X]$ tels que $F = A^2 + B^2$.

b) Supposons de plus que F est **pair** ; si $F(\mathbb{R}) \subset \mathbb{R}_+$, alors il existe A et B dans $\mathbb{R}[X]$ tels que

$$F = A^2(X^2) + X^2 B^2(X^2).$$

c) Si $F(\mathbb{R}_+) \subset \mathbb{R}_+$, il existe A et B dans $\mathbb{R}[X]$ tels que

$$F = A^2(X) + XB^2(X).$$

Exercice 8 : Décomposer $F = X^4 + X^3 + X^2 + X + 1$ dans $\mathbb{R}[X]$ (cf. exercice 5, utiliser $X + \frac{1}{X}$) et en déduire que F est irréductible dans $\mathbb{Q}[X]$ sans avoir besoin d'utiliser les résultats de l'exercice 4 du § VII.3.

Exercice 9 : Pour quelles valeurs de l'entier n le nombre $n^4 + 4$ est-il premier ? Même question pour $n^4 + 4^n$.

§ VII.7 CONGRUENCES DANS $K[X]$. ANNEAUX QUOTIENTS

Congruences entre polynômes

Soit K un corps commutatif et $P \in K[X]$. Reprenant la terminologie du § IV.1, nous dirons que deux polynômes A et B dans $K[X]$ sont **congrus modulo P** , ce que nous écrirons

$$(1) \quad A \equiv B \pmod{P} \quad \text{ssi } A - B \text{ est multiple de } P.$$

Une relation du type (1) s'appelle une **congruence modulo P** (si $P = 0$, la relation (1) est l'égalité dans $K[X]$). Si P est de degré 0 les éléments de $K[X]$ sont tous congrus entre eux, ce qui ne présente aucun intérêt).

Il est clair que la congruence mod (P) est aussi celle modulo (λP) pour tout $\lambda \in K^*$. On peut écrire (1) sous la forme $A - B \in \mathfrak{p}$, où \mathfrak{p} est l'idéal engendré par P dans $K[X]$, et pour cette raison on peut aussi l'appeler **congruence modulo \mathfrak{p}** . La vérification des propriétés qui suivent est immédiate.

(CG1) D'abord, pour P fixé, la congruence mod (P) est une **relation d'équivalence** dans $K[X]$. C'est une conséquence du fait que \mathfrak{p} est un *sous-groupe additif* de $K[X]$. Les classes de cette relation d'équivalence sont appelées **classes de congruence mod (P)** .

(CG2) Pour A_1, A_2, B_1, B_2 dans $K[X]$:

$$(A_1 \equiv A_2 \pmod{P} \text{ et } B_1 \equiv B_2 \pmod{P}) \Rightarrow (A_1 + B_1 \equiv A_2 + B_2 \pmod{P}).$$

C'est la compatibilité de la congruence mod (P) avec l'addition.

(CG3) Pour A_1, A_2, B_1, B_2 dans $K[X]$:

$$(A_1 \equiv A_2 \pmod{P} \text{ et } B_1 \equiv B_2 \pmod{P}) \Rightarrow (A_1 B_1 \equiv A_2 B_2 \pmod{P}).$$

C'est la compatibilité de la congruence mod (P) avec la multiplication (écrire $A_2 = A_1 + PM$, etc...).

(CG4) Pour tout $k \in \mathbb{N}$ ($A \equiv B \text{ mod } (P)$) \Rightarrow ($A^k \equiv B^k \text{ mod } (P)$) (par récurrence à partir de CG3).

(CG5) Pour tout $C \in K[X] \setminus \{0\}$,

$$A \equiv B \text{ mod } (P) \Leftrightarrow AC \equiv BC \text{ mod } (CP).$$

(CG6) Si $A \equiv B \text{ mod } (P)$, alors $A \equiv B \text{ mod } (Q)$ pour tout diviseur Q de P .

(CG7) Si $Q \neq 0$ divise A , B et P , alors $A \equiv B \text{ mod } (P)$ implique $\frac{A}{Q} \equiv \frac{B}{Q} \text{ mod } \left(\frac{P}{Q}\right)$.

(CG8) Si $A \equiv B \text{ mod } (P)$, alors pour toute racine α de P dans K , on a : $A(\alpha) = B(\alpha)$, propriété nouvelle par rapport aux congruences dans \mathbb{Z} .

Soit γ une classe de congruence mod (P) et A_0 un élément fixé dans γ . Par définition,

$$(2) \quad \gamma = \{A_0 + MP\}_{M \in K[X]}.$$

Si P est de degré ≥ 1 , la relation (2) montre que le théorème de la division euclidienne peut s'énoncer ainsi :

THÉORÈME VII.7.1

|| Soit $P \in K[X]$ un polynôme de degré $n \geq 1$. Dans chaque classe de congruence $\gamma \text{ mod } (P)$ il y a un, et un seul, élément de $K_{n-1}[X]$.

En conséquence, si nous notons \mathcal{A}_P l'ensemble des classes de congruence de $K[X] \text{ mod } (P)$, l'application $f : \mathcal{A}_P \rightarrow K_{n-1}[X]$ qui associe à chaque classe $\gamma \in \mathcal{A}_P$ l'unique élément de $\gamma \cap K_{n-1}[X]$ est **bijective**.

Si $A \in \gamma$, le polynôme $f(\gamma)$ est évidemment le **reste dans la division euclidienne de A par P** .

Les propriétés (CG1) à (CG8) sont à la base d'un *calcul des congruences* très analogue à celui développé dans \mathbb{Z} au § IV.1. En les combinant aux résultats du § VII.4 on obtient un puissant outil de calcul.

Exemple 1 : On recherche pour quelles valeurs de l'entier naturel n le polynôme de $K[X]$: $F_n = X^{2^n} + X^n + 1$ est divisible par $F_1 = X^2 + X + 1$.

On peut écrire $1 \equiv 1 \text{ mod } (F_1)$, $X \equiv X \text{ mod } (F_1)$, puis

$$X^2 \equiv -X - 1 \text{ mod } (F_1), X^3 \equiv -X^2 - X \equiv 1 \text{ mod } (F_1)$$

et une périodicité apparaît clairement. Formons le tableau suivant, où la première ligne donne la classe de n modulo 3 dans \mathbb{Z} , et les suivantes les classes correspondantes de X^n et $X^{2n} \bmod (F_1)$

n	$\bar{0}$	$\bar{1}$	$\bar{2}$
X^n	1	X	$-X - 1$
X^{2n}	1	$-X - 1$	X

On voit bien que $(\forall n \in \mathbb{N})$, si $n \equiv 0 \bmod (3)$,

$$1 + X^n + X^{2n} \equiv 3 \bmod (F_1)$$

et si $n \not\equiv 0 \bmod (3)$, $1 + X^n + X^{2n} \equiv 0 \bmod (F_1)$,

ce qui donne la réponse à la question posée.

Evidemment si $F_n \in \mathbb{C}[X]$ il est aussi rapide de chercher pour quelles valeurs de n : $j^{2n} + j^n + 1 = 0$, mais ici nous avons une réponse valable pour un corps K quelconque.

Exemple 2 : On prend $K = \mathbb{C}$ et $P = X^2 + aX + b \in \mathbb{C}[X]$. Si α est une racine de P , qu'on ne cherchera pas à calculer, exprimer α^4 en fonction de α .

Raisonnons $\bmod (P)$. On a $X^2 \equiv -aX - b$,

$$X^3 \equiv -a(-aX + b) - bX = (a^2 - b)X + ab,$$

$$X^4 \equiv (a^2 - b)(-aX - b) + abX = (2ab - a^3)X + b^2$$

et par suite, en utilisant le fait que $P(\alpha) = 0$, $\alpha^4 = (2ab - a^3)\alpha + b^2$.

Soit maintenant L une extension de K et X une indéterminée sur L ; si A, B, P sont des éléments de $K[X]$, la congruence $A \equiv B \bmod (P)$ dans $K[X]$ et la congruence $A \equiv B \bmod (P)$ dans $L(X)$ sont des relations équivalentes, en raison de l'unicité de la division euclidienne dans $K[X]$ et dans $L[X]$. C'est le cas par exemple pour $K = \mathbb{R}$ et $L = \mathbb{C}$, où la remarque précédente pourra être très utile.

L'anneau quotient de $K[X]$ par un idéal

Fixons dans $K[X]$ un idéal non nul $\mathfrak{a} \neq K[X]$, et soit P son générateur normalisé qui est donc de degré $d \geq 1$. Nous poserons

$$(3) \quad P = X^d + c_1 X^{d-1} + \cdots + c_d.$$

Notons \mathcal{A} l'ensemble des classes de congruence $\bmod (P)$ dans $K[X]$. Les propriétés (CG2) et (CG3) montrent qu'on définit deux lois de composition internes $+$ et \times sur \mathcal{A} en posant, si $\gamma \in \mathcal{A}$ et $\delta \in \mathcal{A}$:

$$\gamma + \delta = \text{classe de congruence mod } (P) \text{ commune à tous les } A + B \quad (A \in \gamma, B \in \delta)$$

$$\gamma \times \delta = \text{classe de congruence mod } (P) \text{ commune à tous les } AB \quad (A \in \gamma, B \in \delta)$$

Du fait que \mathfrak{a} est un idéal de $K[X]$, on déduit immédiatement que : \mathcal{A} est un groupe abélien pour la loi $+$, dont l'élément neutre est la classe de 0, c'est-à-dire \mathfrak{a} , et : la loi \times est associative, commutative et distributive par rapport à la loi $+$; de plus l'élément $1_{\mathcal{A}}$ = classe de congruence de 1 mod (P) est élément neutre de la loi \times . Bref, les lois d'addition $+$ et de multiplication \times définies ci-dessus munissent l'ensemble \mathcal{A} d'une structure d'anneau commutatif.

L'anneau ainsi obtenu est appelé **anneau quotient** de $K[X]$ par l'idéal $\mathfrak{a} = PK[X]$, et noté $K[X]/\mathfrak{a}$: c'est un anneau *non nul* car $\mathfrak{a} \neq K[X]$.

Le groupe additif de cet anneau $K[X]/\mathfrak{a}$ n'est autre, par construction même, que le *groupe additif quotient* de $(K[X], +)$ par son sous-groupe \mathfrak{a} (cf. § V.7).

L'application canonique $\varphi : K[X] \rightarrow K[X]/\mathfrak{a}$ est un **homomorphisme d'anneaux**, en vertu des propriétés (CG2) et (CG3) ; φ est surjectif, et son noyau $\text{Ker}(\varphi)$ est l'idéal \mathfrak{a} .

Il est facile de munir l'anneau $K[X]/\mathfrak{a}$ d'une structure naturelle de **K -algèbre**. En effet fixons $\lambda \in K$; si $A \in K[X]$, $B \in K[X]$ et si $A \equiv B \pmod{(P)}$, alors $\lambda A \equiv \lambda B \pmod{(P)}$. On définit donc une loi externe de domaine K sur l'anneau $K[X]/\mathfrak{a}$ en posant, pour tout $\gamma \in K[X]/\mathfrak{a}$, $\lambda \gamma$ = classe de congruence des $(\lambda A)_{A \in \gamma}$. Il est élémentaire de vérifier que cette loi externe a toutes les qualités requises pour définir sur $K[X]/\mathfrak{a}$ une structure de K -algèbre.

Puisque K est un corps, l'application canonique $K \rightarrow K[X]/\mathfrak{a}$, $\lambda \mapsto \lambda \cdot 1_{\mathcal{A}}$ est injective, et donc la K -algèbre $K[X]/\mathfrak{a}$ admet K comme sous-anneau.

THÉORÈME VII.7.2

Soit \mathfrak{a} l'idéal engendré par le polynôme P défini par (3), et soit :

$\varphi : K[X] \rightarrow K[X]/\mathfrak{a} = \mathcal{A}$ l'homomorphisme canonique.

(I) La suite $(1_{\mathcal{A}}, \varphi(X), \varphi(X^2), \dots, \varphi(X^{d-1}))$ est une base du K -ev $K[X]/\mathfrak{a}$.

(II) Il y a équivalence entre les trois propriétés suivantes :

- a) l'anneau $K[X]/\mathfrak{a}$ est intègre
- b) l'anneau $K[X]/\mathfrak{a}$ est un corps
- c) le polynôme P est irréductible dans $K[X]$.

Démonstration :

(I) Montrons d'abord que $1_{\mathcal{A}}, \varphi(X), \dots, \varphi(X^{d-1})$ sont linéairement indépendants sur K . Soit donc

$$\lambda_0 1_{\mathcal{A}} + \lambda_1 \varphi(X) + \dots + \lambda_{d-1} \varphi(X^{d-1}) = 0_{\mathcal{A}}$$

une combinaison linéaire nulle, avec $\lambda_i \in K$ pour $i \in \llbracket 0, d-1 \rrbracket$. Comme le polynôme

$$\lambda_0 + \lambda_1 X + \dots + \lambda_{d-1} X^{d-1} \in \mathfrak{a} = PK[X]$$

et que P est de degré d , on en déduit que tous les λ_i sont nuls.

Montrons ensuite que $1_{\mathcal{A}}, \varphi(X), \dots, \varphi(X^{d-1})$ engendrent le K -ev \mathcal{A} : si $\gamma \in \mathcal{A}$, soit A_0 l'unique élément de $\gamma \cap K_{d-1}[X]$ (cf. théorème VII.7.1). Alors

$$A_0 = \lambda_0 + \lambda_1 X + \dots + \lambda_{d-1} X^{d-1} \quad \text{pour des } \lambda_i \in K, \text{ d'où}$$

$$\gamma = \varphi(A_0) = \lambda_0 1_{\mathcal{A}} + \dots + \lambda_{d-1} \varphi(X^{d-1}).$$

(II) $(b) \Rightarrow (a)$ est évident. Supposons $K[X]/\mathfrak{a}$ intègre et soit $P = Q_1 Q_2$ dans $K[X]$; alors $\varphi(Q_1) \varphi(Q_2) = 0_{\mathcal{A}}$, d'où $\varphi(Q_1) = 0_{\mathcal{A}}$ ou $\varphi(Q_2) = 0_{\mathcal{A}}$, c'est-à-dire Q_1 ou Q_2 est multiple de P , ce qui entraîne la propriété (c). Mais alors, soit $\gamma \in \mathcal{A} \setminus \{0\}$ et A un représentant de γ dans $K[X]$. On a : $A \not\equiv 0 \pmod{P}$, donc A et P sont premiers entre eux puisque P est irréductible dans $K[X]$. Le *théorème de Bezout* fournit alors U et V dans $K[X]$ tels que $UA + VP = 1$, d'où

$$\varphi(U) \varphi(A) + \varphi(V) \varphi(P) = 1_{\mathcal{A}}.$$

Mais $\varphi(P) = 0_{\mathcal{A}}$, donc $\varphi(U) \varphi(A) = 1_{\mathcal{A}}$, donc γ est *inversible* dans \mathcal{A} et son inverse est $\varphi(U)$, ce qui donne d'ailleurs une méthode pratique pour le calcul de cet inverse. On a donc prouvé que tout élément non nul de \mathcal{A} est inversible, et comme \mathcal{A} est un anneau non nul, $(c) \Rightarrow (b)$. Comme $(a) \Rightarrow (c)$ et $(b) \Rightarrow (a)$ il en résulte bien que les trois propriétés sont équivalentes. ■

On remarquera que $\alpha = \varphi(X)$ est toujours un zéro de P dans \mathcal{A} , car

$$P(\alpha) = \varphi(P(X)) = 0_{\mathcal{A}}.$$

De plus si l'une des conditions (a), (b) ou (c) du théorème VII.7.2 est satisfaite, alors α est **élément algébrique sur K dans le corps $K[X]/\mathfrak{a}$** et le polynôme $P(Y)$ est le **polynôme minimal** de α sur K dans l'algèbre des polynômes $L[Y]$, où $L = K[X]/\mathfrak{a}$. En effet $P(\alpha) = 0$ comme on vient de le voir ; et si $F \in K[Y]$ et $F(\alpha) = 0$, alors

$$F(\varphi(X)) = 0 = \varphi(F(X)),$$

puisque φ est un homomorphisme de K -algèbres. Donc

$$F(X) \in \text{Ker}(\varphi) = \mathfrak{a} = PK[X], \quad \text{d'où le résultat.}$$

Nous retrouverons par une autre voie le théorème VII.7.2 après l'étude des espaces vectoriels de dimension finie. Pour l'instant observons sa similitude avec le théorème IV.4.4 : cette similitude n'est pas due au hasard, mais tient au fait que \mathbb{Z} et $K[X]$ sont tous deux des anneaux principaux.

Voici une importante conséquence du théorème VII.7.2 :

THÉORÈME VII.7.3

|| Soit $F \in K[X]$ un polynôme non constant. Il existe au moins **une extension**
 || L de K telle que, dans l'anneau de polynômes $L[Y]$, le polynôme
 || $F(Y)$ soit **dissocié**.

Démonstration :

On procède par récurrence sur $d = \deg(F)$, à partir de $d = 1$ pour lequel le théorème est évident. Supposons le vrai pour l'entier $d - 1$, avec $d \geq 2$, et pour tout corps commutatif. Considérons un facteur P , irréductible dans $K[X]$, de F , d'où $F = GP$, $G \in K[X]$. Soit Ω le corps $K[X]/PK[X]$: c'est une extension de K . Si Z est une indéterminée sur Ω , substituons Z à X ; on obtient : $F(Z) = G(Z)P(Z)$. D'autre part l'image canonique α de X dans Ω est racine de P dans Ω , comme nous l'avons remarqué juste après la démonstration du théorème VII.7.2. Donc $P(Z) = (Z - \alpha)Q(Z)$ avec $Q \in \Omega[Z]$, d'où $F(Z) = (Z - \alpha)G(Z)Q(Z)$. Le polynôme $H(Z) = G(Z)$;

$\Omega[Z]$ et il a pour degré $d - 1$. L'hypothèse de récurrence entraîne l'existence d'une extension L de Ω telle que, si Y est une indéterminée sur L , le polynôme $H(Y)$ soit dissocié dans $L[Y]$. Substituons Y à Z dans $F(Z)$; on obtient : $F(Y) = (Y - \alpha) H(Y)$, donc $F(Y)$ est dissocié dans $L[Y]$ puisque $H(Y)$ l'est. ■

Une extension L telle que $F(Y)$ soit dissocié dans $L[Y]$ s'appelle un **corps de dissociation de F** . Il n'y a pas unicité d'une telle extension puisque, si L en est une, toute extension L' en est une autre. Cependant on peut prouver qu'il en existe une « la plus petite possible », mais cela dépasse le niveau de cet ouvrage (voir [20], [27]).

Anneau quotient d'un anneau commutatif par un idéal

Le passage au quotient pour un anneau commutatif quelconque à partir d'un idéal arbitraire s'effectue de manière très analogue à ce qui vient d'être fait pour $K[X]/PK[X]$. En effet, dans ce qui précède, nous n'avons pas utilisé le fait que tous les idéaux de $K[X]$ sont principaux, ni même le fait que $K[X]$ est un anneau intègre. La seule chose qui a vraiment servi est la structure d'idéal de l'ensemble $\mathfrak{a} = PK[X]$.

Soit donc A un anneau commutatif et \mathfrak{a} un idéal de A . La relation binaire sur A définie par $x \mathcal{R} y$ ssi $x - y \in \mathfrak{a}$ (x et y dans A) est une **relation d'équivalence** sur A , appelée **congruence modulo \mathfrak{a}** et notée $x \equiv y \pmod{\mathfrak{a}}$. Nous avons déjà vu au § V.7 que l'ensemble quotient \mathcal{A} , noté A/\mathfrak{a} , de A par cette relation d'équivalence, est muni d'une structure naturelle de groupe, appelée **groupe quotient de A par le sous-groupe additif \mathfrak{a}** . L'addition dans ce groupe est définie par : si $\gamma \in \mathcal{A}$ et $\delta \in \mathcal{A}$, $\gamma + \delta =$ classe de congruence mod \mathfrak{a} commune à tous les $x + y$ ($x \in \gamma, y \in \delta$). L'élément nul est $\mathfrak{a} =$ classe de 0, noté $0_{\mathcal{A}}$ (ou simplement 0). L'application canonique $\varphi : A \rightarrow A/\mathfrak{a}, x \mapsto$ classe de x est un homomorphisme surjectif de groupes, de noyau \mathfrak{a} .

La congruence mod \mathfrak{a} possède la propriété suivante :

Si $x \equiv y \pmod{\mathfrak{a}}$ et si $x' \equiv y' \pmod{\mathfrak{a}}$ (x, y, x', y' dans A), alors $xx' \equiv yy' \pmod{\mathfrak{a}}$, ce qui se vérifie en écrivant $x' = x + \lambda, y' = y + \mu$ avec $\lambda \in \mathfrak{a}, \mu \in \mathfrak{a}$, d'où $x' y' = xy + z$ avec $z = \lambda y + \mu x + \lambda \mu$ et $z \in \mathfrak{a}$ justement *parce que \mathfrak{a} est un idéal*.

Cette propriété permet de munir le groupe additif quotient A/\mathfrak{a} d'un produit ainsi défini : si $\gamma \in A/\mathfrak{a}, \delta \in A/\mathfrak{a}$, $\gamma\delta =$ classe de congruence commune à tous les xy , pour $x \in \gamma, y \in \delta$.

On vérifie alors facilement que *ce produit munit le groupe additif A/\mathfrak{a} d'une structure d'anneau commutatif*. L'élément unité est $1_{\mathcal{A}} = \varphi(1)$ et $\varphi : A \rightarrow A/\mathfrak{a}$ est un homomorphisme d'anneaux. On pose :

DÉFINITION VII.7.1

$\{$ Avec les notations ci-dessus, l'anneau commutatif A/\mathfrak{a} s'appelle
 $\{$ l'**anneau quotient** de A par l'idéal \mathfrak{a} , et $\varphi : A \rightarrow A/\mathfrak{a}$ s'appelle
 $\{$ l'**homomorphisme canonique**.

En particulier $A/\{0\}$ est naturellement isomorphe à A , et A/\mathfrak{a} est l'anneau nul ssi $\mathfrak{a} = A$.

On se souvient (cf. § III.7) que pour tout homomorphisme d'anneaux $f: A \rightarrow B$, le noyau $\text{Ker}(f)$ est un idéal de A .

THÉORÈME VII.7.4

Soit A et B deux anneaux commutatifs et $f: A \rightarrow B$ un homomorphisme d'anneaux ; $\varphi: A \rightarrow A/\text{Ker}(f)$ désigne l'homomorphisme canonique et $j: \text{Im}(f) \rightarrow B$ l'injection canonique.
Il existe une unique application $\bar{f}: A/\text{Ker}(f) \rightarrow \text{Im}(f)$ telle que $f = j \circ \bar{f} \circ \varphi$, et cette application est un **isomorphisme d'anneaux**.

Démonstration :

On sait déjà, d'après le § V.7, que \bar{f} existe et est unique et que c'est un isomorphisme de groupes. Il reste donc à prouver que $\bar{f}(\alpha\beta) = \bar{f}(\alpha)\bar{f}(\beta)$ pour tous α, β dans $A/\text{Ker}(f)$, mais cela résulte du fait que f est un homomorphisme d'anneaux et de la définition du produit dans $A/\text{Ker}(f)$. ■

Exemple 3 : Soit A un anneau commutatif ; l'application naturelle $f: \mathbb{Z} \rightarrow A, n \mapsto n \cdot 1_A$ est un homomorphisme d'anneaux. Son image I est l'anneau engendré par 1_A dans A , qui est évidemment le plus petit (au sens de l'inclusion) sous-anneau de A . Son noyau \mathcal{N} est un idéal de \mathbb{Z} , donc $\mathcal{N} = q\mathbb{Z}$ pour un unique $q \in \mathbb{N}$. On pose :

DÉFINITION VII.7.2

Si A est un anneau commutatif, on appelle **indicateur de torsion de A** l'entier $q \in \mathbb{N}$ tel que $q\mathbb{Z}$ soit le noyau de l'homomorphisme $n \mapsto n \cdot 1_A$ de \mathbb{Z} dans A .

L'indicateur de torsion est nul ssi f est injectif. Dans ce cas, \mathbb{Z} peut être considéré comme un sous-anneau de A .

Si l'indicateur de torsion q est ≥ 1 , le théorème VII.7.4 montre que l'image $I = \text{Im}(f)$ est canoniquement isomorphe à l'anneau $\mathbb{Z}/q\mathbb{Z}$.

Dans ce cas, on a : $qx = 0$ pour tout $x \in A$.

Bien évidemment si A est un corps, on reconnaît dans l'indicateur de torsion la *caractéristique* de ce corps (on se souvient alors que cette caractéristique est ou bien nulle, ou bien un nombre premier).

Revenons à la démonstration de la deuxième partie du théorème VII.7.2. Les seules choses qui ont été utilisées sont : le *théorème de Bezout*, le fait qu'un polynôme irréductible dans $K[X]$ est premier avec tout polynôme qu'il ne divise pas, et le *théorème de Gauss* VII.2.5. On a vu que toutes ces propriétés restent vraies dans un anneau principal (fin du § VII.2).

Nous obtenons donc le théorème plus général suivant :

THÉOREME VII.7.5

Soit A un anneau principal, et $p \in A$ un élément non nul et non inversible. Les propriétés suivantes sont équivalentes :

- a) l'anneau quotient A/pA est intègre
- b) l'anneau quotient A/pA est un corps
- c) l'élément p est irréductible.

Exercice 1 : On prend $K = \mathbb{C}$. Pour quels $n \in \mathbb{N}$ le polynôme $(X+1)^n - X^n - 1$ est-il divisible par $X^2 + X + 1$? Pour quels $m \in \mathbb{N}$ le polynôme $(X-1)^{m+1} + X^{2m-1}$ est-il divisible par $X^2 - X + 1$? Pour quels $m \in \mathbb{N}$ le polynôme

$$1 - X^m + X^{2m} - X^{3m} + X^{4m}$$

est-il divisible par $1 - X + X^2 - X^3 + X^4$?

Exercice 2 : Soit P un polynôme de $\mathbb{C}[X]$ tel que $P(0) \neq 0$ et ayant toutes ses racines distinctes. Montrer qu'il existe $A \in \mathbb{C}[X]$ tel que $A^2 \equiv X \pmod{(P)}$.

Indication : Poser $A_i = \prod_{j \neq i} (X - \alpha_j)$, les α_j désignant les racines de P , et chercher A sous la forme $\sum_i a_i A_i$.

Exercice 3 : Soit $P \in K[X]$. Montrer que si $P(X^n)$ est divisible par $X - 1$, alors $P(X^n)$ est divisible par $X^n - 1$, n désignant un entier ≥ 1 .

Exercice 4 : Soit $P \in K[X]$. Démontrer que le polynôme $P(P(X)) - X$ est divisible par $P(X) - X$.

Exercice 5 : Soit $P \in \mathbb{Q}[X]$. Montrer que si P admet sur \mathbb{R} la racine $a + b\sqrt{m}$ (a, b, m rationnels, m non carré parfait), il admet aussi la racine $a - b\sqrt{m}$ au même ordre de multiplicité. Même question avec la racine complexe $a + ib\sqrt{m}$.

Exercice 6 : Chercher le polynôme minimal P du nombre algébrique $a = \sqrt[3]{2} + \sqrt[3]{4}$. Exprimer ensuite toutes les racines de P en fonction de $\sqrt[3]{2}$, de j et de \bar{j} .

Exercice 7 : On désigne par P le polynôme $X^5 - 1$ et par $\zeta = e^{2i\pi/5}$ l'une des racines primitives 5-ième de l'unité. Soit E le \mathbb{Q} -ev engendré dans \mathbb{C} par les racines de P .

Montrer que $(1, \zeta, \zeta^2, \zeta^3)$ est une base de E . Montrer que E est un sous-anneau de \mathbb{C} . Soit $A \in \mathbb{Q}[X] \setminus \{0\}$ de degré ≤ 3 . Montrer que A est premier avec

$$1 + X + X^2 + X^3 + X^4.$$

En déduire que $A(\zeta)$ est inversible dans E et finalement que E est un sous-corps de \mathbb{C} .

Exercice 8 : On considère dans l'anneau $A = \mathbb{Q}[X]$ la congruence modulo $X^3 - 3$. Vérifier que l'anneau quotient obtenu est un corps. Donner les inverses des classes $\overline{X-1}$, $\overline{X^4+1}$.

Exercice 9 : Dans l'anneau $\mathbb{R}[X]$ des polynômes à coefficients réels, on considère la congruence modulo $(X^2 + 1)$. Montrer que l'anneau quotient $\mathbb{R}[X]/(X^2 + 1)$ obtenu est un corps que nous noterons K . On définit alors une application $f : K \rightarrow \mathbb{C}$ de la façon suivante : si $aX + b$ est le représentant de degré ≤ 1 de la classe de congruence $\gamma \in K$, alors on pose $f(\gamma) = ai + b \in \mathbb{C}$. Montrer que f est un **isomorphisme de corps**. (Cauchy).

Si l'on remplace \mathbb{R} par \mathbb{Q} dans cette construction on obtient un corps isomorphe à $\mathbb{Q}(i)$, sous-corps de \mathbb{C} .

Pour quelles valeurs du nombre premier p obtient-on encore un corps à partir de l'anneau $\mathbb{Z}/p\mathbb{Z}[X]$?

Exercice 10 : On considère le polynôme $P = X^3 - 3X + 1$ qui est irréductible sur \mathbb{Q} dont on admettra ici que les trois racines sur \mathbb{C} sont réelles. Soit u l'une de ces racines

a) Soit E le \mathbb{Q} -ev engendré par $1, u$ et u^2 (c'est donc un sous-ev de \mathbb{R}). Montrer que E est un sous-anneau de \mathbb{R} . Montrer que $1, u$ et u^2 sont \mathbb{Q} -linéairement indépendants. Montrer que E est un corps.

b) Montrer que $u^2 - 2$ est une racine de P . En déduire que les trois racines de P sont dans E .

c) On considère l'application \mathbb{Q} -linéaire $\sigma : E \rightarrow E$ définie par $\sigma(1) = 1, \sigma(u) = u^2 - 2, \sigma(u^2) = (u^2 - 2)^2$. Montrer que σ induit une permutation sur l'ensemble des racines de P . En déduire que σ^3 est l'application Id_E . Montrer que σ est bijective et que, pour tout x de E , la relation $\sigma(x) = x$ équivaut à $x \in \mathbb{Q}$.

Exercice 11 : Soit P et G des polynômes de degrés respectifs $m \geq 1$ et $n \geq 1$ dans $K[X]$. On considère l'application $\varphi : K[X] \rightarrow K_{n-1}[X]$ qui à tout polynôme F associe le reste \bar{F} de la division euclidienne de FP par G . Démontrer que φ est une application K -linéaire.

Soit $H \in \text{Im}(\varphi)$ et H_1 le reste de $MH \bmod (G)$, avec M quelconque dans $K[X]$. Montrer que $H_1 \in \text{Im}(\varphi)$. En déduire que φ est bijective ssi $1 \in \text{Im}(\varphi)$. En déduire que φ est bijective ssi P et G sont premiers entre eux. Calculer alors $\varphi^{-1}(R)$ où $R \in K_{n-1}[X]$.

Exercice 12 : Si $\alpha \in \mathbb{C}$ désignons par S l'ensemble $\{A(\alpha)\}_{A \in \mathbb{Q}[X]}$. S est un sous-anneau de \mathbb{C} et un sous \mathbb{Q} -ev de \mathbb{C} . Montrer que si B est un sous-anneau de \mathbb{C} contenant à la fois \mathbb{Q} et $\{\alpha\}$, alors $S \subset B$. Dans la suite S sera noté $\mathbb{Q}(\alpha)$.

On suppose que α est algébrique de degré n , de polynôme minimal

$$P_\alpha = p_0 + p_1 X + \cdots + p_{n-1} X^{n-1} + X^n.$$

Montrer que $\mathcal{B}_\alpha = (1, \alpha, \dots, \alpha^{n-1})$ est une base du \mathbb{Q} -ev $\mathbb{Q}(\alpha)$. Si $\alpha \neq 0, \frac{1}{\alpha} \in \mathbb{Q}(\alpha)$; écrire les composantes de $\frac{1}{\alpha}$ dans \mathcal{B}_α .

Montrer que l'application $\sigma_\alpha : x \mapsto \alpha x$ est un endomorphisme du \mathbb{Q} -ev \mathbb{C} et que sa restriction ω_α à $\mathbb{Q}(\alpha)$ est un endomorphisme de $\mathbb{Q}(\alpha)$; et un isomorphisme si $\alpha \neq 0$. Soit β un élément de $\mathbb{Q}(\alpha)$. Montrer que β est algébrique de degré $\leq n$.

Montrer que $\mathbb{Q}(\alpha)$ est isomorphe à l'anneau quotient $\mathbb{Q}[X]/P_\alpha$ et que c'est un corps. On pourra prendre comme exemple $\alpha = j^3 \sqrt{2}$ et l'on constatera que le polynôme minimal P_α n'est pas dissocié sur $\mathbb{Q}(\alpha)$. Essayer de construire une extension de $\mathbb{Q}(\alpha)$ dans laquelle P_α serait dissocié.

Exercice 13 : Pour $n \geq 2$ on désigne par P le polynôme $1 + X + \cdots + X^{n-1}$ et par ζ le nombre $e^{2i\pi/n} \in \mathbb{C}$. On note A l'anneau quotient $\mathbb{R}[X]/P$ muni de sa structure de \mathbb{R} -ev. Montrer que dans $\mathbb{R}[X]$ on a $X^r \equiv 1 \bmod (P)$ ssi r est multiple de n .

Montrer qu'il existe une application h et une seule : $A \rightarrow \mathbb{C}$, qui est \mathbb{R} -linéaire, qui est un homomorphisme d'anneaux et qui vérifie $h(\bar{X}) = \zeta$ (\bar{X} étant la classe de X).

Montrer que h est injective pour $n = 2$ et bijective pour $n = 3$. Que peut-on dire de l'anneau A dans ces deux cas ?

Montrer que, pour $n \geq 4$, h n'est pas injective et que l'anneau A n'est pas intègre.

Exercice 14 : Montrer que le nombre réel $2 \cos \frac{2\pi}{7}$ est racine du polynôme

$$P = X^3 + X^2 - 2X - 1 \in \mathbb{Q}[X]$$

et que ce polynôme est irréductible sur \mathbb{Q} . Construire une extension de \mathbb{Q} dans laquelle P soit dissocié.

Chapitre VIII

FRACTIONS RATIONNELLES. NOTIONS SUR LES SÉRIES FORMELLES

Dans tout ce chapitre, K désigne un corps commutatif arbitraire.

§ VIII.1 LE CORPS $K(X)$

La notion de *corps des fractions d'un anneau intègre* a été développée de façon générale au § II.6. Or justement l'anneau $K[X]$ des polynômes à une indéterminée sur le corps commutatif K est un anneau intègre. On peut donc construire un corps commutatif \mathcal{F} à partir de $K[X]$. Rappelons les principales étapes de cette construction. On commence par former l'ensemble quotient de $K[X] \times (K[X] \setminus \{0\})$ par la relation d'équivalence

$$\mathcal{R} : (A_1, B_1) \mathcal{R} (A_2, B_2) \quad \text{ssi} \quad A_1 B_2 = A_2 B_1.$$

On munit ensuite cet ensemble quotient \mathcal{F} des deux lois d'addition et de multiplication :

$$(\forall F_1 \in \mathcal{F}, \forall F_2 \in \mathcal{F})$$

$$F_1 + F_2 = \text{classe mod } (\mathcal{R}) \text{ commune à tous les } \frac{A_1 B_2 + A_2 B_1}{B_1 B_2} \quad ((A_i, B_i) \in F_i^2)$$

$$F_1 F_2 = \text{classe mod } (\mathcal{R}) \text{ commune à tous les } \frac{A_1 A_2}{B_1 B_2} \quad ((A_i, B_i) \in F_i^2).$$

L'image canonique F dans \mathcal{F} de $(A, B) \in K[X] \times (K[X] \setminus \{0\})$ se note A/B ou $\frac{A}{B}$ et on dit que la fraction $\frac{A}{B}$ représente F ; A est appelé *numérateur*, B *dénominateur* de cette fraction, et l'on écrit $F = \frac{A}{B}$, notation qui se justifie par le fait que F , dans le corps \mathcal{F} , est le produit de $\frac{A}{1} = A$ par l'inverse de B qui n'est autre que $\frac{1}{B}$ d'après la loi de multiplication ci-dessus.

L'application $A \mapsto \frac{A}{1_K}$ est un homomorphisme injectif d'anneaux, de l'anneau $K[X]$ dans le corps \mathcal{F} , à l'aide duquel on identifie $K[X]$ à un sous-anneau de \mathcal{F} .

DÉFINITION VIII.1.1

Le corps des fractions de l'anneau de polynômes $K[X]$ s'appelle **corps des fractions rationnelles à une indéterminée sur K** et se note $K(X)$.

Puisque K s'identifie à un sous-anneau de $K[X]$, c'est *a fortiori* un sous-corps de $K(X)$. Ainsi le corps $K(X)$ qui est une extension du corps K se trouve naturellement muni d'une **structure de K -algèbre** qui joue un rôle important.

Le sous-corps K de $K(X)$ s'appelle sous-corps des *fractions constantes*.

Formes irréductibles

Soit $F \in K(X) \setminus \{0\}$; nous appellerons **représentant irréductible** de F (ou « forme irréductible » de F) toute fraction A/B telle que $A/B = F$ et que A et B soient **premiers entre eux**.

THÉORÈME VIII.1.1

Si $F \in K(X) \setminus \{0\}$, F possède des représentants irréductibles. Soit A_0/B_0 l'un quelconque d'entre eux ; alors les représentants irréductibles de F sont les $\lambda A_0/\lambda B_0$, où $\lambda \in K^*$, et les représentants de F sont les CA_0/CB_0 , où $C \in K[X] \setminus \{0\}$.

Démonstration :

Soit A/B un représentant de F . Par hypothèse $A \neq 0$ et $B \neq 0$. Soit D le pgcd (non nul) de A et B . On a $A = A_0 D$, $B = B_0 D$ avec $\text{pgcd}(A_0, B_0) = 1$. Par suite $A/B = A_0/B_0 = F$, et A_0/B_0 est bien un représentant irréductible de F dont l'existence est ainsi prouvée.

Il est clair que, pour tout $C \in K[X] \setminus \{0\}$, $A_0 C/B_0 C$ est bien un représentant de F . Réciproquement cherchons la forme d'un représentant quelconque A/B de F . Il est tel que $A \neq 0$, $B \neq 0$ et $AB_0 = A_0 B$ par hypothèse ; B_0 , premier avec A_0 , doit diviser B (théorème de Gauss VII.2.5), d'où $B = CB_0$ et ensuite $A = CA_0$, avec $C \in K[X] \setminus \{0\}$. De plus $\text{pgcd}(A, B) = C$, d'où $\text{pgcd}(A, B) = 1$ ssi $C \in K^*$, ce qui donne la forme des représentants irréductibles. ■

Nous pouvons donc parler de **la forme irréductible P/Q** de $F \in K[X] \setminus \{0\}$, étant sous-entendu qu'en réalité le couple (P, Q) figurant dans cette expression n'est défini qu'à la multiplication près par une constante $\lambda \in K^*$.

Remarque 1 : Ce théorème VIII.1.1 est analogue au théorème IV.2.6 ; ce n'est pas un hasard, mais la conséquence du fait que \mathbb{Z} et $K[X]$ sont tous deux des anneaux factoriels.

Valuations. Degré

Nous noterons \mathcal{J} l'ensemble des polynômes $P \in K[X]$ irréductibles et normalisés. La décomposition en facteurs irréductibles dans $K[X]$ nous permet d'obtenir la **décomposition des fractions rationnelles en facteurs irréductibles** :

THÉOREME VIII.1.2

Soit une fraction rationnelle $F \in K(X) \setminus \{0\}$. Alors,

(I) il existe une constante $u_F \in K^*$ et une famille $(\alpha_P)_{P \in \mathcal{J}}$ à support fini, où $\alpha_P \in \mathbb{Z}$ pour tout $P \in \mathcal{J}$, telle que F puisse s'écrire sous la forme

$$F = u_F \prod_{P \in \mathcal{J}} P^{\alpha_P},$$

(II) et il y a unicité d'une telle décomposition.

Démonstration :

Existence : soit A et B les deux termes d'un représentant quelconque A/B de la fraction $F \in K(X)^*$. D'après le théorème VII.3.1 on peut décomposer chacun des polynômes A et B en produit de facteurs irréductibles : $A = a_0 \prod_{P \in \mathcal{J}} P^{v_P(A)}$ et $B = b_0 \prod_{P \in \mathcal{J}} P^{v_P(B)}$ d'où l'existence d'au moins une décomposition de F du type indiqué, avec $\alpha_P \in \mathbb{Z}$.

Unicité : Si l'on considère deux décompositions de F de ce type, par division membre à membre on obtiendrait

$$(1) \quad 1 = u \prod_{P \in \mathcal{J}} P^{\beta_P}.$$

Si les β_P sont dans \mathbb{N} , cette relation (1) entraîne $u = 1$ et $\beta_P = 0$ d'après l'unicité de la décomposition des polynômes de $K[X] \setminus \{0\}$ en facteurs irréductibles. Envisageons l'hypothèse où l'ensemble \mathcal{N} des $P \in \mathcal{J}$ tels que $\beta_P < 0$ soit non vide. En multipliant les deux membres de (1) par $D = \prod_{P \in \mathcal{N}} P^{-\beta_P}$, on obtient $D = u \prod_{P \in \mathcal{J} \setminus \mathcal{N}} P^{\beta_P}$, ce qui donnerait deux

écritures distinctes du polynôme non constant D comme produit de facteurs irréductibles, en contradiction avec le théorème VII.3.1. Donc $\mathcal{N} = \emptyset$, d'où $\beta_P = 0$ pour tout $P \in \mathcal{J}$ et $u = 1$. ■

Remarque 2 : On s'est servi dans la démonstration précédente du fait que dans $K(X)$ la division par un élément non nul est parfaitement définie, c'est-à-dire de la structure de groupe multiplicatif de $K(X)^*$. Une question qui se pose naturellement est d'essayer d'élucider la structure de ce groupe. Pour cela notons, comme d'habitude, $\mathbb{Z}^{(\mathcal{J})}$ l'ensemble des *familles à support fini* $(\alpha_P)_{P \in \mathcal{J}}$, où $\alpha_P \in \mathbb{Z}$ pour tout $P \in \mathcal{J}$. Cet ensemble peut être muni d'une structure naturelle de *groupe abélien*, qu'on note additivement, la loi interne de ce groupe étant définie par l'addition ordinaire sur chaque coordonnée : si

$$\alpha = (\alpha_P)_{P \in \mathcal{J}} \quad \text{et} \quad \beta = (\beta_P)_{P \in \mathcal{J}},$$

alors $\alpha + \beta = (\alpha_P + \beta_P)_{P \in \mathcal{J}}$ (cf. § V.1). Il suffit alors de munir l'ensemble $K^* \times \mathbb{Z}^{(\mathcal{J})}$ de sa structure de *groupe produit* du groupe multiplicatif K^* par le groupe $\mathbb{Z}^{(\mathcal{J})}$ que nous venons de définir pour obtenir le résultat suivant :

L'application $J : K^* \times \mathbb{Z}^{(\mathcal{J})} \rightarrow K(X)^*$, $(u, (\alpha_P)_{P \in \mathcal{J}}) \mapsto u \prod_{P \in \mathcal{J}} P^{\alpha_P}$ dont

le théorème VIII.1.2 montre que c'est une bijection, est aussi un **isomorphisme de groupes**. En effet si $F = J(u, \alpha)$ et $G = J(v, \beta)$ on a évidemment $J(uv, \alpha + \beta) = FG$, conséquence du fait que pour chacun des facteurs R figurant dans la décomposition de F ou de G on a $R^{m+n} = R^m \cdot R^n$ pour $m \in \mathbb{Z}$, $n \in \mathbb{Z}$.

Remarque 3 : Soit $F \in K(X)^*$. Ecrivons F sous la forme

$$F = J(u, \alpha) = u \prod_{P \in \mathcal{J}} P^{\alpha_P}.$$

Introduisons les ensembles $\mathcal{P} = \{P \in \mathcal{J} \mid \alpha_P \geq 0\}$

et $\mathcal{N} = \{P \in \mathcal{J} \mid \alpha_P < 0\}$ et posons : $N = u \prod_{P \in \mathcal{P}} P^{\alpha_P}$, et : $D = \prod_{P \in \mathcal{N}} P^{-\alpha_P}$.

Alors $\frac{N}{D}$ est un représentant irréductible de F . En effet il est clair d'une part que $\frac{N}{D} = F$, et d'autre part que N et D sont premiers entre eux.

Notons également que pour que $F \in K(X)^*$ soit égale à un *polynôme* de $K[X] \setminus \{0\}$, il faut et il suffit que $\alpha_P \in \mathbb{N}$ pour tout $P \in \mathcal{J}$.

Remarque 4 : Prouver un théorème d'existence et d'unicité est une chose. Trouver une décomposition **effective** d'une fraction $F \in K(X)^*$ en facteurs irréductibles en est une autre, qui relève du domaine de l'Analyse numérique, et se heurte à deux obstacles : en dehors de cas simples (K corps fini F_q , $K = \mathbb{R}$ ou $K = \mathbb{C}$) il n'est déjà pas si facile d'expliciter les éléments de l'ensemble \mathcal{J} ; et même si on les connaît tous, il est hors de question de les essayer un par un (la même difficulté se présente déjà pour décomposer un entier assez grand en facteurs premiers). C'est dans

problèmes pratiques que l'aide d'un ordinateur puissant peut être très appréciée.

Revenons à la théorie pour laquelle, de manière analogue au § VII.1, il nous paraît commode d'utiliser l'ensemble $\overline{\mathbb{Z}} = \mathbb{Z} \cup \{-\infty, +\infty\}$ dans lequel on prolonge l'ordre et l'addition de \mathbb{Z} , c'est-à-dire que, par convention :

$$\forall n \in \mathbb{Z}, \quad -\infty < n < +\infty ; \quad n + (+\infty) = +\infty ; \quad n + (-\infty) = -\infty ; \\ -\infty \leq -\infty ; \quad +\infty \leq +\infty$$

mais il reste entendu que la somme $(+\infty) + (-\infty)$ n'est pas définie.

DÉFINITION VIII.1.2

Soit $P \in \mathcal{P}$ et $F \in K(X)^*$. On appelle **P -valuation de F** , et on note $\text{val}_P(F)$, l'unique entier $\alpha_P \in \mathbb{Z}$ exposant de P dans la décomposition en facteurs irréductibles de F ; on convient que $\text{val}_P(0) = +\infty$.

Cette définition nous montre immédiatement que, lorsque $F \in K[X]$, $\text{val}_P(F)$ est la P -valuation qui a été définie au § VII.1 pour les polynômes : la fonction val_P prolonge à la K -algèbre $K(X)$ la P -valuation des polynômes. En utilisant les propriétés déjà connues, on obtient facilement :

THÉORÈME VIII.1.3

Soit $P \in \mathcal{P}$. Pour tous $F \in K(X)$ et $G \in K(X)$, on a :

(I) $\text{val}_P(F + G) \geq \min(\text{val}_P(F), \text{val}_P(G))$

et, si $\text{val}_P(F) \neq \text{val}_P(G)$ il y a égalité

(II) $\text{val}_P(FG) = \text{val}_P(F) + \text{val}_P(G)$.

La remarque 3 montre que, si N/D est un représentant irréductible de $F \in K(X)^*$, alors :

si $\text{val}_P(F) \geq 0$, c'est la P -valuation de N .

et si $\text{val}_P(F) < 0$, $-\text{val}_P(F)$ est la P -valuation de D .

On retrouve bien, dans chaque cas, $\text{val}_P(F) = \text{val}_P(N) - \text{val}_P(D)$, en conformité avec le théorème VIII.1.3.

D'après le théorème VIII.1.1, on voit que si $F \in K(X)^*$, la différence $\deg(A) - \deg(B) \in \mathbb{Z}$ ne dépend pas du choix d'un représentant A/B de F , mais seulement de F . Cela conduit à poser :

DÉFINITION VIII.1.3

Soit $F \in K(X)^*$. On appelle **degré de F** , et on note $\deg(F)$, l'entier rationnel d tel que $d = \deg(A) - \deg(B)$ pour tout représentant A/B de F . On convient que $\deg(0) = -\infty$.

Le degré des fractions rationnelles prolonge le degré des polynômes, ce qui conduit immédiatement aux propriétés suivantes :

THÉORÈME VIII.1.4

$$\left\{ \begin{array}{l} \text{Pour tous } F \in K(X) \text{ et } G \in K(X), \text{ on a :} \\ \text{(I)} \quad \deg(F + G) \leq \max(\deg(F), \deg(G)) \\ \text{et, si } \deg(F) \neq \deg(G), \text{ il y a égalité} \\ \text{(II)} \quad \deg(FG) = \deg(F) + \deg(G). \end{array} \right.$$

Zéros et pôles

DÉFINITION VIII.1.4

$\left\{ \begin{array}{l} \text{Soit } F \in K(X)^*. \text{ On appelle } \mathbf{zéro de } F \text{ tout } a \in K \text{ tel que} \\ \text{val}_{X-a}(F) \geq 1, \text{ et } \mathbf{pôle de } F \text{ tout } a \in K \text{ tel que } \text{val}_{X-a}(F) \leq -1. \\ \text{L'entier } |\text{val}_{X-a}(F)| \text{ est appelé, selon le cas, la } \mathbf{multiplicité du zéro} \\ \text{ } a \text{ (resp. du pôle } a) \text{ dans } F. \end{array} \right.$

Les zéros (resp. les pôles) de multiplicité 1 sont dits *simples*, ceux de multiplicité 2 sont dits *doubles*, etc.

Soit N/D un représentant irréductible de $F \in K(X)^*$. D'après l'étude qui précède la définition VIII.1.3, **les zéros de F sont les zéros de N , et les pôles de F sont les zéros de D .** Si $a \in K$ est un zéro (resp. un pôle) de F , sa **multiplicité dans F est celle dans N (resp. celle comme zéro de D).**

En particulier, F n'a qu'un nombre *fini* de zéros et de pôles, et si $F \in K[X]$, F n'a pas de pôle dans K .

Un cas particulier très important est celui où le *corps K est algébriquement clos*. Dans ce cas, $F \in K(X)^*$ s'écrit $F = u \frac{N}{D}$, avec $u \in K^*$,

$$N = \prod_{a \in \mathcal{P}} (X - a)^{\nu_a}, \quad D = \prod_{a \in \mathcal{N}} (X - a)^{\mu_a},$$

en notant \mathcal{P} l'ensemble des zéros de F , \mathcal{N} l'ensemble des pôles, ν_a la multiplicité du zéro a , et μ_a celle du pôle a . On voit ici que $F \in K[X]$ ssi F n'a pas de pôle. De plus, $\frac{uN}{D}$ est une forme irréductible de F , et on a :

$$\deg(F) = \sum_{a \in \mathcal{P}} \nu_a - \sum_{a \in \mathcal{N}} \mu_a = \sum_{a \in K} \text{val}_{X-a}(F).$$

Un représentant A/B de $F \in K(X)^*$ est irréductible ssi A et B n'ont *aucune racine commune*.

Exercice 1 : Le corps de base est \mathbb{C} . Suivant les valeurs de m et n dans \mathbb{N}^* , et de $\lambda \in \mathbb{C}$, donner la forme irréductible de $\frac{X^m + \lambda^m}{X^n + \lambda^n} \in \mathbb{C}(X)$.

Exercice 2 : Le corps de base est \mathbb{C} . Suivant les valeurs de $n \in \mathbb{N}^*$, mettre sous forme irréductible la fraction

$$F = \frac{(X^{5^n} + X^{2^n} - 2)(X^3 + (2 - 4\lambda^2)X - 4\lambda)}{(X^3 - 1)(X^4 + 4)}.$$

Exercice 3 : Vérifier l'égalité $\frac{X^2 + X\sqrt{3} + 1 - i\sqrt{3}}{X^2 + 1} = \frac{X^2 + 2X\sqrt{3} + 4}{X^2 + X\sqrt{3} + 1 + i\sqrt{3}}$ dans $\mathbb{C}(X)$.

Exercice 4 : Mettre sous forme irréductible les fractions de $\mathbb{C}(X)$

a) $\frac{X^3 + (1 - i)X^2 + (1 - i)X - i}{X^3 - iX^2 + X - i};$

b) $\frac{X^3 - 3X + 2}{X^4 - 5X^2 + 4}.$

Exercice 5 : On suppose que K est le corps des fractions d'un anneau intègre A . Montrer que $K(X)$ est le corps des fractions de l'anneau $A[X]$.

Exercice 6 : On fixe $P \in \mathcal{J}$ ($P \in K[X]$, P irréductible et normalisé). Montrer que l'ensemble $\mathcal{F}_P = \{F \in K(X) \mid \text{val}_P(F) \geq 0\}$ est une sous- K -algèbre de $K(X)$, et que \mathcal{F}_P possède un et un seul idéal maximal \mathfrak{M}_P (cf. § III.7, exercice 3), qui est $\mathfrak{M}_P = \{F \in K(X) \mid \text{val}_P(F) \geq 1\}$, et que cet idéal est principal, engendré par P . Enfin démontrer que l'anneau quotient $\mathcal{F}_P/\mathfrak{M}_P$ est canoniquement isomorphe au corps $K[X]/PK[X]$ (cf. théorème VII.7.2).

Exercice 7 : Montrer que l'ensemble $\mathcal{F}_\infty = \{F \in K(X) \mid \deg(F) \leq 0\}$ est un sous-anneau du corps $K(X)$, qui possède un unique idéal maximal $\mathfrak{M}_\infty = \{F \in K(X) \mid \deg(F) \leq -1\}$. Montrer que l'anneau quotient $\mathcal{F}_\infty/\mathfrak{M}_\infty$ est canoniquement isomorphe au corps K .

Exercice 8 : a) Si $F \in K(X)$, on pose $\text{val}_\infty(F) = -\deg(F)$. Vérifier

$$\text{val}_\infty(FG) = \text{val}_\infty(F) + \text{val}_\infty(G) \text{ et } \text{val}_\infty(F + G) \geq \min(\text{val}_\infty(F), \text{val}_\infty(G))$$

pour F et G dans $K(X)^*$.

b) Soit $\varphi : K(X) \rightarrow \mathbb{Z}$ telle que $\varphi(0) = +\infty$, $\varphi(\lambda) = 0$ pour $\lambda \in K^*$,

$$\varphi(F) \in \mathbb{Z} \text{ pour } F \neq 0, \quad \text{Im}(\varphi) \neq \{0\} \quad \text{et} \quad (\forall F \in K(X), \forall G \in K(X))$$

$$\varphi(FG) = \varphi(F) + \varphi(G), \quad \varphi(F + G) \geq \min(\varphi(F), \varphi(G)).$$

Montrer que $A = \{F \in K(X) \mid \varphi(F) \geq 0\}$ est un sous-anneau de $K(X)$, et que $\mathfrak{M} = \{F \in K(X) \mid \varphi(F) \geq 1\}$ est un idéal maximal de A . Prouver que $A \setminus \mathfrak{M}$ est l'ensemble des éléments inversibles de A et que $F \in K(X) \setminus A \Rightarrow \frac{1}{F} \in \mathfrak{M}$.

c) Si $X \in A$, démontrer que $\mathfrak{M} \cap K[X] = PK[X]$ pour un $P \in \mathcal{J}$, puis, qu'il existe $k \in \mathbb{N}^*$ tel que $\varphi(F) = k \text{val}_P(F)$ pour tout $F \in K(X)$.

d) Si $X \notin A$, montrer qu'il existe $k \in \mathbb{N}^*$ tel que $\varphi(F) = k \text{val}_\infty(F)$ pour tout $F \in K(X)$.

Exercice 9 : Soit $F \in K(X)$. On suppose trouvés P_1, P_2, \dots, P_n dans $K[X]$ tels que $F^n + P_1 F^{n-1} + \dots + P_n = 0$ avec $n \geq 1$. Montrer que $F \in K[X]$.

Exercice 10 : Soit $F, G_1, G_2, \dots, G_n \in \mathbb{C}(X)$ avec $n \geq 1$. On suppose que :

$$F^n + G_1 F^{n-1} + \dots + G_n = 0.$$

Montrer que l'ensemble des pôles de F est contenu dans l'union des ensembles de pôles des G_i .

Exercice 11 : Soit \mathcal{E} une partie non vide de \mathcal{J} dans $K[X]$. On considère l'ensemble $V_{\mathcal{E}} = \{F \in K(X) \mid (\forall P \in \mathcal{E}) \text{val}_P(F) \geq 0\}$.

- a) Vérifier que $V_{\mathcal{J}}$ est une sous- K -algèbre de $K(X)$ qui contient $K[X]$. Quels sont ses éléments inversibles ?
- b) Démontrer que $V_{\mathcal{J}}$ est un anneau principal. Préciser ses éléments irréductibles.
- c) Montrer directement, sans utiliser b), que l'anneau $V_{\mathcal{J}}$ est factoriel.

Exercice 12 : Chercher $F \in K(X)$ telle que $F^2 = X$; conclure : $K(X)$ n'est jamais algébriquement clos.

§ VIII.2 DÉCOMPOSITION EN ÉLÉMENTS SIMPLES

Nous continuons à noter \mathcal{J} l'ensemble des polynômes irréductibles et normalisés dans $K[X]$.

LEMME 1

Soit $D \in K[X]$ un polynôme non constant de degré d , $\alpha \in \mathbb{N}^*$ et $F \in K[X]$. Alors il existe une (et une seule) suite $(G_0, G_1, \dots, G_{\alpha-1}, R)$ de polynômes telle que $G_i \in K_{d-1}[X]$ pour tout i , et

$$(1) \quad F = G_0 + G_1 D + \dots + G_{\alpha-1} D^{\alpha-1} + RD^{\alpha}.$$

Démonstration :

Si F est nul et de la forme (1), on reconnaît que la division euclidienne du second membre par D donnant G_0 comme reste, il s'ensuit que $G_0 = 0$; par récurrence on aura $G_1 = G_2 = \dots = G_{\alpha-1} = 0$ et enfin $R = 0$, d'où l'unicité. L'existence se montre aussi par récurrence, sur l'entier α , à partir de $\alpha = 1$ pour laquelle on reconnaît la division euclidienne de F par D . Supposons alors l'existence prouvée pour tout polynôme F avec l'entier $\alpha - 1 \geq 1$, la division de F par D donne $F = G_0 + DF_1$, avec $G_0 \in K_{d-1}[X]$ et $F_1 \in K[X]$. Mais par hypothèse il existe $G_1, \dots, G_{\alpha-1}$ dans $K_{d-1}[X]$ tels que

$$F_1 = G_1 D + \dots + G_{\alpha-1} D^{\alpha-2} + RD^{\alpha-1} \quad (R \in K[X]),$$

d'où finalement F sous la forme (1). ■

LEMME 2

Soit A, B_1, B_2, \dots, B_n dans $K[X] \setminus \{0\}$ avec A, B_1, B_2, \dots, B_n deux à deux premiers entre eux et $n \geq 2$. Alors il existe des polynômes A_1, A_2, \dots, A_n tels que $\frac{A}{B_1 B_2 \dots B_n} = \sum_{i=1}^n \frac{A_i}{B_i}$, chaque fraction $\frac{A_i}{B_i}$ étant irréductible.

Démonstration :

Une récurrence immédiate sur n , compte tenu du fait que chaque B_i est premier avec le produit de tous les autres B_j , montre qu'il suffit de prouver ce lemme pour $n = 2$.

Soit donc U_1, U_2 dans $K[X]$ tels que

$$(2) \quad U_1 B_1 + U_2 B_2 = 1$$

(théorème de Bezout). Alors $\frac{A}{B_1 B_2} = \frac{A(U_1 B_1 + U_2 B_2)}{B_1 B_2} = \frac{AU_2}{B_1} + \frac{AU_1}{B_2}$.

Posant $A_1 = AU_2$ et $A_2 = AU_1$, on voit bien que A_1 est premier B_1 (puisque A et U_2 le sont) et que A_2 est premier avec B_2 , d'où le lemme. ■

THÉORÈME VIII.2.1

La famille constituée d'une part par les $(X^n)_{n \geq 0}$ et d'autre part par les $\left(\frac{X^k}{P^l}\right)$ où $P \in \mathcal{J}$, $l \in \mathbb{N}^*$, $k \in \mathbb{N}$ et $k < \deg(P)$ est une base du K -espace vectoriel $K(X)$. Autrement dit : pour tout $F \in K(X)$, il existe une et une seule famille à support fini $(E, (A_{l,P})_{P \in \mathcal{J}, l \in \mathbb{N}^*})$ d'éléments de $K[X]$ telle que $(\forall P \in \mathcal{J}) A_{l,P} \in K_{\deg(P)-1}[X]$ et

$$(3) \quad F = E + \sum_{P \in \mathcal{J}} \left(\sum_{l \in \mathbb{N}^*} \frac{A_{l,P}}{P^l} \right).$$

Démonstration :

Unicité : il s'agit de prouver que si $F = 0$ dans une relation du type (3), alors $E = 0$ et tous les $A_{l,P}$ sont nuls. Considérons une partie finie \mathcal{S} de \mathcal{J} telle que $A_{l,P} = 0$ pour tout l et tout $P \in \mathcal{J} \setminus \mathcal{S}$. Pour $P \in \mathcal{S}$, soit n_P le plus petit des entiers $k \in \mathbb{N}^*$ tels que $A_{l,P} = 0$ pour $l \geq k$ (les n_P et \mathcal{S} existent parce que la famille $(A_{l,P})$ est à support fini). Il s'agit de montrer que tous les $(n_P)_{P \in \mathcal{S}}$ valent 1. Par l'absurde, supposons qu'il existe un $P \in \mathcal{S}$ tel que $n_P \geq 2$, d'où $A_{n_P-1,P} \neq 0$. Multiplions les deux membres de (3) par $P^{n_P-1} \prod_{Q \in \mathcal{S}, Q \neq P} P^{n_Q}$. On obtient une relation du type

$$(4) \quad A_{n_P-1,P} \times B + C P^{n_P-1} = 0 \quad \text{avec} \quad B = \prod_{Q \in \mathcal{S}, Q \neq P} Q^{n_Q}.$$

Mais $A_{n_P-1,P} \neq 0$ et il est premier avec P puisque P est irréductible et que $\deg(A_{n_P-1,P}) < \deg(P)$. D'autre part le facteur P ne figure pas dans B , ce qui conduit manifestement dans (4) à une contradiction puisque $n_P - 1 \geq 1$. On a bien prouvé que tous les $A_{l,P}$ sont nuls, et finalement $E = 0$.

Existence : Soit $F \in K(X) \setminus \{0\}$. Si F est un polynôme

$E = F$ et tous les $A_{l,P}$ nuls, d'où (3). Si $F \notin K[X]$, soit \mathcal{S} l'ensemble non vide des $P \in \mathcal{J}$ tels que $\text{val}_P(F) < 0$. On a $F = \frac{A}{\prod_{P \in \mathcal{S}} P^{\alpha_P}}$ avec

$\alpha_P = -\text{val}_P(F)$ pour $P \in \mathcal{S}$, et avec $A \in K[X] \setminus \{0\}$ et A premier avec $B = \prod_{P \in \mathcal{S}} P^{\alpha_P}$. Le lemme 2 permet d'écrire F sous la forme $F = \sum_{P \in \mathcal{S}} \frac{A_P}{P^{\alpha_P}}$, avec A_P premier avec P pour tout $P \in \mathcal{S}$.

Pour chaque $P \in \mathcal{S}$, le lemme 1 permet d'écrire

$$\frac{A_P}{P^{\alpha_P}} = \frac{G_{0,P}}{P^{\alpha_P}} + \frac{G_{1,P}}{P^{\alpha_P-1}} + \cdots + \frac{G_{\alpha_P-1,P}}{P} + R_P,$$

avec $R_P \in K[X]$ et $G_{i,P} \in K_{\deg(P)-1}[X]$ pour $i \leq \alpha_P - 1$.

Finalement, en rassemblant dans E la somme des R_P pour $P \in \mathcal{S}$, et en posant $A_{l,P} = 0$ si $l > \alpha_P$ ou si $P \notin \mathcal{S}$, $A_{l,P} = G_{\alpha_P-l,P}$ si $P \in \mathcal{S}$ et $l \leq \alpha_P$, on a bien écrit F sous la forme (3). Il est important de remarquer que $G_{0,P} \neq 0$ pour $P \in \mathcal{S}$, car A_P est premier avec P . ■

Par définition, la décomposition de $F \in K(X)$ sous la forme (3) du théorème VIII.2.1 s'appelle **décomposition en éléments simples de F sur K** .

Le polynôme E dans l'écriture (3) s'appelle la **partie entière** de F . Pour chaque $P \in \mathcal{J}$, la somme $\sum_{l \geq 1} \frac{A_{l,P}}{P^l}$ s'appelle la **partie P -fractionnaire** de F qu'on peut noter par exemple $\mathcal{F}_P(F)$.

La fraction F se réduit à un polynôme ssi elle est égale à sa partie entière. Si F n'est pas un polynôme, soit $\frac{N}{D}$ sa forme irréductible avec $D = \prod_{P \in \mathcal{S}} P^{\alpha_P}$.

Il résulte de la démonstration du théorème VIII.2.1 que les $P \in \mathcal{J}$ tels que la partie P -fractionnaire $\mathcal{F}_P(F)$ soit non nulle sont les $P \in \mathcal{S}$. De plus, pour chaque $P \in \mathcal{S}$, on a

$$(5) \quad \mathcal{F}_P(F) = \sum_{l \geq 1} \frac{A_{l,P}}{P^l} \quad \text{avec} \quad A_{l,P} = 0 \quad \text{si} \quad l > \alpha_P \quad \text{et} \quad A_{\alpha_P,P} \neq 0$$

comme nous l'avons remarqué ci-dessus. Enfin, en multipliant les deux membres de (3) par D , on obtient évidemment $N = ED + T$ avec $\deg(T) < \deg(D)$, ce qui prouve que **la partie entière E n'est autre que le quotient dans la division euclidienne de N par D** et ce qui achève de déterminer avec précision les différents termes figurant dans la formule (3).

Remarque 1 : Dans le cas où $F = \frac{N}{D}$ n'est pas un polynôme, ayant mis D sous la forme $\prod_{P \in \mathcal{S}} P^{\alpha_P}$, si l'on se contente d'écrire F sous la forme (non

unique) $F = \sum_{P \in \mathcal{S}} \frac{A_P}{P^{\alpha_P}}$, on constate que la partie P -fractionnaire de F est égale, pour tout $P \in \mathcal{S}$, à celle de $\frac{A_P}{P^{\alpha_P}}$.

Parties polaires. Résidus

Reprenons $F \in K(X) \setminus \{0\}$. Les $a \in K$ tels que la partie $(X - a)$ -fractionnaire de F soit non nulle sont évidemment les **pôles** de F . Si a est un pôle, de multiplicité α , la partie $(X - a)$ -fractionnaire $\mathcal{F}_{X-a}(F)$ s'appelle la **partie polaire** de F relative au pôle a . D'après l'étude précédente, on a :

$$(6) \quad \mathcal{F}_{X-a}(F) = \frac{C_{1,a}}{X-a} + \frac{C_{2,a}}{(X-a)^2} + \cdots + \frac{C_{\alpha,a}}{(X-a)^\alpha}, \quad \text{avec } C_{i,a} \in K$$

pour tout $i \in \llbracket 1, \alpha \rrbracket$ et $C_{\alpha,a} \neq 0$.

DÉFINITION VIII.2.1

Soit a un pôle de $F \in K(X) \setminus \{0\}$. On appelle **résidu de F en a** , et on note $\text{Res}(F, a)$, le scalaire $C_{1,a} \in K$, coefficient de $\frac{1}{X-a}$ dans le développement de la partie polaire de F en le pôle a sous la forme (6).

Si le calcul de ce résidu n'est pas toujours facile (cas d'un pôle multiple), il est un cas où ce calcul est particulièrement aisé :

THÉORÈME VIII.2.2

Soit $\frac{N}{D}$ la forme irréductible d'une fraction $F \in K(X) \setminus \{0\}$. Si a est un **pôle simple** de F , alors $D'(a) \neq 0$ et $\text{Res}(F, a) = \frac{N(a)}{D'(a)}$.

Démonstration :

Par hypothèse on a : $D(X) = (X - a) D_1(X)$ avec $D_1 \in K[X]$ et $D_1(a) \neq 0$. Par dérivation, on en déduit :

$$D'(X) = (X - a) D_1'(X) + D_1(X), \quad \text{d'où } D'(a) = D_1(a) \neq 0.$$

Ecrivons $F(X) = \frac{C_1}{X-a} + G(X)$, où $C_1 = \text{Res}(F, a)$ et $G \in K(X)$. La partie polaire de F en le pôle a se réduit à $\frac{C_1}{X-a}$ puisque a est pôle simple. Donc a n'est pas pôle de G (cf. remarque 1). Multipliant

prenant la valeur en a , on obtient : $N(a) = C_1 D_1(a)$, car le polynôme $DG = (X - a) D_1 G$ s'annule en A . D'où, puisque $D_1(a) = D'(a) \neq 0$: $C_1 = \frac{N(a)}{D'(a)}$. ■

Lorsque le dénominateur N de la forme irréductible de F est **dissocié sur K** (ce qui est toujours le cas lorsque K est *algébriquement clos*), les seules parties fractionnaires dans la décomposition en éléments simples sont ses parties polaires, ce qui nous incite à pousser plus loin leur étude en nous dotant de nouveaux moyens.

Division suivant les puissances croissantes

THÉORÈME VIII.2.3

Soit A et B dans $K[X]$ avec $\text{val}(B) = 0$ (i.e. $B(0) \neq 0$). Pour tout $n \in \mathbb{N}$, il existe un unique couple $(Q_n, R_n) \in K_n[X] \times K[X]$ tel que

$$(7) \quad A = BQ_n + X^{n+1} R_n.$$

Démonstration :

Pour démontrer l'unicité prenons $A = 0$; X^{n+1} divise BQ_n et il est premier avec B puisque $B(0) \neq 0$. Donc il divise Q_n , mais Q_n étant de degré $\leq n$, il en résulte que $Q_n = 0$, et enfin $R_n = 0$.

Démontrons l'existence par récurrence sur n . Pour $n = 0$, posons $Q_0 = \frac{A(0)}{B(0)}$. Alors $Q_0 \in K$ et $A - BQ_0$ est divisible par X , d'où $R_0 \in K[X]$ tel que :

$$(8) \quad A = BQ_0 + XR_0.$$

Supposons l'existence du couple (Q_n, R_n) acquise pour n et utilisons (8) avec $A = R_n$. Il vient $R_n = Bq_{n+1} + XR_{n+1}$ où $q_{n+1} \in K$ et $R_{n+1} \in K[X]$, d'où $A = BQ_{n+1} + X^{n+2} R_{n+1}$ en posant

$$Q_{n+1} = Q_n + q_{n+1} X^{n+1} \in K_{n+1}[X],$$

et $R_{n+1} \in K[X]$. La propriété étudiée est bien héréditaire, d'où le théorème. ■

Cette démonstration fournit un algorithme de calcul de Q_n et R_n tout à fait analogue à celui de la division euclidienne.

Pour $n \in \mathbb{N}$ donné, l'écriture (7) s'appelle **division suivant les puissances croissantes de A par B à l'ordre n** . Dans cette division, Q_n est le **quotient à l'ordre n** et R_n le **reste à l'ordre n** . On peut utiliser une disposition pratique analogue à celle de la division euclidienne, mais en écrivant le :

A et de B par ordre de degré croissant, ce qui explique le nom donné à cette division.

Donnons un exemple de cette disposition en cherchant le développement limité à l'ordre 7 de la fonction $X \mapsto \operatorname{tg} X$ au voisinage de 0 (cf. cours d'Analyse).

Exemple 1 : $K = \mathbb{C}$, $n = 7$,

$$A = X - \frac{1}{6}X^3 + \frac{1}{120}X^5 - \frac{1}{5040}X^7, \quad B = 1 - \frac{1}{2}X^2 + \frac{1}{24}X^4 - \frac{1}{720}X^6.$$

X	$-\frac{1}{6}X^3 + \frac{1}{120}X^5 - \frac{1}{5040}X^7$ ($= -A$)	$1 - \frac{1}{2}X^2 + \frac{1}{24}X^4 - \frac{1}{720}X^6$ ($= B$)
$(A - BQ_1)$	$\frac{1}{3}X^3 - \frac{1}{30}X^5 + \frac{1}{840}X^7$	$X + \frac{1}{3}X^3 + \frac{2}{15}X^5 + \frac{17}{315}X^7$ ($= Q_7$)
$(Q_1 = Q_2, R_1 = XR_2)$		
$(A - BQ_3)$	$\frac{2}{15}X^5 - \frac{4}{315}X^7$	
$(A - BQ_5)$	$\frac{17}{315}X^7$	
$(A - BQ_7 = X^8R_7)$	$\frac{331}{15120}X^9 - \frac{13}{6300}X^{11} + \frac{17}{226800}X^{13}$	

Le quotient à l'ordre 7 est $Q_7 = X + \frac{1}{3}X^3 + \frac{2}{15}X^5 + \frac{17}{315}X^7$, et le reste à l'ordre 7 est $R_7 = \frac{331}{15120}X - \frac{13}{6300}X^3 + \frac{17}{226800}X^5$, mais nous avons entouré de pointillés les termes qui seraient inutiles si l'on n'avait en vue que le calcul du *quotient*, ce qui arrive fréquemment.

Application au calcul d'une partie polaire

THÉORÈME VIII.2.4

Soit a un pôle de multiplicité $\alpha \geq 2$ de la fraction $F \in K(X) \setminus \{0\}$, dont la forme irréductible est $\frac{N}{D}$, avec $D = (X - a)^\alpha Q$, $Q(a) \neq 0$.

Posons $X - a = Y$, où Y est considéré comme une nouvelle indéterminée sur K . Divisons $N(a + Y)$ par $Q(a + Y)$, suivant les puissances croissantes de Y , à l'ordre $\alpha - 1$:

$$(9) \quad N(a + Y) = Q(a + Y) \times (c_0 + c_1 Y + \cdots + c_{\alpha-1} Y^{\alpha-1}) + Y^\alpha R(Y),$$

où $c_i \in K$ et $R \in K[Y]$. Alors la partie polaire de F relative au pôle a est

$$\mathcal{P}_{X-a}(F) = \frac{c_0}{(X-a)^\alpha} + \frac{c_1}{(X-a)^{\alpha-1}} + \cdots + \frac{c_{\alpha-1}}{X-a}.$$

Démonstration :

Revenant à X dans (9) et divisant les deux

$(X - a)^\alpha Q = D$, on obtient $F(X) = \frac{c_0}{(X - a)^\alpha} + \dots + \frac{c_{\alpha-1}}{X - a} + \frac{R(X - a)}{Q(X)}$; mais comme $Q(a) \neq 0$, a n'est pas pôle de la fraction $\frac{R(X - a)}{Q(X)}$, et le théorème résulte donc de la remarque 1. ■

Exemple 2 : Le corps de base est \mathbb{C} . Si $n \in \mathbb{N}$, $n \geq 2$, décomposer en éléments simples $F = \frac{1}{X^n - 1}$.

F est sous forme irréductible. Elle possède n pôles simples qui sont les éléments de μ_n (racines n -ièmes de l'unité). La partie entière est nulle, et le théorème VIII.2.2 donne immédiatement :

$$F = \sum_{\zeta \in \mu_n} \frac{A_\zeta}{X - \zeta} \quad \text{avec} \quad A_\zeta = \frac{1}{n\zeta^{n-1}} = \frac{\zeta}{n}, \text{ d'où la réponse.}$$

Exemple 3 : Le corps de base est \mathbb{C} . Décomposer en éléments simples : $F = \frac{X^7}{(X - 1)^3(X^3 + 1)}$. En déduire une décomposition de F dans $\mathbb{R}[X]$.

Solution : la fraction est donnée sous forme irréductible. Commençons par déterminer sa partie entière en divisant X^7 par

$$(X^3 - 3X^2 + 3X - 1)(X^3 + 1) = X^6 - 3X^5 + \dots,$$

ce qui donne, si nous ne désirons pas utiliser le reste, $E = X + 3$ sans qu'on ait besoin de développer entièrement le dénominateur ; F possède trois pôles simples qui sont $\omega_1 = -1$, $\omega_2 = -j$ et $\omega_3 = -j^2$ et un pôle triple : 1. Il peut être avantageux de commencer par chercher la partie polaire relative au pôle triple qui est de la forme $\frac{c_0}{(X - 1)^3} + \frac{c_1}{(X - 1)^2} + \frac{c_2}{X - 1}$ et s'obtient

« globalement » en utilisant le théorème VIII.2.3, ce qui amène à poser $X - 1 = Y$ et à diviser, suivant les puissances croissantes de Y , le polynôme $A = (1 + Y)^7$ par le polynôme $Q = 1 + (1 + Y)^3$, à l'ordre 2, ce qui ne nécessite que la connaissance des termes de degré ≤ 2 dans A et dans Q . Voici cette division

$$\begin{array}{r|l} 1 + 7Y + 21Y^2 + \dots & 2 + 3Y + 3Y^2 + \dots \\ \frac{11}{2}Y + \frac{39}{2}Y^2 + \dots & \frac{1}{2} + \frac{11}{4}Y + \frac{45}{8}Y^2 \\ \frac{45}{4}Y^2 + \dots & \end{array}$$

Il est parfois intéressant de pouvoir utiliser le reste de cette division, mais ici ce serait inutilement long car les autres pôles étant simples, la

correspondante se calcule très simplement en utilisant le théorème VIII.2.2.

On a les résidus $r_k = \frac{(\omega_k)^7}{D'(\omega_k)}$. Mais, compte tenu de $\omega_k^3 = -1$, cela donne

$$D'(\omega_k) = 3(\omega_k - 1)^2 (\omega_k^3 + 1) + (\omega_k - 1)^3 3 \omega_k^2 = 3 \omega_k^2 (\omega_k - 1)^3,$$

d'où $r_k = \frac{\omega_k^5}{3(\omega_k - 1)^3} = \frac{-1}{3} \frac{\omega_k^2}{(\omega_k - 1)^3}$, ce qui donne $r_1 = \frac{+1}{24}$, puis pour les autres $\omega_k^2 - \omega_k + 1 = 0$, d'où $r_k = -\frac{1}{3} \frac{1}{\omega_k^4} = -\frac{1}{3} \omega_k^2$, ce qui donne $r_2 = -\frac{1}{3} j^2$, $r_3 = -\frac{1}{3} j$. Finalement,

$$F = X + 3 + \frac{1}{2(X-1)^3} + \frac{11}{4(X-1)^2} + \frac{45}{8(X-1)} + \frac{1}{24(X+1)} - \frac{1}{3} \left[\frac{j^2}{X+j} + \frac{j}{X+j^2} \right].$$

Pour avoir une décomposition de F sur \mathbb{R} il suffit de *regrouper les parties polaires complexes conjuguées*, ce qui donne ici en regroupant les deux derniers termes $\frac{+X+1}{3(X^2-X+1)}$. Il reste à vérifier le résultat obtenu en donnant à X une ou deux valeurs commodes, par exemple ici $X = 0$:

$$0 = 3 - \frac{1}{2} + \frac{11}{4} - \frac{45}{8} + \frac{1}{24} + \frac{1}{3}, \text{ ce qui est satisfaisant.}$$

Exemple 4 : Le corps de base est \mathbb{R} . Décomposer en éléments simples la fraction $F = \frac{X^4 + X^2 + 1}{(X^3 - 1)^3}$.

Il faut se méfier car la fraction n'est pas donnée sous forme irréductible. La première précaution consiste à la « simplifier » :

$$F = \frac{X^2 - X + 1}{(X-1)^3(X^2 + X + 1)^2} = \frac{N}{D}.$$

On voit déjà que la partie entière est nulle. Le dénominateur D possède dans $\mathbb{R}[X]$ un facteur double $(X^2 + X + 1)^2$, mais $X^2 + X + 1$ est irréductible. Nous commençons comme dans l'exemple précédent à rechercher la partie polaire relative au pôle triple 1. En posant $X = 1 + Y$, il y a lieu d'effectuer la division de $A = 1 + Y + Y^2$ par

$$Q = (3 + 3Y + Y^2)^2 = 9 + 18Y + 15Y^2 + \dots,$$

ce qui donne

$$\begin{array}{r|l}
 1 + Y + Y^2 & 9 + 18Y + 15Y^2 \\
 - Y - \frac{2}{3}Y^2 & \frac{1}{9} - \frac{1}{9}Y + \frac{4}{27}Y^2 \\
 \frac{4}{3}Y^2 &
 \end{array}$$

L'idée la plus simple est peut-être de retrancher cette partie polaire $\frac{1/9}{(X-1)^3} - \frac{1/9}{(X-1)^2} + \frac{4/27}{X-1}$ à la fraction F pour obtenir une fraction F_1 qui ne possède plus le pôle 1. Cela donne

$$\begin{aligned}
 & \frac{X^2 - X + 1 - \left(+\frac{4}{27}X^2 - \frac{11}{27}X + \frac{10}{27} \right) (X^4 + 2X^3 + 3X^2 + 2X + 1)}{(X-1)^3 (X^2 + X + 1)^2} = \\
 & = \frac{-4X^6 + 3X^5 + 5X^3 + 15X^2 - 36X + 17}{27(X-1)^3 (X^2 + X + 1)^2}
 \end{aligned}$$

et on a la satisfaction, en effectuant la division du numérateur par $(X-1)^3 = X^3 - 3X^2 + 3X - 1$ de voir que le reste est bien nul. Quant au quotient, on trouve $-4X^3 - 9X^2 - 15X - 17$. Il reste donc à décomposer $F_1 = -\frac{4X^3 + 9X^2 + 15X + 17}{27(X^2 + X + 1)^2}$ sur \mathbb{R} . Une simple division euclidienne

$$\begin{array}{r|l}
 4X^3 + 9X^2 + 15X + 17 & X^2 + X + 1 \\
 5X^2 + 11X & 4X + 5 \\
 6X + 12 &
 \end{array}$$

donne les deux éléments simples :

$$F_1 = \frac{-1}{27} \left(\frac{4X + 5}{X^2 + X + 1} + \frac{6X + 12}{(X^2 + X + 1)^2} \right).$$

Finalement :

$$\begin{aligned}
 F &= \frac{X^2 - X + 1}{(X-1)^3 (X^2 + X + 1)^2} = \frac{1/9}{(X-1)^3} - \frac{1/9}{(X-1)^2} + \\
 &+ \frac{4/27}{X-1} - \frac{1}{27} \frac{6X + 12}{(X^2 + X + 1)^2} - \frac{1}{27} \frac{4X + 5}{X^2 + X + 1}
 \end{aligned}$$

et on n'oublie pas de vérifier que $F(0) = -1 = \frac{-1}{9} - \frac{1}{9} - \frac{4}{27} - \frac{12}{27} - \frac{5}{27}$, ce qui est rassurant.

Une autre idée aurait pu être, après le calcul de la partie polaire relative au pôle 1, d'écrire

$$F = \frac{1/9}{(X-1)^3} - \frac{1/9}{(X-1)^2} + \frac{4/9}{X-1} + \frac{AX + B}{(X^2 + X + 1)^2} + \frac{CX + D}{X^2}$$

l'élément simple le plus facile à obtenir étant celui où le facteur irréductible $P = X^2 + X + 1$ est à la puissance la plus élevée. En effet, en multipliant les deux membres par D , et en calculant modulo $X^2 + X + 1$, il vient :

$X^2 - X + 1 \equiv (AX + B)(X - 1)^3$, soit $-2X \equiv (AX + B)(6X + 3)$,
soit encore $-2X \equiv (6B - 3A)X - 6A + 3B$, c'est-à-dire $3B - 6A = 0$
et $6B - 3A = -2$, ce qui redonne bien $A = \frac{-2}{9}$ et $B = \frac{-4}{9}$ (on notera
que le calcul serait exactement le même si l'on faisait $X = j$ après
multiplication par D , mais cela obligerait à considérer F comme une fraction
dans $\mathbb{C}(X)$, ce que l'énoncé semble interdire). Il ne reste plus alors que les
constantes C et D à déterminer, ce que l'on fait souvent (en anticipant sur le
§ VIII.3) en donnant deux valeurs convenablement choisies à l'indéterminée
 X , par exemple $X = 0$ et une valeur très pratique : $X \rightarrow \infty$ (après
multiplication des deux membres par X), ce qui donne $0 = \frac{4}{27} + C$, d'où C ,
et enfin $-1 = \frac{-1}{9} - \frac{1}{9} - \frac{4}{27} - \frac{4}{9} + D$ qui donne $D = \frac{-5}{27}$. On voit l'importance
de ne pas commettre d'erreur dans les calculs initiaux car elle se
répercuterait dans les calculs suivants.

Exemple 5 : Le corps de base est \mathbb{Q} . Décomposer en éléments simples :

$$F = \frac{N}{D} = \frac{X - 1}{(X^3 + X + 1)(X^2 + 1)^2}.$$

On constate que $E = 0$ et que la fraction est donnée sous forme irréductible. Dans $\mathbb{Q}[X]$ les polynômes $A = X^2 + 1$ et $B = X^3 + X + 1$ sont irréductibles (car de degrés respectifs 2 et 3, et sans racine dans \mathbb{Q}). On peut donc écrire la décomposition cherchée sous la forme :

$$F = \frac{aX + b}{(X^2 + 1)^2} + \frac{cX + d}{X^2 + 1} + \frac{\alpha X^2 + \beta X + \gamma}{X^3 + X + 1}.$$

On pourrait réduire au même dénominateur et identifier, mais ce n'est pas une bonne idée car il y aurait 7 inconnues à trouver, et résoudre un système linéaire de 7 équations à 7 inconnues n'est pas très pratique. Mieux vaut profiter du fait que $X^3 + X + 1 - X(X^2 + 1) = 1$, d'où en élevant au carré $B^2 - 2XBA + X^2A^2 = 1$, et par suite :

$$\begin{aligned} F &= \frac{N[B(B - 2XA) + X^2A^2]}{BA^2} \\ &= \frac{N(B - 2XA)}{A^2} + \frac{NX^2}{B} = \frac{N(1 - XA)}{A^2} + \frac{NX^2}{B}. \end{aligned}$$

Le premier terme a pour numérateur

$$(X - 1)(1 - XA) = (X - 1) - (X^2 - X)A$$

que l'on a intérêt à développer comme polynôme en A car $X^2 = A - 1$, d'où

$$N(1 - XA) = (X - 1) - (A - 1 - X)A = (X - 1) + (X + 1)A - A^2.$$

Quant au second terme, en effectuant la division euclidienne de $(X - 1)X^2$ par $X^3 + X + 1$, on obtient un quotient 1 et un reste $-X^2 - X - 1$. Finalement

$$F = \frac{X - 1}{A^2} + \frac{X + 1}{A} - 1 + 1 - \frac{X^2 + X + 1}{B},$$

d'où le résultat cherché

$$F = \frac{X - 1}{(X^2 + 1)^2} + \frac{X + 1}{X^2 + 1} - \frac{X^2 + X + 1}{X^3 + X + 1}$$

qu'il est très facile de contrôler.

Exercice 1 : Décomposer en éléments simples sur \mathbb{C} les fractions rationnelles suivantes :

- a) $F = \frac{1}{(X^2 - 1)^2}$ (méthode rapide : commencer par $\frac{1}{X^2 - 1}$ et élever au carré).
 b) $F = \frac{1}{(X - 1)(X - 2) \dots (X - n)}, n \in \mathbb{N}^* ;$
 c) $F = \frac{1}{(X - 1)^2 (X - 2)^2 \dots (X - n)^2}, n \in \mathbb{N}^* .$
 d) $\frac{1}{T_n(X)}$ et $\frac{1}{U_n(X)}$, où T_n et U_n sont les polynômes de Tchebychev (cf. § VII.4).
 e) $\frac{1}{E_n(X)}$ pour $n \in \{2, 3, 4, 5\}$, où E_n est un polynôme d'Euler (cf. § VII.5).
 f) $F = \frac{1}{X \prod_{k=1}^n (X^2 - k^2)} ;$ g) $F = \frac{1}{(X + 1)^7 - X^7 - 1} .$

Exercice 2 : Décomposer en éléments simples sur \mathbb{R} les fractions rationnelles suivantes :

- a) $F = \frac{X^2 - X + 1}{(X - 1)^3 (X^2 + X + 1)^2 (X^4 + 4)} ;$ b) $F = \frac{1}{(X^n - 1)^2} ;$
 c) $F = \frac{1}{(X^n + 1)^2} ;$ d) $F = \frac{X^m}{X^n + 1}, m \in \mathbb{N}^*, n \in \mathbb{N}^* ;$
 e) $\frac{1}{(X^{2n} - 1)^2} ;$ f) $\frac{X^5}{(X^4 - 1)^2} ;$ g) $\frac{X^2 + X + 4}{(X - 1)^4 (X^2 + X + 1)} .$

Exercice 3 : Décomposer en éléments simples sur \mathbb{R} les fractions rationnelles suivantes :

- a) $F = \frac{1}{X^m(1 - X)^n}, m \in \mathbb{N}^*, n \in \mathbb{N}^* ;$ b) $\frac{1}{(X^2 - 1)^n} ;$

$$\begin{aligned} c) & \frac{X^4 + 1}{(X-1)(X^2 - X + 1)^n}; & d) & \frac{X^5 - X^3 - X^2}{X^2 - 1}; \\ e) & \frac{4X^3}{(X^2 + 1)^2}; & f) & \frac{X^6 - X^2 + 1}{(X-1)^3}. \end{aligned}$$

Exercice 4 : Décomposer en éléments simples sur \mathbb{R} les fractions rationnelles suivantes :

$$\begin{aligned} a) & F = \frac{X^2 + X + 1}{(X^2 - 1)(X^2 + 1)}; & b) & F = \frac{X}{X^4 + X^2 + 1}; \\ c) & \frac{X^5 - X + 1}{(X^2 + 1)^n}; & d) & \frac{X^7 + 5}{(X^2 + X + 1)^2 (X + 2)^3}; \\ e) & \frac{X^5 + 64}{(X^2 + 2X + 4)^3}; & f) & \frac{1}{(X + 1)(X + 2)^6}. \end{aligned}$$

Exercice 5 : Décomposer en éléments simples sur $K = \mathbb{Z}/5\mathbb{Z}$ les fractions rationnelles suivantes :

$$\begin{aligned} a) & \frac{X - \bar{1}}{(X + \bar{1})^2 (X + \bar{2})}; & b) & \frac{\bar{1}}{(X^2 + \bar{1})(X^2 + \bar{2})}; \\ c) & \frac{X - \bar{2}}{(X^2 + \bar{4})(X^2 + \bar{3})}; & d) & \frac{\bar{4}X + \bar{2}}{X^3 + \bar{2}X^2 + \bar{4}X + \bar{3}}; \end{aligned}$$

et décomposer $\frac{\bar{2}X^5 + \bar{3}X^3 + \bar{6}X^2 + 4}{(X^2 + \bar{1})^2}$ sur $\mathbb{Z}/7\mathbb{Z}$.

Exercice 6 : On veut décomposer $F = \frac{1}{(X^3 - 1)^3}$ sur \mathbb{C} ; a) chercher la partie polaire relative au pôle 1 ; b) utiliser $F(X) = F(jX) = F(j^2X)$ pour achever.

Exercice 7 : Réduire au même dénominateur et simplifier les fractions suivantes :

$$\begin{aligned} a) & F = \sum_{\zeta \in \mu_n} \frac{1}{X - \zeta}; & b) & F = \sum_{\zeta \in \mu_n} \frac{1}{(X - \zeta)^2}; \\ c) & F = \sum_{\zeta \in \mu_n} \frac{\zeta^k (X - \omega)}{(X - \zeta)^2}, \text{ avec } \omega = e^{i\pi/n} \text{ et } k \in \mathbb{N}. \end{aligned}$$

Exercice 8 : Simplifier les sommes de fractions rationnelles suivantes :

$$\begin{aligned} a) & \sum_{k=0}^n \frac{1}{(X+k)(X+k+1)}; & b) & \sum_{k=0}^n \frac{1}{(X+k)(X+k+1)(X+k+2)}; \\ c) & \sum_{k=0}^n \frac{2^k}{(X+2^k)(X+2^{k+1})}. \end{aligned}$$

Exercice 9 : Montrer que si P est irréductible et A non multiple de P , et si l'on écrit $\frac{A}{P}$ comme somme de fractions rationnelles $F_1 + F_2 + \dots + F_n$, l'une au moins des F_i doit avoir un dénominateur multiple de P . Peut-on affirmer la même chose pour $\frac{A}{P^m}$?

Exercice 10 : Trouver un procédé permettant de représenter un nombre rationnel comme « somme » de fractions du type $\frac{a}{p^n}$ (p premier, $n \in \mathbb{N}^*$, $0 \leq a < p$). Par exemple $\frac{1}{6} = \frac{1}{2} - \frac{1}{3}$. Effectuer cette décomposition pour $\frac{1887}{5400}$ et pour $\frac{122}{1323}$.

Exercice 11 : Le corps de base K est de caractéristique nulle. On considère $F \in K(X)^*$, mise sous forme irréductible $F = \frac{P}{Q}$.

a) Soit a un pôle double de F . Montrer que la partie polaire de F en a est

$$\frac{A}{(X-a)^2} + \frac{B}{X-a} \quad \text{avec} \quad A = \frac{2P(a)}{Q''(a)}, \quad B = \frac{2}{3} \times \frac{3P'(a)Q''(a) - P(a)Q'''(a)}{[Q''(a)]^2}.$$

b) Application : Décomposer $\frac{1}{[H(X)]^2}$, où $H(X) = \prod_{i=1}^n (X - x_i)$, les $x_i \in K$ étant tous distincts.

c) Calculer de même la partie polaire d'une fraction F en un pôle triple a .

Exercice 12 : Si $\theta \in \mathbb{R}$ calculer les dérivées successives de

$$F = \frac{1}{1 - 2X \cos \theta + X^2} \quad \text{et} \quad G = \frac{1 - X \cos \theta}{1 - 2X \cos \theta + X^2}.$$

Exercice 13 : Soit $F = \frac{1}{(1-X)(1-X^2)(1-X^5)} \in \mathbb{C}(X)$.

a) Calculer la décomposition en éléments simples de F .

b) Décomposer F en éléments simples dans $\mathbb{R}(X)$ en explicitant avec des radicaux la décomposition de $1 - X^5$ dans $\mathbb{R}[X]$.

c) Même problème avec $G = \frac{1}{(1-X^2)(1-X^3)(1-X^5)}$.

Exercice 14 : Dédurre le théorème VIII.2.2 du théorème d'existence et unicité de la division euclidienne, en utilisant la nouvelle indéterminée $T = \frac{1}{X}$, et expliquer ainsi le parallélisme parfait des deux algorithmes de division.

Exercice 15 : Soit a_1, a_2, \dots, a_n des éléments distincts dans un corps commutatif K . On pose $P = (X - a_1)(X - a_2) \dots (X - a_n)$. Pour b_1, b_2, \dots, b_n dans K , étudier la fraction $F = \sum_{i=1}^n \frac{b_i}{P'(a_i)(X - a_i)}$ et en déduire une nouvelle démonstration du théorème d'interpolation de Lagrange.

Exercice 16 : Déterminer les constantes p et q pour que les résidus de la fraction rationnelle $F = \frac{X^2 + pX + q}{(X^2 - 1)^2}$ aux pôles 1 et -1 soient nuls.

Même question avec $G = \frac{1}{X^2(X-1)^2(X^2 + pX + q)}$ et les résidus nuls en 0 et 1.

§ VIII.3 FONCTIONS RATIONNELLES. DÉRIVATION

Considérons une fraction rationnelle $F \in K(X)$ et une extension L de K .

Si $F = 0$, on convient que F définit la fonction nulle de L dans L .

Si $F \neq 0$, écrivons F sous sa forme irréductible $F = \frac{N}{D}$; soit $x \in L$: nous

dirons que F est **définie en x** ssi $D(x) \neq 0$, et si c'est le cas, l'élément $\frac{N(x)}{D(x)} \in L$ sera appelé **la valeur de F en x** et noté $\tilde{F}_L(x)$, ou $F_L(x)$, ou même tout simplement $F(x)$ si aucune confusion ne peut en résulter

$\mathcal{D}_L(F) = \{x \in L \mid D(x) \neq 0\}$ sera appelé **domaine de définition de F dans L** , et la fonction $\mathcal{D}_L(F) \rightarrow L, x \mapsto F(x)$ sera appelée la **fonction rationnelle définie sur L par F** . (Si $F = 0$, on a $\mathcal{D}_L(F) = L$.)

Soit $F \in K(X)$ et $G \in K(X)$, alors $\mathcal{D}_L(F) \cap \mathcal{D}_L(G) \subset \mathcal{D}_L(F + G)$, $\mathcal{D}_L(F) \cap \mathcal{D}_L(G) \subset \mathcal{D}_L(FG)$, et la définition même de l'addition et de la multiplication dans $K(X)$ montre :

$$(1) \quad \forall x \in \mathcal{D}_L(F) \cap \mathcal{D}_L(G) \quad \begin{cases} (F + G)(x) = F(x) + G(x) \\ (FG)(x) = F(x) G(x) . \end{cases}$$

Un cas particulier essentiel est celui où $L = K$. Dans ce cas $\mathcal{D}_K(F)$ est tout simplement le **complémentaire dans K de l'ensemble des pôles de F** .

Il est parfois commode de rajouter à K un élément noté ∞ ($\infty \notin K$).

On note $\tilde{K} = K \cup \{\infty\}$ ⁽¹⁾, et on convient que $F \in K(X)$ prend la valeur ∞ en chacun de ses pôles, que $F(\infty) = 0$ si $\deg(F) \leq -1$, que $F(\infty) = \infty$ si $\deg(F) \geq 1$ et $F(\infty) = \frac{N}{D}$ quotient des coefficients dominants de N et de D si $\deg(F) = 0$ et $F = \frac{N}{D}$.

Si $F \in K(X)$ est *constante*, égale à $\lambda \in K$, elle est partout définie (même sur \tilde{K}), et prend partout la valeur λ .

Exemple 1 : Supposons $x \in L$ transcendant sur K . Puisque $P(x) \neq 0$ pour tout $P \in K[X]^*$, il s'ensuit que toute fraction $F \in K(X)$ est définie en x .

L'application $F \mapsto F(x)$ laisse fixes les éléments de K , donc en vertu des relations (1), cette application est un *isomorphisme du corps $K(X)$ dans le corps L* (cf. théorème II.6.3), et cet isomorphisme est un *homomorphisme de K -algèbres*. Le corps image de cet isomorphisme se note naturellement $K(x)$. Par exemple $\mathbb{Q}(\pi)$ désigne l'ensemble des nombres réels qui peuvent se mettre sous la forme $\frac{a_0 + a_1 \pi + \dots + a_n \pi^n}{b_0 + b_1 \pi + \dots + b_m \pi^m}$, avec des a_i et des b_j rationnels. $\mathbb{Q}(\pi)$ est isomorphe à $\mathbb{Q}(X)$.

Exemple 2 : Prenons une indéterminée Y sur le corps L . Alors Y est transcendant sur K . D'après l'exemple 1, l'application $F \mapsto F(Y)$ est un isomorphisme de $K(X)$ dans $L(Y)$. Il envoie X sur Y et est K -linéaire. A l'aide de cet isomorphisme, $K(X)$ s'identifie à un sous-corps de $L(Y)$: cette identification consiste à « considérer les fractions rationnelles à coefficients dans K comme des fractions à coefficients dans L ». Quand cela n'entraîne pas de confusion, on procède à cette identification sans même prendre la peine de changer la lettre X . Mais s'il advenait que L soit non seulement une extension de K mais aussi une extension de $K(X)$, cela amènerait des contresens et doit donc être évité.

⁽¹⁾ Note : Bien entendu, cet élément ∞ n'a absolument rien à voir avec les éléments $+\infty$ et $-\infty$ qu'on a rajoutés à \mathbb{Z} pour définir la valuation et le degré d'ur

Exemple 3 : Prenons $L = K(X)$. Alors toute $G \in K(X)$ non constante est un élément *transcendant* sur K . En effet, soit $\frac{P}{Q}$ la forme irréductible de G et soit $D \in K[X]$, D non constant, de degré d ($d \geq 1$). Posons $D = c_0 X^d + c_1 X^{d-1} + \dots + c_d$ ($c_i \in K$, $c_0 \in K^*$). On a :

$$D \left(\frac{P}{Q} \right) = \frac{1}{Q^d} (c_0 P^d + c_1 P^{d-1} Q + \dots + c_d Q^d).$$

Si $D \left(\frac{P}{Q} \right) = 0$, cela entraîne que Q divise P^d , d'où, puisqu'on a supposé P et Q premiers entre eux, Q est une constante non nulle. Des raisons de degré montrent ensuite que P est constant, mais alors $G = \frac{P}{Q}$ serait constant, contrairement à l'hypothèse. Donc $D(G) \neq 0$. D'après l'exemple 1, cela entraîne que toute fraction $F \in K(X)$ est définie en G . Cet exemple débouche sur une importante définition :

DÉFINITION VIII.3.1

Soit $G \in K(X)$ non constante, et $F \in K(X)$. On appelle **composée de F par G** , et on note $F \circ G$, ou $F(G)$, la valeur de F en G .

Lorsque $G \in K(X)$ est fixée (non constante), d'après l'exemple 1, l'application $F \mapsto F \circ G$ est un isomorphisme du corps $K(X)$ dans lui-même, qui est de plus K -linéaire. Si on prend $G = X$, on a $F \circ G = F(X) = F$ et l'isomorphisme se réduit à l'identité de $K(X)$, ce qui légitime l'écriture $F(X)$ pour F . Bien entendu, si $F \in K[X]$ et $G \in K[X]$ (G non constant), on retrouve la notion de composée de deux polynômes vue au § VII.4.

DÉFINITION VIII.3.2

Soit \mathcal{D} une partie non vide du corps K . Une fonction $f : \mathcal{D} \rightarrow K$ est dite **rationnelle** ssi il existe $F \in K(X)$ telle que F soit partout définie sur \mathcal{D} et que $f(x) = F(x)$ pour tout $x \in \mathcal{D}$.

Une conséquence immédiate de (1) est que si \mathcal{D} est fixée, les fonctions rationnelles définies sur \mathcal{D} forment une sous- K -algèbre de la K -algèbre $\mathcal{F}(\mathcal{D}, K)$ des fonctions de \mathcal{D} dans K .

La notation $F \circ G$ n'est pas choisie au hasard : en effet, si la fraction G n'est pas constante et si l'image de G est entièrement contenue dans le domaine de définition de F , alors à la fraction $F \circ G$ est associée une fonction rationnelle qui est justement la composée des fonctions rationnelles respectivement associées à F et à G .

Pour que $f : \mathcal{D} \rightarrow K$ soit rationnelle, il faut et il suffit qu'il existe $P \in K[X]$, $Q \in K[X]$, avec Q sans zéro sur \mathcal{D} , tels que ($\forall x \in \mathcal{D}$) $f(x) = \frac{P(x)}{Q(x)}$.

C'est lorsque \mathcal{D} est *infini* que la notion de fonction rationnelle prend tout son sens :

THÉORÈME VIII.3.1

Soit \mathcal{D} une partie **infinie** du corps K , et soit $\Phi_{\mathcal{D}}$ l'ensemble des $F \in K(X)$ **partout définies sur \mathcal{D}** . Alors $\Phi_{\mathcal{D}}$ est une sous- K -algèbre de $K(X)$, et l'application qui associe, à toute $F \in \Phi_{\mathcal{D}}$, la fonction rationnelle $\mathcal{D} \rightarrow K$, $x \mapsto F(x)$, est **injective**, et définit un **isomorphisme de K -algèbres** entre $\Phi_{\mathcal{D}}$ et la K -algèbre des fonctions rationnelles sur \mathcal{D} .

Démonstration :

La seule chose nouvelle à prouver est l'injectivité annoncée, tout le reste découlant de (1) et de l'étude précédente. Supposons donc que $F \in \Phi_{\mathcal{D}}$ définisse sur \mathcal{D} la fonction rationnelle nulle ; écrivons $F = \frac{P}{Q}$ avec P et Q dans $K[X]$, et Q sans zéro sur \mathcal{D} . Par hypothèse, si $x \in \mathcal{D}$, on a donc $Q(x) \neq 0$ et $F(x) = \frac{P(x)}{Q(x)} = 0$, d'où $P(x) = 0$. Puisque \mathcal{D} est infini, il s'ensuit que $P = 0$, d'où $F = 0$. ■

Exemple 4 : On prend $K = \mathbb{C}$. Soit $F \in \mathbb{C}(X) \setminus \mathbb{C}[X]$, mise sous forme irréductible $F = \frac{N}{D}$; elle définit une fonction rationnelle $f : \mathcal{D} \rightarrow \mathbb{C}$, où \mathcal{D} est le complémentaire de l'ensemble des pôles de F . L'image de f est soit \mathbb{C} , soit \mathbb{C} privé d'un point, comme le montre l'étude de l'équation $\{N(z) - \lambda D(z) = 0, D(z) \neq 0\}$ pour $\lambda \in \mathbb{C}$.

Soit a un pôle de F , de multiplicité α . Étudions f au voisinage de a , puisqu'elle est définie pour $|z - a| < r$, $z \neq a$, dès que $r > 0$ est choisi \leq (distance de a aux autres pôles s'il y en a), r quelconque si a est le seul pôle. Pour cela on décompose F sur \mathbb{C} et on isole la partie polaire $\mathcal{F}_a(X)$ de F relative au pôle a : $F = G + \mathcal{F}_a(X)$, d'où $f = g + \varphi_a$, g et φ_a étant les fonctions correspondantes à G et \mathcal{F}_a . Or, g admet une limite finie quand $z \rightarrow a$, $z \neq a$. Et $\varphi_a(z) = \frac{c_\alpha}{(z-a)^\alpha} + \dots + \frac{c_1}{z-a}$ avec $c_\alpha \neq 0$ (théorème VIII.2.4). Écrivant :

$$\varphi_a(z) = \frac{c_\alpha}{(z-a)^\alpha} \left[1 + \frac{c_{\alpha-1}}{c_\alpha} (z-a) + \dots + \frac{c_1}{c_\alpha} (z-a)^{\alpha-1} \right],$$

on voit que

$$\varphi_a(z) \underset{z \rightarrow a}{\sim} \frac{c_\alpha}{(z-a)^\alpha},$$

d'où $|f(z)| \rightarrow +\infty$ quand $z \rightarrow a$, $z \neq a$. Pour des précisions, voir exercice 4.

Dérivation

Soit $F \in K(X)$ et deux de ses représentants : $F = \frac{A_1}{B_1} = \frac{A_2}{B_2}$. Alors

$$(2) \quad \frac{A'_1 B_1 - A_1 B'_1}{B_1^2} = \frac{A'_2 B_2 - A_2 B'_2}{B_2^2},$$

ce qui se montre facilement en dérivant l'égalité $A_1 B_2 = A_2 B_1$, ce qui donne $A'_1 B_2 - A_2 B'_1 = A'_2 B_1 - A_1 B'_2$, puis en multipliant les deux membres par $B_1 B_2$, compte tenu de $A_1 B_2 = A_2 B_1$, et enfin en divisant par $B_1^2 B_2^2$.

DÉFINITION VIII.3.3

Si $F \in K(X)$, on appelle **K-dérivée** de F , ou **dérivée**, et on note F' la fraction de $K(X)$ telle que, pour tout représentant $\frac{A}{B}$ de F , on ait

$$F' = \frac{A' B - A B'}{B^2}.$$

Les propriétés les plus importantes de la dérivée sont résumées dans le :

THÉOREME VIII.3.2

- (I) La dérivation des fractions rationnelles **prolonge** la dérivation des polynômes, et est K -linéaire.
 (II) Pour F et G dans $K(X)$, $(FG)' = F' G + F G'$.
 (III) Si $F \in K(X)^*$, $\left(\frac{1}{F}\right)' = -\frac{F'}{F^2}$.
 (IV) Si G est non constante dans $F(X)$, alors pour toute $F \in K(X)$
- $$(F \circ G)' = (F' \circ G) \times G'.$$

Démonstration (abrégée) :

L'assertion (I), facile, sera laissée au lecteur.

(III) se déduit de (II) en posant $G = \frac{1}{F}$. (II) se démontre en choisissant des représentants $\frac{A}{B}$ de F et $\frac{C}{D}$ de G , et en se ramenant à des calculs sur des polynômes. Il reste l'assertion (IV). Fixons G . Compte tenu de la K -linéarité de la dérivation et de la possibilité de décomposer F en une somme d'éléments simples, il suffit de prouver (IV) avec $F = X^n$ ($n \in \mathbb{N}^*$) et de prouver que, si elle est vraie avec $F \in K(X)^*$, elle reste vraie avec $\frac{1}{F}$.

Si $F = X^n$ ($n \in \mathbb{N}^*$), $F \circ G = G^n$, $F' = nX^{n-1}$ et $(F \circ G)' = nG^{n-1} G'$ par récurrence immédiate à partir de (II). L'assertion (IV) est bien prouvée dans ce cas.

Si (IV) est vraie avec $F \neq 0$, on a : $\left(\frac{1}{F}\right) \circ G = \frac{1}{F \circ G}$, $\left(\frac{1}{F}\right)' = -\frac{F'}{F^2}$, puis d'après (III)

$$\left(\frac{1}{F \circ G}\right)' = -\frac{(F \circ G)'}{(F \circ G)^2} = \frac{-(F' \circ G) \times G'}{(F \circ G)^2}.$$

L'assertion (IV) est bien prouvée là aussi. ■

Bien sûr on a intérêt, pour dériver $F \in K(X)^*$, à choisir un représentant irréductible $F = \frac{N}{D}$, d'où $F' = \frac{N'D - ND'}{D^2}$. On voit que si $F' \neq 0$, la forme irréductible $F' = \frac{P}{Q}$ de F' est telle que Q divise D^2 . Donc on ne fait pas apparaître par dérivation de nouveaux facteurs irréductibles au dénominateur. En particulier, *l'ensemble des pôles de F' est contenu dans l'ensemble des pôles de F* . De plus :

THÉORÈME VIII.3.3

|| Supposons K de caractéristique nulle. Si a est un pôle d'ordre $\alpha \geq 1$ de $F \in K(X)^*$, alors a est un pôle **d'ordre $\alpha + 1$** de F' .

Démonstration :

Décomposons F en éléments simples et dérivons : on voit que la partie entière de F' est la dérivée de celle de F . Pour tout $P \in \mathcal{J}$, la partie P -fractionnaire de F' est la dérivée de celle de F . Or la partie $(X - a)$ -fractionnaire de F s'écrit

$$\mathcal{F}_{X-a}(F) = \frac{c_0}{(X-a)^\alpha} + \frac{c_1}{(X-a)^{\alpha-1}} + \cdots + \frac{c_{\alpha-1}}{X-a} \quad \text{avec } c_0 \in K^*,$$

d'où

$$\mathcal{F}_{X-a}(F') = \frac{-\alpha c_0}{(X-a)^{\alpha+1}} + \cdots - \frac{c_{\alpha-1}}{(X-a)^2}.$$

Puisque la caractéristique est nulle, et puisque $c_0 \neq 0$, on a $-\alpha c_0 \neq 0$, et cela établit bien, d'après les remarques qui suivent le théorème VIII.2.1, que a est pôle d'ordre $\alpha + 1$ de F' . ■

Remarque 1 : D'après cette démonstration on voit que F' n'a pas de pôle simple et que **tous les résidus de F' sont nuls**. Ces résultats restent valables, même si K est de caractéristique p (mais certains pôles risquent de disparaître par dérivation lorsque la caractéristique est non 1

Etudions maintenant l'application K -linéaire $K(X) \rightarrow K(X)$, $F \mapsto F'$, dans l'hypothèse où **la caractéristique de K est nulle**. Elle envoie les constantes sur 0. Réciproquement, soit $F \in K(X)^*$ telle que $F' = 0$. Peut-on affirmer que F est constante ? (Nous avons déjà vu que cette affirmation serait fausse si K est de caractéristique p (exemple : X^p a une dérivée nulle dans $\mathbb{Z}/p\mathbb{Z}$)). Ecrivons F sous forme irréductible : $F = \frac{P}{Q}$, d'où $F' = \frac{P'Q - PQ'}{Q^2}$. Par hypothèse $P'Q - PQ' = 0$. Mais nous avons pris P et Q premiers entre eux. Le théorème de Gauss entraîne que P divise P' , ce qui oblige P à être constant, car la caractéristique est nulle. De même Q divise Q' , donc Q est constant et finalement F est constante. Ainsi, **lorsque la caractéristique de K est nulle, $F' = 0$ ssi F est constante** ($F \in K(X)$). De plus :

THÉORÈME VIII.3.4

|| Supposons K **algébriquement clos** et de **caractéristique nulle**.
 || L'image de l'application K -linéaire $K(X) \rightarrow K(X)$, $F \mapsto F'$ est
 || l'ensemble des fractions $G \in K(X)$ dont tous les résidus sont nuls.

· *Démonstration :*

On a déjà vu dans la remarque 1 que toute dérivée a tous ses résidus nuls. Il s'agit d'examiner la réciproque compte tenu des hypothèses sur K . Soit donc $G \in K(X)$ ayant tous ses résidus nuls. Ecrivons la décomposition en éléments simples de G : $G = E + \sum_{a \in K} \mathcal{F}_a$, où

$E \in K[X]$ et $\mathcal{F}_a = \sum_{n \geq 1} \frac{\lambda_{n,a}}{(X-a)^n}$ pour $a \in K$, la famille de scalaires

$(\lambda_{n,a})$ étant à support fini et telle que par l'hypothèse $\lambda_{1,a} = 0$ pour tout a .

Il est alors facile de vérifier que $\mathcal{F}_a = \mathcal{G}'_a$ avec $\mathcal{G}_a = \sum_{n \geq 2} \frac{1}{1-n} \frac{\lambda_{n,a}}{(X-a)^{n-1}}$.

De même si $E = \sum_{k \geq 0} c_k X^k$, il est évident que $E = P'$ avec

$$P = \sum_{k \geq 0} \frac{c_k}{k+1} X^{k+1}, \text{ d'où } G = F' \text{ avec } F = P + \sum_{a \in K} \mathcal{G}_a. \quad \blacksquare$$

De manière générale, les fractions $F \in K(X)$ ayant une dérivée donnée $G \in K(X)$ s'appellent les **primitives** de G dans $K(X)$ (ou : **primitives rationnelles** de G). L'étude précédente montre que si K est de caractéristique nulle, lorsqu'il existe une primitive rationnelle à une $G \in K(X)$, cette primitive est unique à l'addition près d'une constante. Si de plus K est algébriquement clos, le théorème VIII.3.4 permet de reconnaître quelles sont les $G \in K(X)$ qui ont des primitives rationnelles.

Dérivation pour $K = \mathbb{R}$ ou $K = \mathbb{C}$

Soit \mathcal{D} un intervalle non vide et non réduit à un point d.

$f: \mathcal{D} \rightarrow \mathbb{C}$ une fonction rationnelle, définie par $F \in \mathbb{C}(X)$ (cela suppose expressément que F n'admet aucun pôle sur \mathcal{D}). Alors F' n'a aucun pôle sur \mathcal{D} . Nous verrons en Analyse (tome 2) que f est dérivable sur \mathcal{D} et que $f'(x) = F'(x)$ pour tout $x \in \mathcal{D}$.

En particulier une fonction rationnelle f définie sur \mathcal{D} est continue sur \mathcal{D} , donc admet des primitives au sens de l'Analyse. Si f est définie à partir de la fraction $F \in \mathbb{C}(X)$, ce qui précède prouve que **ces primitives seront rationnelles ssi F admet dans $\mathbb{C}(X)$ des primitives rationnelles**, c'est-à-dire ssi F a tous ses résidus nuls ; et si c'est le cas, une primitive de f au sens de l'Analyse est fournie par la fonction $\mathcal{D} \rightarrow \mathbb{C}$, $x \mapsto H(x)$ où H est une primitive rationnelle de F .

De même, soit \mathcal{U} un ouvert non vide de \mathbb{C} et $f: \mathcal{U} \rightarrow \mathbb{C}$ une fonction rationnelle, définie par $F \in \mathbb{C}(X)$ n'ayant aucun pôle dans \mathcal{U} . Alors f est \mathbb{C} -dérivable sur \mathcal{U} , et sa \mathbb{C} -dérivée est la fonction $\mathcal{U} \rightarrow \mathbb{C}$, $z \mapsto F'(z)$ (voir le tome 3 d'Analyse).

Exercice 1 : Soit $F \in \mathbb{C}(X) \setminus \mathbb{C}[X]$ et \mathcal{P} l'ensemble des pôles de F . Étudier avec soin l'image de la fonction rationnelle $\mathbb{C} \setminus \mathcal{P} \rightarrow \mathbb{C}$, $z \mapsto F(z)$.

Exercice 2 : On se place dans le plan d'Argand-Cauchy. Soit $F \in \mathbb{C}[X]$ un polynôme non constant.

- Si toutes les racines de F sont dans un même demi-plan ouvert \mathcal{H} de \mathbb{C} , les racines de F' appartiennent à \mathcal{H} (indication : utiliser la fraction $\frac{F'}{F}$).
- Si toutes les racines de F sont dans un même demi-plan fermé de frontière Δ , et si F' a des racines sur Δ , alors F a des racines sur Δ .
- Si les racines de F sont dans un polygone convexe fermé de \mathbb{C} , celles de F' y sont aussi.
- Si les racines de F sont simples et ont pour images les sommets d'un polygone convexe, les racines de F' sont à l'intérieur de ce polygone.

Exercice 3 : Trouver tous les couples (F, G) d'éléments non constants de $\mathbb{C}(X)$ tels que $F \circ G \in \mathbb{C}[X]$. Indication : utiliser l'exercice 1 pour borner le nombre de pôles de F .

Exercice 4 : Soit a un pôle d'ordre $\alpha \geq 1$ d'une fraction rationnelle $F \in \mathbb{C}(X)$. Montrer qu'il existe $\eta > 0$ et $M > 0$ tels que, pour tout $t \in \mathbb{C}$ vérifiant $|t| \geq M$, l'équation $\{z \neq 0, |z - a| < \eta, F(z) = t\}$ possède exactement α racines.

Exercice 5 : $F \in K(X)$ est dite *homographique* ssi elle est de la forme $F = \frac{aX + b}{cX + d}$ avec $ad - bc \neq 0$ (ce qui signifie qu'elle est non constante). On note $\text{PGL}(K)$ l'ensemble de ces fractions.

- Montrer que la loi $(F, G) \mapsto F \circ G$ définit sur $\text{PGL}(K)$ une structure de groupe.
- A chaque matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}(2, K)$ on associe la fraction homographique $F_M = \frac{aX + b}{cX + d}$. Montrer que l'application $\text{GL}(2, K) \rightarrow \text{PGL}(K)$, $M \mapsto F_M$ est un homomorphisme de groupes, surjectif. Quel est son noyau ? Cas particulier où K est algébriquement clos.
- Vérifier qu'on définit une action à gauche du groupe $\text{PGL}(K)$ sur l'ensemble $\tilde{K} = K \cup \{\infty\}$ en associant à $F = \frac{aX + b}{cX + d}$ la bijection $h_F: \tilde{K} \rightarrow \tilde{K}$, telle que : si $c = 0$, $\infty \mapsto \infty$, $z \mapsto \frac{az + b}{d}$ si $z \in K$ et si $c \neq 0$, $\infty \mapsto \frac{a}{c}$, $\frac{-d}{c} \mapsto \infty$ et $z \mapsto \frac{az + b}{cz + d}$ si $z \in K$.

En d'autres termes, h_F est la *fonction rationnelle* définie par F , et étendue à \tilde{K} . Prouver que cette action à gauche est fidèle. (Cette action est dite naturelle et permet d'identifier $\text{PGL}(K)$ à un groupe de permutations de \tilde{K}).

d) Prouver que, quels que soient z_1, z_2, z_3 , éléments distincts dans \tilde{K} , quels que soient t_1, t_2, t_3 , éléments distincts de \tilde{K} , il existe une et une seule $h \in \text{PGL}(K)$ telle que $h(z_i) = t_i$ pour $i \in \{1, 2, 3\}$.

Exercice 6 : On prend $K = \mathbb{C}$.

a) Soit $G \in \mathbb{C}(X)$ une fraction non constante : $G = \frac{P}{Q}$. Montrer qu'il existe $\nu \in \mathbb{N}^*$ et une partie finie \mathcal{S}_G de \mathbb{C} telle que, pour tout $t \in \mathbb{C} \setminus \mathcal{S}_G$, l'ensemble $\{z \in \mathbb{C} \mid G \text{ est définie en } z \text{ et } G(z) = t\}$ ait exactement ν éléments, chaque racine de $P - tQ$ étant *simple*. L'entier ν sera appelé *indice* de G . Prouver que si $t \in \mathcal{S}_G$, l'ensemble $\{z \in \mathbb{C} \mid G \text{ est définie en } z \text{ et } G(z) = t\}$ a au plus ν éléments.

b) Montrer que l'indice ν de $G \in \mathbb{C}(X)$ non constante vaut 1 ssi G est *homographique* (cf. exercice 5).

c) Soit $G \in \mathbb{C}(X)$ non constante ; pour que l'isomorphisme $F \mapsto F(G)$ de $\mathbb{C}(X)$ dans $\mathbb{C}(X)$ soit un automorphisme, il faut et il suffit que G soit homographique.

d) Montrer que les automorphismes de la \mathbb{C} -algèbre $\mathbb{C}(X)$ sont exclusivement ceux trouvés en c).

Exercice 7 : Soit $A, B, y_1, y_2, \dots, y_N, c_1, c_2, \dots, c_N$ des réels, les c_k non nuls, les y_k distincts ($N \in \mathbb{N}^*$). On considère la fraction $F \in \mathbb{R}(X)$ donnée par :

$$F = AX + B - \sum_{k=1}^N \frac{c_k}{X - y_k}.$$

Montrer que, pour que F possède $N + 1$ zéros réels simples et *entrelacés* avec les pôles y_k , il faut et il suffit que A, B et les c_k soient tous de même signe.

Exercice 8 : Soit F, G, H dans $K(X)$ avec G et H non constantes. Démontrer en détail que $F \circ (G \circ H) = (F \circ G) \circ H$.

Exercice 9 : Le corps de base est \mathbb{C} . On donne $F \in \mathbb{C}(X)$ non constante. On suppose trouvé $k \in \mathbb{N}^*$ tel que $F(\zeta X) = F(X)$ pour tout $\zeta \in \mu_k$. Prouver qu'il existe $G \in \mathbb{C}(X)$ telle que $F(X) = G(X^k)$.

Exercice 10 : Le corps de base est \mathbb{C} . Trouver tous les réels λ tels que la fonction $x \mapsto \text{tg}(\lambda \text{Arctg } x)$ (qui est définie au voisinage de 0 dans \mathbb{R}) soit rationnelle.

Exercice 11 : Soit $m \in \mathbb{N}^*$.

a) Montrer qu'il existe $F \in \mathbb{C}(X)$ unique telle que $\cotg(2m+1)x = F(\text{tg } x)$ pour tout $x \in \mathbb{R}$ tel que les deux membres soient définis.

b) En calculant, pour $x \in \mathbb{R}$, le produit $P_m = \prod_{k=0}^{2m} \cotg\left(x + \frac{k\pi}{2m+1}\right)$, déterminer les pôles de F , puis décomposer F en éléments simples.

c) En déduire : $(\forall x \in \mathbb{R} \setminus \pi\mathbb{Z})$

$$\cotg x = \frac{(-1)^m}{(2m+1) \text{tg} \frac{x}{2m+1}} + \sum_{k=1}^{2m} \frac{(-1)^m (2m+1) \text{tg} \frac{x}{2m+1}}{\cos^2 \frac{k\pi}{2m+1} (2m+1)^2 \text{tg}^2 \frac{x}{2m+1} - (2m+1)^2 \sin^2 \frac{k\pi}{2m+1}}.$$

d) En faisant tendre m vers $+\infty$, en déduire

$$(\forall x \in \mathbb{R} \setminus \pi\mathbb{Z}) \quad \cotg x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2x}{x^2 - k^2 \pi^2} \quad (\text{poser } m = 2p)$$

Exercice 12 : Le corps de base est \mathbb{C} . Soit Γ l'ensemble des six homographies suivantes :

$$e = X, \quad f_1 = \frac{1}{X}, \quad f_2 = 1 - X, \quad f_3 = \frac{1}{1 - X}, \quad f_4 = 1 - \frac{1}{X}, \quad f_5 = \frac{X}{X - 1}.$$

a) Montrer que Γ est un sous-groupe du groupe $\text{PGL}(\mathbb{C})$ défini dans l'exercice 5, et que ce groupe est isomorphe à \mathfrak{S}_3 .

b) On identifie Γ à un groupe de bijections de $\tilde{\mathbb{C}}$ dans $\tilde{\mathbb{C}}$ (cf. exercice 5). Trouver les groupes d'isotropie des divers $z \in \tilde{\mathbb{C}}$. Étudier la figure formée par les points fixes des f_k ($1 \leq k \leq 5$).

c) Montrer que toutes les Γ -orbites de $\tilde{\mathbb{C}}$ ont 6 éléments, sauf un nombre fini d'entre elles que l'on précisera.

d) Soit $\Phi = X^2 - X + 1$, $\Psi = (X + 1)(2X - 1)(X - 2)$, et $\Lambda = \frac{\Phi^3}{\Psi^2}$. Démontrer que

$\Lambda \circ F = \Lambda$ pour toute $F \in \Gamma$.

e) Soit $\Theta \in \mathbb{C}(X)$ telle que $\Theta \circ F = \Theta$ pour toute $F \in \Gamma$. Montrer qu'il existe $G \in \mathbb{C}(X)$ telle que $\Theta = G \circ \Lambda$. *Indication :* Montrer qu'il existe une Γ -orbite ω à 6 éléments sur laquelle Θ et Λ sont définies. Pour une telle orbite, établir que pour $\lambda \in \mathbb{C}$ et $\mu \in \mathbb{C}$ convenables, $\Theta - \mu$ et $\Lambda - \lambda$ s'annulent sur ω , et étudier les numérateurs de ces fractions.

Exercice 13 : Le corps de base est \mathbb{C} . Montrer que les trois homographies $f_1 = iX$, $f_2 = \frac{1}{X}$, $f_3 = \frac{X+1}{X-1}$ engendrent dans $\text{PGL}(\mathbb{C})$ un groupe de cardinal 24 et étudier la structure de ce groupe.

Exercice 14 : a) Trouver les $(\alpha, \beta) \in \mathbb{R}^2$ tels que la fonction $t \mapsto (1-t)^\alpha (1+t)^\beta$ soit rationnelle.

b) Trouver les $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ tels que la fonction

$$t \mapsto (1-t)^\alpha (1+t)^\beta (1+t^2)^\gamma$$

soit rationnelle.

Exercice 15 : n est un entier ≥ 2 et ζ_k désigne le nombre complexe $e^{\frac{2ik\pi}{n}}$.

a) Montrer que l'ensemble des n fractions rationnelles $\frac{X + \zeta_1}{X - \zeta_1}, \frac{X + \zeta_2}{X - \zeta_2}, \dots, \frac{X + \zeta_n}{X - \zeta_n}$ est égal à l'ensemble $\left\{ \frac{\zeta_p X + \zeta_1}{\zeta_p X - \zeta_1}, \dots, \frac{\zeta_p X + \zeta_n}{\zeta_p X - \zeta_n} \right\}$, où p est un entier fixe ($1 \leq p \leq n$).

b) On pose $F = \sum_{k=1}^n \frac{X + \zeta_k}{X - \zeta_k}$. Montrer que $F(z) = F(\zeta_p z)$ pour chaque p ($\forall z \neq \zeta_k$).

c) On pose $P = (X^n - 1)F$. Montrer que $P = aX^n + b$ (a et b à calculer).

Exercice 16 : Le corps de base est \mathbb{C} .

a) Montrer que le stabilisateur de ∞ dans $\text{PGL}(\mathbb{C})$ est le groupe des *similitudes planes directes* $(aX + b)_{a \in \mathbb{C}^*}$. En donner les sous-groupes *finis*.

b) Montrer qu'un sous-groupe fini Γ de $\text{PGL}(\mathbb{C})$ a forcément un point fixe.

c) En déduire tous les sous-groupes *finis* de $\text{PGL}(\mathbb{C})$ (ils sont tous cycliques).

§ VIII.4 NOTIONS SUR LES SÉRIES FORMELLES A UNE INDÉTERMINÉE

On désigne par K un corps commutatif.

Considérons le K -ev $\mathcal{F}(\mathbb{N}, K) = K^{\mathbb{N}}$ des suites, inde:

d'éléments de K . Ce K -ev peut être muni d'une *multiplication* de la manière suivante : si $S = (a_p)_{p \in \mathbb{N}}$ et $T = (b_q)_{q \in \mathbb{N}}$ sont deux éléments de $\mathcal{F}(\mathbb{N}, K)$, on pose :

$$(1) \quad ST = (c_n)_{n \in \mathbb{N}} \quad \text{où, pour tout } n \in \mathbb{N}, \quad \boxed{c_n = \sum_{p+q=n} a_p b_q},$$

ce qui a bien un sens car la somme définissant c_n est finie.

Des vérifications immédiates montrent que cette multiplication munit le K -ev $\mathcal{F}(\mathbb{N}, K)$ d'une **structure de K -algèbre commutative**, pour laquelle l'élément unité est la suite $(\delta_{0k})_{k \in \mathbb{N}} = \delta_0$ ($\delta_{00} = 1_K$, $\delta_{0k} = 0$ si $k \geq 1$). L'élément neutre de l'addition sera noté 0 . On a : $\delta_0 \neq 0$.

Conformément aux conventions générales nous identifierons K au sous-anneau $K\delta_0$ de cette K -algèbre, et l'élément unité sera donc noté 1 au lieu de δ_0 . Les séries éléments de K seront dites **constantes**.

DÉFINITION VIII.4.1

La K -algèbre définie par la multiplication (1) sur le K -ev $\mathcal{F}(\mathbb{N}, K)$ s'appelle **K -algèbre des séries formelles à une indéterminée sur K** . Si $S = (a_n)_{n \in \mathbb{N}}$ est une série formelle, a_n s'appelle le **coefficient de rang n de S** .

On constate que les suites à support fini dans $\mathcal{F}(\mathbb{N}, K)$ forment avec la règle de multiplication (1) une sous- K -algèbre dans laquelle on reconnaît la K -algèbre des polynômes à une indéterminée sur K . Notons X l'indéterminée canonique de cette algèbre de polynômes : alors on convient de noter $K[[X]]$ la K -algèbre $\mathcal{F}(\mathbb{N}, K)$.

DÉFINITION VIII.4.2

Soit $S = (a_p)_{p \in \mathbb{N}} \in K[[X]]$; si $S \neq 0$, on appelle **valuation** de S , et on note $\text{val}(S)$, l'entier $\text{Min} \{p \in \mathbb{N} \mid a_p \neq 0\}$. On convient que $\text{val}(0) = +\infty$.

Il est immédiat que la valuation des séries formelles prolonge celle des polynômes éléments de $K[X]$. De plus :

THÉORÈME VIII.4.1

Soit S et T deux séries formelles sur K . On a :

(I) $\text{val}(S + T) \geq \text{Min}(\text{val}(S), \text{val}(T))$, et si $\text{val}(S) \neq \text{val}(T)$ il y a égalité.

(II) $\text{val}(ST) = \text{val}(S) + \text{val}(T)$.

Démonstration :

La propriété (I) est immédiate. Vérifions (I)

$S \neq 0$ et $T \neq 0$. Soit $p = \text{val}(S)$ et $q = \text{val}(T)$, où $S = (a_i)_{i \in \mathbb{N}}$ et $T = (b_j)_{j \in \mathbb{N}}$. Cela signifie que $a_i = 0$ si $i < p$, que $b_j = 0$ si $j < q$ et que $a_p b_q \neq 0$. D'après (1) le terme de rang n dans ST est $c_n = \sum_{i+j=n} a_i b_j$. On voit que $c_n = 0$ tant que $n < p + q$ tandis que $c_{p+q} = a_p b_q \neq 0$, d'où $\text{val}(ST) = p + q$. ■

Nous voyons en particulier que pour $S \neq 0$, $T \neq 0$, on a $ST \neq 0$, d'où le :

COROLLAIRE

|| La K -algèbre $K[[X]]$ est **intègre**.

Notre but, en présentant ces quelques rudiments sur les séries formelles, est d'en montrer l'incalculable valeur comme *outil de calcul*. Cela nous oblige à introduire des symboles de sommation qui ont l'apparence de sommes infinies mais ne sont en fait que des façons condensées et efficaces de manipuler des calculs qui portent toujours, en dernier ressort, sur des sommes finies.

DÉFINITION VIII.4.3

Soit $(S_\lambda)_{\lambda \in L}$ une famille de séries formelles sur K , $S_\lambda = (a_{\lambda,n})_{n \in \mathbb{N}}$. Cette famille est dite **sommable ssi**, pour tout $n \in \mathbb{N}$, la famille $(a_{\lambda,n})_{\lambda \in L}$ est à **support fini** dans K . Et si c'est le cas, on peut, pour tout $n \in \mathbb{N}$, définir dans K la somme $\sum_{\lambda \in L} a_{\lambda,n} = c_n$, et la série formelle $(c_n)_{n \in \mathbb{N}}$ est alors appelée **somme** de la famille $(S_\lambda)_{\lambda \in L}$, et notée $\sum_{\lambda \in L} S_\lambda$.

Bien entendu une famille *finie* (S_λ) est toujours sommable, et sa somme est sa somme ordinaire.

Exemple 1 : Soit $S = (a_n)_{n \in \mathbb{N}}$ dans $K[[X]]$. La famille $(a_n X^n)_{n \in \mathbb{N}}$ est évidemment sommable, et sa somme est S d'après la définition précédente. Cela justifie la notation usuelle $\sum_{n \in \mathbb{N}} a_n X^n$ pour désigner S ; si $S \in K[X]$, cette écriture représente la somme finie des monômes du polynôme S , et si $S \in K[[X]]$, elle représente la somme $\sum_{n \in \mathbb{N}} a_n X^n$ au sens de la définition

VIII.4.3. Il s'avère que cette manière d'écrire les séries formelles est très performante.

Exemple 2 : Soit $(S_k)_{k \in \mathbb{N}}$ une famille de séries formelles indexées par \mathbb{N} . Si $\text{val}(S_k) \geq k$ pour tout k , alors la famille (S_k) est sommable : en effet, soit $S_k = \sum_{n \in \mathbb{N}} a_{k,n} X^n$. Pour $n \in \mathbb{N}$ fixé, on a par hypothèse $a_{k,n} = 0$ dès que

$k > n$, donc le support des $(a_{k,n})_{k \in \mathbb{N}}$ est contenu dans $\llbracket 0, n \rrbracket$, donc fini, d'où l'assertion.

En particulier, fixons $S \in K[[X]]$, avec $\text{val}(S) \geq 1$ et soit $(b_n)_{n \in \mathbb{N}}$ une suite à valeurs dans K ; alors pour tout n , $\text{val}(b_n S^n) \geq n$ d'après le théorème VIII.4.1, donc la famille $(b_n S^n)_{n \in \mathbb{N}}$ est sommable.

THÉORÈME VIII.4.2

- (I) Toute sous-famille d'une famille sommable dans $K[[X]]$ l'est encore.
- (II) Si la famille $(S_\lambda)_{\lambda \in L}$ est sommable dans $K[[X]]$, alors pour toute $T \in K[[X]]$, la famille $(S_\lambda T)_{\lambda \in L}$ l'est aussi, et on a
- $$\left(\sum_{\lambda \in L} S_\lambda \right) T = \sum_{\lambda \in L} S_\lambda T.$$
- (III) Soit $(S_\lambda)_{\lambda \in L}$ une famille sommable dans $K[[X]]$ et $(J_\alpha)_{\alpha \in A}$ un partage de L . Alors chaque famille $(S_\lambda)_{\lambda \in J_\alpha}$ est sommable, et si T_α est sa somme, la famille $(T_\alpha)_{\alpha \in A}$ est sommable, et on a :
- $$\sum_{\lambda \in L} S_\lambda = \sum_{\alpha \in A} T_\alpha \quad ({}^1).$$

Nous laisserons le soin au lecteur de vérifier toutes ces assertions en utilisant les résultats du § III.1.

L'exemple 2 débouche sur une nouvelle opération dans $K[[X]]$.

DÉFINITION VIII.4.4

Soit $S \in K[[X]]$ telle que $\text{val}(S) \geq 1$, et $T = \sum_{n \in \mathbb{N}} b_n X^n \in K[[X]]$. On appelle **composée de T par S** (ou **superposée de T et S dans cet ordre**), et on note $T \circ S$ ou $T(S)$, la série formelle $\sum_{n \in \mathbb{N}} b_n S^n$. On dit que $T \circ S$ est obtenue par **substitution de S (à l'indéterminée X) dans T** .

Il est évident que $X \circ S = S$ (pour $\text{val}(S) \geq 1$) et que $T \circ X = T$ (pour T quelconque), ce qui légitime l'écriture $T(X)$ pour $T \in K[[X]]$. Dans le cas particulier où S et T sont dans $K[X]$, la composée $T \circ S$ qu'on vient de définir n'est autre que la composée des polynômes T et S déjà étudiée au § VII.4.

Enfin, on vérifie aisément que $\text{val}(T \circ S) \geq \text{val}(T) \times \text{val}(S)$.

(¹) Il faut prendre garde au fait que cette propriété (III) n'admet pas de réciproque simple. Même si les familles $(S_\lambda)_{\lambda \in J_\alpha}$ sont toutes sommables et si la famille $(T_\alpha)_{\alpha \in A}$ est sommable, cela ne prouve pas que la famille $(S_\lambda)_{\lambda \in L}$ soit sommable.

THÉORÈME VIII.4.3

Soit S, T, U dans $K[[X]]$, avec $\text{val}(S) \geq 1$, $\text{val}(T) \geq 1$ (d'où $\text{val}(T \circ S) \geq 1$). Alors :

(I) L'application $\varphi : K[[X]] \rightarrow K[[X]]$, $V \mapsto V \circ S$ est un homomorphisme de K -algèbres.

(II) On a : $U \circ (T \circ S) = (U \circ T) \circ S$ (associativité).

Démonstration :

(I) La K -linéarité de φ est facile à vérifier. Moins évident est le fait que $\varphi(V_1 V_2) = \varphi(V_1) \varphi(V_2)$ pour V_1 et V_2 dans $K[[X]]$.

Posons :

$$V_i = \sum_{n \in \mathbb{N}} a_{i,n} X^n \quad (i \in \{1, 2\}),$$

d'où
$$V_1 V_2 = \sum_{n \in \mathbb{N}} b_n X^n, \quad \text{avec} \quad b_n = \sum_{p+q=n} a_{1,p} a_{2,q}.$$

On a
$$(V_1 V_2) \circ S = \sum_{n \in \mathbb{N}} b_n S^n = \sum_{n \in \mathbb{N}} \left(\sum_{p+q=n} a_{1,p} S^p \times a_{2,q} S^q \right).$$

Mais la famille $(S^{p+q})_{(p,q) \in \mathbb{N}^2}$ est sommable, d'où par la 3^e partie du théorème VIII.4.2 :

$$\sum_{(p,q) \in \mathbb{N}^2} a_{1,p} a_{2,q} S^p S^q = \sum_{q \in \mathbb{N}} a_{2,q} \left(\sum_{p \in \mathbb{N}} a_{1,p} S^p \right) S^q = (V_1 \circ S)(V_2 \circ S),$$

et aussi, par le même théorème :

$$\sum_{(p,q) \in \mathbb{N}^2} a_{1,p} a_{2,q} S^{p+q} = \sum_{n \in \mathbb{N}} \left(\sum_{p+q=n} a_{1,p} a_{2,q} S^{p+q} \right) = (V_1 V_2) \circ S,$$

donc on a bien $(V_1 V_2) \circ S = (V_1 \circ S)(V_2 \circ S)$.

L'associativité (II) se déduit alors sans difficulté. C'est une conséquence immédiate de (I) lorsque U est un polynôme.

Si U est quelconque, pour $n \in \mathbb{N}$, on écrit $U = U_n + R_n$, avec $\text{val}(R_n) \geq n$ et $U_n \in K[X]$. Alors

$$U \circ (T \circ S) - (U \circ T) \circ S = U_n \circ (T \circ S) - (U_n \circ T) \circ S + W_n = W_n,$$

avec

$$W_n = R_n \circ (T \circ S) - (R_n \circ T) \circ S, \quad \text{d'où} \quad \text{val}(W_n) \geq \text{val}(R_n) \geq n.$$

Ainsi, $\text{val}(U \circ (T \circ S) - (U \circ T) \circ S) \geq n$ pour tout n , donc cette valuation vaut $+\infty$, d'où $U \circ (T \circ S) = (U \circ T) \circ S$. ■

Éléments inversibles de $K[[X]]$ **THÉORÈME VIII.4.4**

Soit $S = \sum_{n \in \mathbb{N}} a_n X^n \in K[[X]]$. Pour que S soit inversible dans

|| l'algèbre $K[[X]]$, il faut et il suffit que $a_0 \neq 0$. Autrement dit, les éléments **inversibles** de $K[[X]]$ sont ceux de **valuation nulle**.

Démonstration :

La condition $a_0 \neq 0$ est bien nécessaire. En effet si S est inversible, et si $T = S^{-1} = \sum_{n \in \mathbb{N}} b_n X^n$, de $ST = 1$ on déduit $a_0 b_0 = 1$, d'où $a_0 \neq 0$. Montrons que la condition $a_0 \neq 0$ est bien suffisante. Pour $T = \sum_{n \in \mathbb{N}} b_n X^n$ quelconque dans $K[[X]]$ on a :

$$ST = \sum_{n \in \mathbb{N}} c_n X^n \quad \text{avec} \quad c_0 = a_0 b_0, \dots, c_n = \sum_{p+q=n} a_p b_q, \dots$$

Pour qu'on ait $ST = 1$ il faut et il suffit que les relations suivantes soient satisfaites :

$$(\mathcal{R}_0) \quad a_0 b_0 = 1, \quad (\mathcal{R}_n) \quad a_0 b_n + \sum_{k=1}^n a_k b_{n-k} = 0 \quad \text{pour } n \geq 1.$$

Or, par récurrence, on voit que le « système » de toutes ces équations aux inconnues (b_i) admet une et une seule solution. En effet (\mathcal{R}_0) fournit $b_0 = a_0^{-1}$, et supposant b_0, b_1, \dots, b_{n-1} déjà déterminés pour $n-1 \geq 0$, la relation (\mathcal{R}_n) fournit b_n de manière unique :

$$b_n = -a_0^{-1} \left(\sum_{k=1}^n a_k b_{n-k} \right). \quad \blacksquare$$

Exemple 3 : La série formelle $1 - X$ est inversible. Son inverse est la série $\sum_{n \geq 0} X^n$, ce qui se vérifie immédiatement en effectuant le produit

$$(1 - X) \left(\sum_{n \geq 0} X^n \right) = \sum_{n \geq 0} X^n - \sum_{n \geq 0} X^{n+1} = 1.$$

Exemple 4 : Dans la relation $(1 - X) \left(\sum_{n \geq 0} X^n \right) = 1$, substituons à X une série formelle $S \in K[[X]]$ de *valuation* ≥ 1 . Le théorème VIII.4.3 entraîne : $(1 - S) \left(\sum_{n \geq 0} S^n \right) = 1$, ce qui donne une expression utile de l'inverse de $1 - S$:

(2)

$$(1 - S)^{-1} = \sum_{n \geq 0} S^n.$$

Un calcul du quotient dans $K[[X]]$

Pour $S = \sum_{n \in \mathbb{N}} a_n X^n \in K[[X]]$, et $N \in \mathbb{N}$, nous noterons S_N la somme des $N + 1$ premiers termes de S : $S_N = \sum_{n=0}^N a_n X^n$, de sorte que

$$\text{val}(S - S_N) \geq N + 1.$$

THÉORÈME VIII.4.5

Soit $S \in K[[X]]$, et $T \in K[[X]]$ avec T inversible ; posons $Q = ST^{-1}$. Alors, pour tout $N \in \mathbb{N}$, Q_N est le quotient dans la division suivant les puissances croissantes de S_N par T_N .

Autrement dit, le quotient Q de S par T peut être calculé en effectuant la division suivant les puissances croissantes de S par T « indéfiniment ».

Démonstration :

Fixons $N \in \mathbb{N}$ et posons $R_N = S - S_N$, $\rho_N = T - T_N$, de sorte que $\text{val}(R_N) \geq N + 1$, $\text{val}(\rho_N) \geq N + 1$, $\text{val}(T_N) = \text{val}(T) = 0$. La différence $ST^{-1} - S_N T_N^{-1}$ vaut $(TT_N)^{-1}(R_N T_N - \rho_N S_N)$ et a donc une valuation $\geq N + 1$. Procédons à la division suivant les puissances croissantes de S_N par T_N à l'ordre N : $S_N = Q_N T_N + \sigma_N T_N^{N+1}$, $\sigma_N \in K[X]$, et comme T_N est inversible dans $K[[X]]$, $S_N T_N^{-1} = Q_N + (T_N^{-1} \sigma_N) T_N^{N+1}$ d'où $\text{val}(S_N T_N^{-1} - Q_N) \geq N + 1$, et par suite

$$\text{val}(ST^{-1} - Q_N) \geq \min(\text{val}(ST^{-1} - S_N T_N^{-1}), \text{val}(S_N T_N^{-1} - Q_N)) \geq N + 1,$$

d'où le résultat. ■

Dérivation dans $K[[X]]$

Si $S = \sum_{n \in \mathbb{N}} a_n X^n \in K[[X]]$, on appelle **dérivée** (par rapport à X) de S , et on note S' où $S'(X)$ la série formelle $\sum_{n \in \mathbb{N}} (n + 1) a_{n+1} X^n$.

L'application de *dérivation* : $K[[X]] \rightarrow K[[X]]$, $S \mapsto S'$ est K -linéaire et un calcul élémentaire montre que $(ST)' = S' T + ST'$ pour tous S et T dans $K[[X]]$. On constate que cette dérivation *prolonge celle déjà définie sur* $K[X]$. Soit S et T dans $K[[X]]$ avec $\text{val}(S) \geq 1$. Alors $(T \circ S)' = (T' \circ S) \times S'$. En effet si $T = X^n$ ($n \in \mathbb{N}$) cela résulte par récurrence de la formule de dérivation d'un produit et on passe de là au cas où T est quelconque par un raisonnement analogue à celui du théorème VIII.4.3.

Enfin, si S est inversible, on obtient en dérivant $SS^{-1} = 1$: $(S^{-1})' = -\frac{S'}{S^2}$.

Si K est de caractéristique nulle, il est clair que $S \mapsto S'$ est :

que son noyau est l'ensemble K des séries constantes. Les *primitives* de $S = \sum_{n \geq 0} a_n X^n$ (c.-à-d. les séries dont la dérivée est S) sont évidemment les

séries $C + \sum_{n \geq 0} \frac{a_n}{n+1} X^{n+1}$, où $C \in K$.

Exercice 1 : Soit A un sous-anneau du corps commutatif K .

a) Montrer que les séries formelles $S = \sum_{n \in \mathbb{N}} a_n X^n \in K[[X]]$ telles que $a_n \in A$ pour tout n forment un sous-anneau noté $A[[X]]$ de $K[[X]]$.

b) Si $S \in A[[X]]$, pour que S soit inversible dans $A[[X]]$, il faut et il suffit que a_0 soit inversible dans A .

Exercice 2 : Vérifier que $K[[X]]$ est un anneau principal, et ne possède, aux facteurs inversibles près de K^* , qu'un seul élément irréductible, qui est X . L'anneau $K[[X]]$ est donc factoriel. Quelle est la factorisation en éléments irréductibles de $S \in K[[X]] \setminus \{0\}$? Vérifier enfin que X engendre l'unique *idéal maximal* de $K[[X]]$.

Exercice 3 : Soit \mathcal{F} l'ensemble des suites $S = (a_n)_{n \in \mathbb{Z}}$ d'éléments de K à support borné à gauche, c'est-à-dire pour lesquelles existe N_S tel que $a_n = 0$ pour tout $n < N_S$. Vérifier que \mathcal{F} est un sous- K -ev du K -ev $K^{\mathbb{Z}}$.

Montrer que (1) définit un *produit* sur \mathcal{F} , et que muni de ce produit et de sa structure de groupe additif, \mathcal{F} devient un *corps commutatif*. Ce corps est formé des sommes formelles $\sum_{n \geq N_S} a_n X^n$, où $N_S \in \mathbb{Z}$ et $a_n \in K$ pour tout n : on l'appelle *corps des séries formelles méromorphes*. Vérifier que c'est le corps des fractions de l'anneau intègre $K[[X]]$. On le note $K((X))$. Un élément non nul de $K((X))$ s'écrit de façon unique $X^p S(X)$, avec $p \in \mathbb{Z}$ et $S \in K[[X]]$, $\text{val}(S) = 0$. Prolonger à $K((X))$ la valuation des séries formelles.

Exercice 4 (réversion d'une série formelle).

Soit $S = \sum_{n \geq 1} a_n X^n \in K[[X]]$ une série formelle de valuation ≥ 1 .

a) Montrer que, pour qu'il existe $T \in K[[X]]$ telle que $T \circ S = X$, il faut et il suffit que $\text{val}(S) = 1$, et qu'alors T est unique.

Indication : Si $\text{val}(S) = 1$, écrire $S^k = \sum_{n \geq k} a_{k,n} X^n$ avec $a_{k,k} = a_1^k \neq 0$. Calculer les coefficients de $T \cdot S$ à l'aide des $a_{k,n}$ et des coefficients de T , et raisonner comme dans le théorème VIII.4.4.

b) Montrer que, pour qu'il existe $U \in K[[X]]$ telle que $\text{val}(U) \geq 1$ et $S \circ U = X$, il faut et il suffit que $\text{val}(S) = 1$, et qu'alors U est unique.

Indication : Si $\text{val}(S) = 1$, soit $U = \sum_{n \geq 1} b_n X^n \in K[[X]]$ de valuation ≥ 1 . Constater que $U^k = \sum_{n \geq k} b_{k,n} X^n$ avec $b_{1,n} = b_n$ et $b_{k,n}$ = expression polynomiale en b_1, b_2, \dots, b_{n-1} seuls si $n \geq 2$. Ecrire les coefficients de $S \circ U$ et raisonner comme dans le théorème VIII.4.4.

c) Si $\text{val}(S) = 1$, et si T et U sont calculés tels que $T \circ S = S \circ T = X$, alors $T = U$ (utiliser la partie (II) du théorème VIII.4.3). La série $T = U$ ainsi définie se note $S^{<-1>}$ et s'appelle *réciproque* de S , appellation justifiée par le fait que si $Y = S(X)$, alors $X = T(Y)$.

Exercice 5 : a) Soit $S \in K[[X]] \setminus \{0\}$ avec $\text{val}(S) \geq 1$. Montrer que l'homomorphisme de K -algèbres $\varphi : K[[X]] \rightarrow K[[X]]$, $T \mapsto T \circ S$, est injectif. Prouver que φ est *bijectif* ssi $\text{val}(S) = 1$ (utiliser l'exercice 4).

b) Quels sont les automorphismes de la K -algèbre $K[[X]]$?

§ VIII.5 APPLICATIONS DES SÉRIES FORMELLES

K désigne toujours un corps commutatif.

Fractions rationnelles et séries formelles

Nous désignerons par $K(X)_0$ l'ensemble des fractions rationnelles $F \in K(X)$ qui n'admettent pas 0 pour pôle. Leur forme irréductible est donc (si $F \neq 0$) $F = \frac{P}{Q}$ avec $Q(0) \neq 0$. L'ensemble $K(X)_0$ est une sous- K -algèbre du corps $K(X)$, qui contient l'ensemble des polynômes de $K[X]$. Nous allons voir que cette sous- K -algèbre s'identifie de façon naturelle avec une sous-algèbre de $K[[X]]$.

THÉORÈME VIII.5.1

Il existe un, et un seul, homomorphisme de K -algèbres $\psi : K(X)_0 \rightarrow K[[X]]$ qui prolonge l'injection canonique $K[X] \rightarrow K[[X]]$; cet homomorphisme ψ est **injectif**.

Démonstration (abrégée) :

Si ψ existe, soit $\frac{P}{Q}$ la forme irréductible de $F \in K(X)_0 \setminus \{0\}$. On a $Q(0) \neq 0$, donc Q est inversible dans $K[[X]]$. Alors

$$\psi\left(\frac{P}{Q}\right) \times \psi(Q) = Q \times \psi\left(\frac{1}{Q}\right) \times \psi(P) = (Q \times Q^{-1}) \times P = P, \text{ d'où :}$$

$$\psi\left(\frac{P}{Q}\right) = \psi(F) = P Q^{-1}.$$

Réciproquement si F est un élément de $K(X)_0$ mis sous forme irréductible $\frac{P}{Q}$, comme $Q(0) \neq 0$, Q est inversible dans $K[[X]]$ et l'élément $P Q^{-1}$ de $K[[X]]$ ne dépend pas de la forme irréductible choisie, mais seulement de F . Si on le note $\psi(F)$ et qu'on pose $\psi(0) = 0$, on vérifie que ψ ainsi définie convient. ■

Si $F \in K(X)_0$, la série formelle $\psi(F)$ s'appelle le **développement en série formelle de F** . Les séries formelles ainsi obtenues s'appellent les **séries formelles rationnelles**. Nous allons voir que cela impose pour leurs coefficients des conditions draconiennes, ce qui montre leur rareté dans $K[[X]]$ mais n'enlève rien à leur intérêt. En effet pour que $S \in K[[X]]$ soit rationnelle, il faut et il suffit qu'il existe $Q \in K[X]$ tel que $Q(0) \neq 0$ et que $SQ \in K[X]$, ce qui, en écrivant le coefficient général du produit SQ , conduit au

THÉOREME VIII.5.2

Soit $S = \sum_{n \in \mathbb{N}} a_n X^n \in K[[X]]$. Pour que S soit **rationnelle**, il faut et il suffit qu'il existe $p \in \mathbb{N}^*$ et $\lambda_1, \lambda_2, \dots, \lambda_p$ dans K tels que :

$$\forall n \geq p \quad a_n = \lambda_1 a_{n-1} + \lambda_2 a_{n-2} + \dots + \lambda_p a_{n-p}.$$

Pour exprimer que $S \in K[[X]]$ est le développement en série formelle de $F \in K(X)_0$, nous écrirons simplement : $F = S$.

Exemple 1 : Calcul des polynômes de Tchebychev de seconde espèce. Le corps de base est \mathbb{C} . Pour $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ on considère la fraction $F \in \mathbb{C}(X)_0$ définie par $F = \frac{1}{1 - 2X \cos \theta + X^2}$.

Nous allons calculer de deux façons différentes le développement de F en série formelle :

1^{re} manière. On pose $S = 2X \cos \theta - X^2$, d'où $F = \frac{1}{1 - S}$ et d'après la relation (2) du § VIII.4 il vient :

$$F = \sum_{n \geq 0} S^n = \sum_{n \geq 0} \left(\sum_{k=0}^n (-1)^k \binom{n}{k} 2^{n-k} \cos^{n-k} \theta \times X^{n+k} \right),$$

d'où (par la définition même de la somme de la famille sommable $(S^n)_{n \in \mathbb{N}}$) :

$$F = \sum_{N \in \mathbb{N}} u_N(\theta) X^N, \quad \text{avec} \quad u_N(\theta) = \sum_{\substack{n+k=N \\ 0 \leq k \leq n}} (-1)^k \binom{n}{k} 2^{n-k} \cos^{n-k} \theta,$$

soit

$$(1) \quad u_N(\theta) = \sum_{k \geq 0, 2k \leq N} (-1)^k \binom{N-k}{k} 2^{N-2k} \cos^{N-2k} \theta.$$

2^e manière. On décompose F en éléments simples sur \mathbb{C} : elle possède les deux pôles simples $e^{i\theta}$ et $e^{-i\theta}$ et $F = \frac{1}{(1 - X e^{-i\theta})(1 - X e^{i\theta})}$ prend la forme

$$(2) \quad F = \frac{i}{2 \sin \theta} \left[\frac{e^{-i\theta}}{1 - X e^{-i\theta}} - \frac{e^{i\theta}}{1 - X e^{i\theta}} \right],$$

d'où en développant chaque terme :

$$(3) \quad F = \frac{i}{2 \sin \theta} \sum_{N \geq 0} X^N (e^{-(N+1)i\theta} - e^{(N+1)i\theta}) = \sum_{N \geq 0} X^N \frac{\sin(N+1)\theta}{\sin \theta}.$$

Puisque la série formelle trouvée en (2) représente $F = \sum_{n \geq 0} u_n(\theta) X^n$, les coefficients sont les mêmes (on dit qu'on **identifie** terme à terme), ce qui donne, pour tout $N \in \mathbb{N}$:

$$\frac{\sin(N+1)\theta}{\sin \theta} = \sum_{k \geq 0, 2k \leq N} (-1)^k \binom{N-k}{k} 2^{N-2k} \cos^{N-2k} \theta.$$

On en déduit l'expression exacte du polynôme de Tchebychev de seconde espèce $U_N(X)$ pour $N \in \mathbb{N}$:

$$U_N(X) = \sum_{k \geq 0, 2k \leq N} (-1)^k \binom{N-k}{k} 2^{N-2k} X^{N-2k}.$$

Nous laissons en exercice 1 le calcul analogue des polynômes de Tchebychev de première espèce.

D'une façon générale, pour développer une fraction rationnelle $F \in K(X)_0$ en série formelle, on peut utiliser la relation (2) du § VIII.4. Mais ce n'est vraiment agréable que si S a une forme assez simple comme ci-dessus. Une autre manière de développer F en série formelle particulièrement intéressante, si le dénominateur de la forme irréductible de F est *dissocié sur K* , consiste à décomposer F en éléments simples et à développer chaque élément simple. Pour cela il suffit de savoir développer en série formelle les fractions de la forme $\frac{1}{(1-aX)^k}$ pour $k \in \mathbb{N}^*$, ce qui (en prenant aX comme nouvelle indéterminée) fait l'objet du théorème suivant :

THÉORÈME VIII.5.3

$$\left\| \begin{array}{l} \text{Pour } n \in \mathbb{N}, \text{ le développement en série formelle de la fraction} \\ \frac{1}{(1-X)^{n+1}} \in K(X)_0 \text{ est donné par} \\ \boxed{\frac{1}{(1-X)^{n+1}} = \sum_{p \geq 0} \binom{p+n}{p} X^p} \end{array} \right\|.$$

Démonstration :

Le théorème est évident pour $n = 0$, car alors il se réduit à l'expression déjà connue de l'inverse de $1 - X$ dans $K[[X]]$.

Supposons-le vrai à l'ordre $n \geq 0$. Alors $\frac{1}{(1-X)^{n+2}} = \frac{1}{(1-X)^{n+1}} \times \frac{1}{1-X}$, ce qui donne pour le développement en série formelle de $\frac{1}{(1-X)^{n+2}}$:

$$S = \left(\sum_{p \geq 0} \binom{p+n}{p} X^p \right) \left(\sum_{q \geq 0} X^q \right) = \sum_{N \geq 0} X^N \left(\sum_{p=0}^N \binom{p+n}{p} \right).$$

Mais grâce à la formule de Pascal, la dernière somme

$$\sum_{p=0}^N \binom{p+n}{p} = \sum_{p=0}^N \binom{n+p}{n}$$

se réduit immédiatement à $\binom{N+n+1}{N}$, d'où le résultat. ■

Exemple 2 : Développement en série formelle de

$$F = \frac{1}{(1-X^2)(1-X^3)} \in \mathbb{C}(X)_0.$$

1^{re} manière : On a

$$F = \frac{1}{1-X^2} \times \frac{1}{1-X^3} = \left(\sum_{p \geq 0} X^{2p} \right) \left(\sum_{q \geq 0} X^{3q} \right) = \sum_{N \in \mathbb{N}} u_N X^N,$$

u_N étant le nombre de couples $(p, q) \in \mathbb{N}^2$ tels que $2p + 3q = N$.

2^{ème} manière : On décompose $F = \frac{1}{(1+X)(1-X)^2(1+X+X^2)}$ en éléments simples sur \mathbb{C} , ce qui donne

$$F = \frac{1/4}{1+X} + \frac{1/6}{(1-X)^2} + \frac{1/4}{1-X} + \frac{\frac{1-j^2}{9}}{1-jX} + \frac{\frac{1-j}{9}}{1-j^2X}$$

et le développement en série formelle de F est donc, après regroupement des termes conjugués :

$$F = \sum_{N \geq 0} X^N \left[\frac{1}{4} (-1)^N + \frac{1}{6} (N+1) + \frac{1}{4} + \frac{1}{9} s_N - \frac{1}{9} s_{N+2} \right]$$

avec $s_k = j^k + j^{2k}$ qui vaut 2 si $k \equiv 0 \pmod{3}$ et -1 si $k \not\equiv 0 \pmod{3}$. La comparaison des deux résultats donne le nombre de couples $(p, q) \in \mathbb{N}^2$ qui vérifient $2p + 3q = N$, à savoir :

si $N \equiv 0 \pmod{3}$, c'est-à-dire $N = 3m$,

$$u_N = \frac{1}{4} (1 + (-1)^m) + \frac{1}{6} (3m+1) + \frac{1}{3}$$

si $N \equiv 1 \pmod{3}$, c'est-à-dire $N = 3m+1$,

$$u_N = \frac{1}{4} (1 + (-1)^{m+1}) + \frac{1}{6} (3m+2) - \frac{1}{3}$$

si $N \equiv 2 \pmod{3}$, c'est-à-dire $N = 3m+2$,

$$u_N = \frac{1}{4} (1 + (-1)^m) + \frac{1}{2} (m+1).$$

Le lecteur pourra retrouver cela en effectuant un dénombrement direct.

Il importe de remarquer que le **plongement** de $K(X)_0$ dans $K[[X]]$ est **compatible avec les dérivations**, c'est-à-dire que si $F \in K(X)_0$ admet le développement en série formelle S , alors S' est celui de la *fraction* dérivée F' . En effet, c'est évident lorsque F est un polynôme puisqu'il n'y a rien à développer. Et si F n'est pas polynomiale, on l'écrit $F = \frac{P}{Q}$, P et Q dans $K[X]$, d'où $F' = \frac{P'Q - PQ'}{Q^2}$ dans $K(X)$, et aussi, par les propriétés opératoires de la dérivation dans $K[[X]]$ (qui sont les mêmes que dans $K(X)$), $S' = (PQ^{-1})' = Q^{-2}(QP' - PQ')$ d'où notre assertion. On pourrait retrouver ainsi par exemple le développement en série formelle de $\frac{1}{(1-X)^{n+1}}$ en dérivant n fois celui de $\frac{1}{1-X}$, mais seulement dans le cas où K est de caractéristique nulle.

Séries formelles usuelles

Le corps de base étant \mathbb{C} , on définit les séries formelles suivantes, dont le nom a été bien entendu choisi parce que ce sont les séries formelles associées aux fonctions usuelles de même appellation étudiées en Analyse :

$$\exp(aX) = \sum_{n \geq 0} \frac{a^n}{n!} X^n \quad (a \in \mathbb{C}) \quad \text{dite série exponentielle}$$

$$\operatorname{ch} X = \sum_{n \geq 0} \frac{1}{(2n)!} X^{2n}; \quad \operatorname{sh} X = \sum_{n \geq 0} \frac{1}{(2n+1)!} X^{2n+1}$$

$$\cos X = \sum_{n \geq 0} \frac{(-1)^n}{(2n)!} X^{2n}; \quad \sin X = \sum_{n \geq 0} \frac{(-1)^n}{(2n+1)!} X^{2n+1}$$

$$\operatorname{Log}(1+X) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} X^n$$

$$(1+X)^\alpha = \sum_{n \geq 0} \binom{\alpha}{n} X^n \quad \text{où } \alpha \in \mathbb{C}, \quad \text{et où :}$$

$$\binom{\alpha}{0} = 1 \quad \text{et} \quad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$$

si $n \geq 1$, n entier, dite *série du binôme*, par référence à la formule du binôme (cf. § III.2)

$$\operatorname{Argth} X = \sum_{n \geq 0} \frac{1}{2n+1} X^{2n+1}; \quad \operatorname{Arctg} X = \sum_{n \geq 0} \frac{(-1)^n}{2n+1} X^{2n+1}$$

$$\operatorname{Arcsin} X = \sum_{n \geq 0} \frac{1}{2n+1} \frac{1 \cdot 3 \dots (2n-1)}{2 \cdot 4 \dots 2n} X^{2n+1};$$

$$\operatorname{Argsh} X = \sum_{n \geq 0} \frac{(-1)^n}{2n+1} \frac{1 \cdot 3 \dots (2n-1)}{2 \cdot 4 \dots 2n} X^{2n+1}.$$

Si $\alpha \in \mathbb{N}$ on constate que $(1 + X)^\alpha$ redonne la formule du binôme de Newton, donc la notation $\binom{\alpha}{n}$ pour $\alpha \in \mathbb{C}$ reste cohérente. Pour $\alpha = -n - 1$, avec $n \in \mathbb{N}$, en appliquant la formule ci-dessus à $(1 - X)^\alpha$ on obtient bien le développement en série formelle de la fraction $\frac{1}{(1 - X)^{n+1}}$ trouvée précédemment, et donc la série du binôme $(1 - X)^{-n-1}$ représente cette fraction, ce qui montre encore la cohérence de la notation choisie.

THÉORÈME VIII.5.4

- (I) Pour $a \in \mathbb{C}$ et $b \in \mathbb{C}$, on a :
- $$\exp(aX) \exp(bX) = \exp((a + b)X).$$
- (II) Pour $\alpha \in \mathbb{C}$ et $\beta \in \mathbb{C}$, on a :
- $$(1 + X)^\alpha (1 + X)^\beta = (1 + X)^{\alpha + \beta}.$$
- (III) Pour $\alpha \in \mathbb{C}$, $(1 + X)^\alpha = \exp(\alpha \operatorname{Log}(1 + X))$.
- (IV) $\exp(\operatorname{Log}(1 + X)) = 1 + X$ et $\operatorname{Log}(1 + (\exp(X) - 1)) = X$.
- (V) Pour $a \in \mathbb{C}$ et $b \in \mathbb{C}$,
- $$\operatorname{Log}((1 + aX)(1 + bX)) = \operatorname{Log}(1 + aX) + \operatorname{Log}(1 + bX).$$

Démonstration :

Remarquons que les deux formes de la formule (IV) sont équivalentes, exprimant toutes deux que les séries formelles $\exp(X) - 1$ et $\operatorname{Log}(1 + X)$ sont réciproques l'une de l'autre (cf. exercice 4, § VIII.4). Commençons par l'assertion (I) :

$$\begin{aligned} \exp(aX) \exp(bX) &= \left(\sum_{p \geq 0} \frac{a^p}{p!} X^p \right) \left(\sum_{q \geq 0} \frac{b^q}{q!} X^q \right) = \\ &= \sum_{n \geq 0} \frac{1}{n!} X^n \left(\sum_{p+q=n} \frac{n!}{p! q!} a^p b^q \right) = \sum_{n \geq 0} \frac{1}{n!} (a + b)^n X^n = \exp((a + b)X). \end{aligned}$$

La relation (II) est vraie si $(\alpha, \beta) \in \mathbb{Z}^2$ à cause des remarques précédant le théorème VIII.5.4. Or :

$$(1 + X)^\alpha (1 + X)^\beta - (1 + X)^{\alpha + \beta} = \sum_{n \geq 0} P_n(\alpha, \beta) X^n,$$

avec

$$P_n(\alpha, \beta) = \left[\sum_{p+q=n} \binom{\alpha}{p} \binom{\beta}{q} \right] - \binom{\alpha + \beta}{n}.$$

Fixons n ; la fonction $P_n : \mathbb{C}^2 \rightarrow \mathbb{C}$, $(\alpha, \beta) \mapsto P_n(\alpha, \beta)$ est pc

anticipant sur les notions données au Chap. X), et elle est nulle sur \mathbb{Z}^2 comme on vient de le voir. Donc elle est nulle sur \mathbb{C}^2 (cf. cor. 2 du théorème X.1.2), et par suite tous les $P_n(\alpha, \beta)$ sont nuls, ce qui prouve (II). Pour démontrer (IV), posons $S = \text{Log}(1 + (\exp(X) - 1))$. Le terme constant de S est 0. Pour que S et X soient des séries formelles égales, il suffit donc que leurs dérivées le soient, i.e. que $S' = 1$.

Or les formules de définition de $\exp(X)$ et de $\text{Log}(1 + X)$ montrent que leurs dérivées sont respectivement $\exp(X)$ et $\frac{1}{1+X}$, d'où par dérivation de séries composées, si $T = \exp(X) - 1$: (cf. fin du § VIII.4)

$$S' = \frac{1}{1 + [\exp(X) - 1]} \quad T' = \frac{1}{\exp(X)} \exp(X) = 1,$$

et par suite $S = X$, d'où (IV).

La même méthode s'applique pour la première partie de l'assertion (IV), et pour les assertions (III) et (V). (Pour (III), les formules de définition de $(1 + X)^\alpha$ montrent que

$$\frac{d}{dX} (1 + X)^\alpha = \alpha (1 + X)^{\alpha-1} = (\text{d'après (II)}) = \frac{\alpha}{1 + X} (1 + X)^\alpha =$$

(par dérivation de séries composées) $= \frac{d}{dX} (\exp(\alpha \text{Log}(1 + X)))$; pour (V), on pose : $U = (1 + aX)(1 + bX) - 1 = (a + b)X + abX^2$ et on dérive les deux membres). ■

Des formules (I) et (II) on déduit les conséquences immédiates :

$$(\exp(aX))^{-1} = \exp(-aX), \quad ((1 + X)^\alpha)^{-1} = (1 + X)^{-\alpha}$$

et également, pour $k \in \mathbb{N}^*$, le fait que si $S = (1 + X)^{\frac{1}{k}}$, on a $S^k = 1 + X^{(1)}$.

Exemple 3 : Montrons, pour terminer ce chapitre, comment les séries formelles peuvent être utiles pour résoudre des problèmes combinatoires (on en a déjà eu un aperçu avec l'exemple 2) : soit à calculer le nombre D_n de *dérangements* dans \mathfrak{S}_n (une permutation $\sigma \in \mathfrak{S}_n$ est appelée **dérangement**, ou permutation « sans rencontre » si elle n'a *pas de point fixe*, c'est-à-dire si $\sigma(k) \neq k$ pour tout $k \in \llbracket 1, n \rrbracket$). On convient que $D_0 = 1$ et $D_1 = 0$.

Un partage de \mathfrak{S}_n obtenu en classant les permutations suivant le nombre

(¹) De manière générale, du Th. VIII.5.4 on déduit aisément que toutes les relations algébriques vérifiées par les fonctions puissance, trigonométrique ou exponentielle de l'Analyse restent vraies avec les séries formelles de même appellation.

de leurs points fixes donne tout de suite, pour $n \geq 2$, la relation :

$$D_n + \binom{n}{1} D_{n-1} + \cdots + \binom{n}{n} D_0 = n!$$

Mais comment tirer parti de cette relation pour obtenir une expression de D_n ? Voici un moyen simple ; en divisant par $n!$, la relation précédente s'écrit :

$$(4) \quad 1 = \sum_{p=0}^n \frac{1}{p!} D_p \times \frac{1}{(n-p)!},$$

et cela reste exact pour $n = 0$ et pour $n = 1$.

Introduisons alors la série formelle $S = \sum_{n \geq 0} \frac{1}{n!} D_n X^n \in \mathbb{C}[[X]]$. Les relations (4), vraies pour tout $n \in \mathbb{N}$, traduisent, d'après la règle du produit de deux séries formelles, l'unique équation :

$$S \times \exp(X) = \sum_{n \geq 0} X^n,$$

et on reconnaît à droite la série $\frac{1}{1-X}$. On obtient donc $S = \frac{1}{1-X} \cdot \exp(-X)$. Il suffit maintenant de développer ce produit pour obtenir, avec un minimum de calculs, la valeur explicite de D_n :

$$(\forall n \in \mathbb{N}) \quad D_n = n! \left(\sum_{p=0}^n \frac{(-1)^p}{p!} \right).$$

Comparée à d'autres méthodes permettant de résoudre le même problème, par exemple celle qui consiste à obtenir d'abord la relation de récurrence

$$(5) \quad D_n = (n-1)(D_{n-1} + D_{n-2})$$

qui n'a rien d'évident, et à résoudre ensuite l'équation (5), il est certain que la méthode indiquée ci-dessus est avantageuse.

Exercice 1 : Soit $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. Le corps de base est \mathbb{C} . Calculer de deux manières le développement en série formelle de la fraction $F = \frac{1 - X \cos \theta}{1 - 2X \cos \theta + X^2}$, et en déduire le calcul complet des polynômes de Tchebychev de 1^{re} espèce (T_n) définis par $(\forall \theta \in \mathbb{R}) \cos n\theta = T_n(\cos \theta)$.

Exercice 2 : Soit des entiers $\alpha_1, \alpha_2, \dots, \alpha_p$ ($p \geq 2, 1 \leq \alpha_1 < \alpha_2 < \dots < \alpha_p$). Le corps de base est \mathbb{C} et on considère la fraction $F = \frac{1}{(1 - X^{\alpha_1})(1 - X^{\alpha_2}) \dots (1 - X^{\alpha_p})}$ dont on note $S = \sum_{n \geq 0} L_n(\alpha_1, \dots, \alpha_p) X^n$ le développement en série formelle.

a) Vérifier que $L_n(\alpha_1, \dots, \alpha_p)$ est le nombre de p -uples $(k_1, k_2, \dots, k_p) \in \mathbb{N}^p$ tels que $\alpha_1 k_1 + \alpha_2 k_2 + \dots + \alpha_p k_p = n$.

b) Calculer effectivement S en décomposant F en éléments simples (et en regroupant les termes conjugués) dans les cas suivants : 1) $p = 3$, $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 5$ (nombre de façons de payer n francs avec des pièces de 1 F, 2 F et 3 F) ; 2) $p = 3$, $\alpha_1 = 2$, $\alpha_2 = 3$, $\alpha_3 = 5$ (nombre de façons de constituer un stock de 500 litres d'huile avec des bidons de 2 litres, 3 litres et 5 litres).

Exercice 3 : Développer en série formelle les fonctions rationnelles suivantes, le corps de base étant \mathbb{C} :

$$a) F = \frac{1}{1 + X + X^2 + \dots + X^{n-1}} \quad (n \in \mathbb{N}^*).$$

$$b) F = \frac{1}{(1 - aX)^p (1 - bX)^q}, \quad a \neq b, p \in \mathbb{N}^*, q \in \mathbb{N}^*.$$

$$c) F = \frac{1}{(1 - X^p)(1 - X^q)} \quad \text{avec } p \text{ et } q \text{ premiers entre eux dans } \mathbb{N}^* \text{ (cf. exercice 2)}.$$

Exercice 4 : a) Soit $k \in \mathbb{N}^*$. Montrer que si $S \in \mathbb{C}[[X]]$ est une série donnée de valuation nulle, il existe exactement k éléments $T \in \mathbb{C}[[X]]$ tels que $T^k = S$. On commencera par traiter le cas où $S = 1 + X$.

b) Soit $A \in \mathbb{C}[[X]]$, $B \in \mathbb{C}[[X]]$, a et b les termes constants de A et B . Montrer que si $a^2 - 4b \neq 0$, alors l'équation $T^2 + AT + B = 0$ où la série inconnue $T \in \mathbb{C}[[X]]$ possède exactement deux solutions.

Exercice 5 : Le corps de base est \mathbb{C} . Soit $S = (1 + X)^\alpha$ et $T = \exp(\alpha \operatorname{Log}(1 + X))$, avec $\alpha \in \mathbb{C}$. Montrer que S et T ont le terme constant 1 et que toutes deux vérifient dans $\mathbb{C}[[X]]$ l'équation suivante : $(1 + X)Y' - \alpha Y = 0$. En déduire que $S = T$.

Exercice 6 : Soit A, B, C dans $\mathbb{C}[[X]]$ avec $\operatorname{val}(A) = 0$. Montrer que pour chaque $b \in \mathbb{C}$ il existe un et un seul élément $S \in \mathbb{C}[[X]]$ tel que $AS' + BS = C$, avec terme constant de $S = b$ (se ramener au cas où $A = 1$). Généraliser ce résultat avec une équation de la forme $A_0 S'' + A_1 S' + A_2 S = B$ (les A_i et B dans $\mathbb{C}[[X]]$, $\operatorname{val}(A_0) = 0$) ou d'ordre $k \geq 3$.

Application : Montrer que pour calculer la série formelle $(\operatorname{Arcsin} X)^2$ il est plus commode de résoudre l'équation « différentielle » $(1 - X^2)Y'' - XY' = 2$ que d'effectuer la multiplication. De même pour $(\operatorname{Arg sh} X)^2$, $\operatorname{Log}^2(1 - X)$ et leurs dérivées (former les équations correspondantes et les résoudre).

Exercice 7 (nombre d'involutions dans \mathfrak{S}_n).

Pour $n \in \mathbb{N}$, soit u_n le nombre de permutations σ dans \mathfrak{S}_n telles que $\sigma^2 = \operatorname{Id}$. On pose $u_0 = u_1 = 1$.

a) Montrer : $(\forall n \geq 2) u_n = u_{n-1} + (n-1)u_{n-2}$ (il suffit de classer les involutions en deux sous-ensembles : celles qui fixent n et celles qui dérangent n).

b) On considère la série formelle $S = \sum_{n \geq 0} \frac{u_n}{n!} X^n \in \mathbb{C}[[X]]$. Montrer que $S' - (1 + X)S = 0$

et en déduire que $S = \exp\left(X + \frac{X^2}{2}\right)$ (cf. exercice 6). En déduire l'expression de u_n (réponse :

$$u_n = \sum_{p+2q=n} \frac{n!}{p! q! 2^q}.$$

Exercice 8 : On définit la suite de Fibonacci ⁽¹⁾ $(F_n)_{n \in \mathbb{N}}$ à valeurs dans \mathbb{N} par $F_0 = F_1 = 1$ et

⁽¹⁾ Léonard de Pise (1180-1250), plus connu sous le nom de *Fibonacci*, marchand italien ayant beaucoup voyagé, a écrit en 1202 un « Liber abaci » où sa fameuse suite intervient dans un problème concernant la prolifération d'un couple de lapins.

$F_n = F_{n-1} + F_{n-2}$ pour $n \geq 2$. Au-delà de propriétés arithmétiques faciles telles que : F_n et F_{n+1} sont premiers entre eux, $F_{n+1} \cdot F_{n-1} - F_n^2 = (-1)^{n+1}$, cherchons à évaluer F_n . Pour cela on pose $S = \sum_{n \geq 0} F_n X^n \in \mathbb{C}[[X]]$.

- a) Montrer que $S = \frac{1}{1-X-X^2}$ et en déduire, par décomposition en éléments simples de cette fraction, la valeur exacte de F_n (réponse : $\frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right]$).
- b) Soit $T = \sum F_n^2 X^n$. Montrer que $T = \frac{1-X}{1-2X-2X^2+X^3}$.
- c) Montrer que $F_{n-1}^2 + F_n^2 = F_{2n}$.
- d) Montrer qu'il existe une infinité de valeurs de F_n se terminant par 4 zéros.

Exercice 9 (problème des parenthésages de Catalan).

On pose $P_0 = 0$, $P_1 = 1$; si $n \geq 2$, on note P_n le nombre de manières de parenthéser une suite de n termes égaux à un même élément x pour obtenir leur produit relativement à une loi de composition interne la plus générale, et donc présumée non associative. Par exemple $P_2 = 1$; $P_3 = 2$, les deux parenthésages possibles étant $((xx)x)$ et $(x(xx))$; $P_4 = 5$:

$((xx)x)x$, $((xx)(xx))$, $((x(xx))x)$, $(x((xx)x))$ et $(x(x(xx)))$.

- a) Montrer que $P_n = \sum_{1 \leq k \leq n-1} P_k P_{n-k}$ pour tout $n \geq 2$.
- b) On pose $S = \sum_{n \geq 0} P_n X^n \in \mathbb{C}[[X]]$. Déduire du a) que $S^2 - S + X = 0$. Utiliser l'exercice 4b) pour en déduire $S = \frac{1}{2} [1 - (1 - 4X)^{1/2}]$, et à partir de là : $(\forall n \geq 2) P_n = \frac{1}{n} \binom{2n-2}{n-1}$.
- On pourra vérifier l'exactitude de ce résultat en montrant qu'il satisfait bien la relation de récurrence.

Exercice 10 : Pour $n \geq 2$, soit S_n le nombre des permutations du groupe alterné \mathcal{U}_n dont le nombre de points fixes est pair, et U_n le nombre de dérangements dans \mathcal{U}_n . On pose $U_0 = 1$, $U_1 = 2$, $S_0 = 1$, $S_1 = 0$. Soit $\Phi_1 = \sum_{p \geq 0} \frac{U_p}{p!} X^p \in \mathbb{C}[[X]]$ et $\Psi_1 = \sum_{p \geq 0} \frac{1}{p!} S_p X^p$.

On note aussi T_n le nombre des $\sigma \in \mathcal{S}_n \setminus \mathcal{U}_n$ dont le nombre de points fixes est impair, et A_p le nombre de dérangements dans $\mathcal{S}_p \setminus \mathcal{U}_p$. On pose $A_0 = 0$, $A_1 = 0$, $T_0 = 0$, $T_1 = 0$. Soit $\Phi_2 = \sum_{p \geq 0} \frac{A_p}{p!} X^p$ et $\Psi_2 = \sum_{p \geq 0} \frac{T_p}{p!} X^p$. Montrer que

$$\Phi_1 = \frac{1}{2} \frac{2-X^2}{1-X} \exp(-X), \quad \Psi_1 = \frac{1}{4} \frac{2-X^2}{1-X} (1 + \exp(-2X)),$$

$$\Phi_2 = \frac{1}{2} \frac{X^2}{1-X} \exp(-X) \quad \text{et} \quad \Psi_2 = \frac{1}{4} \frac{X^2}{1-X} (1 - \exp(-2X)).$$

En déduire U_n , S_n , T_n , A_n .

Exercice 11 : Le corps de base est \mathbb{C} . Montrer qu'il n'existe aucun système C_0, C_1, \dots, C_n d'éléments de $\mathbb{C}[[X]]$ ($n \geq 1$, $C_0 \neq 0$, $C_n \neq 0$) tel que la série formelle $S = \sum_{n \geq 0} \frac{1}{n!} X^n = \exp X$ vérifie dans $\mathbb{C}[[X]]$ $C_0 S^n + C_1 S^{n-1} + \dots + C_n = 0$. Même question avec $S = \text{Log}(1+X)$.

Exercice 12 : On considère l'ensemble E des suites de $2n$ chiffres constituées uniquement de 0 et de 1.

a) Quel est le nombre a_n de ces suites où ne figurent jamais deux zéros consécutifs ? Même question avec trois zéros consécutifs.

b) Quel est le nombre u_n de ces suites où il y a autant de 0 que de 1. En déduire le nombre v_n de telles suites où l'égalité entre le nombre de 0 et de 1 se produit pour la première fois au rang $2n$.

Indication : Il y a une relation simple entre $S = \sum_{n \geq 0} u_n X^n$ et $T = \sum_{n \geq 0} v_n X^n$.

Exercice 13 (série de Lambert et fonction de Möbius).

Le corps de base est \mathbb{C} . Pour chaque $n \in \mathbb{N}^*$ on considère la série formelle $S_n = X^n + X^{2n} + X^{3n} + \dots = X^n(1 - X^n)^{-1}$ de valuation n . A toute suite $(a_n)_{n \in \mathbb{N}^*}$ on associe d'une part la série formelle $F(X) = \sum_{n \geq 1} a_n X^n$, et d'autre part la famille de séries formelles $(a_n S_n)_{n \in \mathbb{N}^*}$. Montrer que cette famille est sommable. Sa somme $G(X)$ est appelée *série de Lambert* associée à $(a_n)_{n \in \mathbb{N}^*}$. On l'écrira $G(X) = \sum_{n \geq 1} b_n X^n$.

a) Démontrer que $G(X) = \sum_{m \geq 1} F(X^m)$.

b) On définit la *suite de Möbius* $(\mu_n)_{n \in \mathbb{N}^*}$ par le fait que la série de Lambert qui lui est associée est $G(X) = X$. Après avoir prouvé que $b_n = \sum_{d|n} a_d$ et que $a_n = \sum_{d|n} \mu(d) b_{n/d}$, en déduire la valeur de $\mu(n)$ pour $n = 1$, puis pour $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$. Montrer que la fonction $\mu : \mathbb{N}^* \rightarrow \mathbb{Z}$, $n \mapsto \mu(n)$ est *multiplicative*, en ce sens que si a et b sont premiers entre eux, $\mu(ab) = \mu(a) \mu(b)$.

c) Si $a_n = 1$ pour tout $n \in \mathbb{N}$, montrer que b_n représente le nombre de diviseurs de n . Si $a_n = \varphi(n)$ (indicateur d'Euler), calculer explicitement $G(X) (= X(1 - X)^{-2})$.

Exercice 14 (séries de dérivation).

Soit K un corps commutatif de caractéristique 0 ; on notera \mathcal{E} la K -algèbre $\text{Hom}_K(K[X])$ et D l'élément de \mathcal{E} tel que $D(F) = F'$ pour $F \in K[X]$. Pour toute série formelle $S \in K[[X]]$, $S = \sum_{k \in \mathbb{N}} a_k X^k$, et tout polynôme $F \in K[X]$, on pose : $S(D) \cdot F = \sum_{k \in \mathbb{N}} a_k F^{(k)} = \sum_{k \in \mathbb{N}} a_k D^k(F)$, somme qui a un sens puisque $D^k(F) = 0$ dès que $k > \deg(F)$.

a) Montrer que $S \mapsto S(D)$ définit un homomorphisme de K -algèbres, injectif, de $K[[X]]$ dans \mathcal{E} . On notera \mathcal{S}_D l'image de cet homomorphisme ; si $S = \sum_{k \in \mathbb{N}} a_k X^k$, on écrira

aussi $S(D) = \sum_{k \in \mathbb{N}} a_k D^k$. Ainsi \mathcal{S}_D est une sous- K -algèbre de \mathcal{E} isomorphe à $K[[X]]$.

b) Soit $h \in K$. On lui associe $T_h \in \mathcal{E}$ tel que $T_h(F(X)) = F(X + h)$ pour $F \in K[X]$. Vérifier que $T_h = S_h(D)$, où $S_h = \exp(hX)$. On écrira : $T_h = \exp(hD)$.

c) Soit $f \in \mathcal{E}$ tel que $f \circ D = D \circ f$. Pour $k \in \mathbb{N}$, on pose : $a_k = \frac{1}{k!} [f(X^k)](0)$. Démontrer que $f = S(D)$ avec $S = \sum_{k \in \mathbb{N}} a_k X^k$, d'où $f \in \mathcal{S}_D$.

d) Soit $f \in \mathcal{E}$ tel que $(\forall h \in K) f \circ T_h = T_h \circ f$. Démontrer que $f \circ D = D \circ f$, et en déduire que $f \in \mathcal{S}_D$.

Exercice 15 (nombres de Bernoulli ⁽¹⁾).

Le corps de base est \mathbb{C} . La série formelle $\frac{\exp(X) - 1}{X} = \sum_{k \geq 0} \frac{1}{(k+1)!} X^k$ étant inversible, on note $\sum_{n \geq 0} \frac{b_n}{n!} X^n$ son inverse $\beta(X) = \frac{X}{\exp(X) - 1} = \sum_{n \geq 0} \frac{b_n}{n!} X^n$. Les nombres $(b_n)_{n \geq 0}$ s'appellent *nombres de Bernoulli* : ils sont rationnels puisque $\frac{\exp(X) - 1}{X} \in \mathbb{Q}[[X]]$.

a) Ecrire la relation de récurrence permettant le calcul de b_n en fonction des précédents et en déduire que $|b_n| \leq n! (\forall n \in \mathbb{N})$.

⁽¹⁾ Jakob Bernoulli (1654-1705) est le plus illustre représentant d'une famille de mathématiciens suisses. Son nom reste attaché à une famille de polynômes, à une équation différentielle, à la lemniscate, mais sur sa tombe à Bâle, c'est une spirale logarithmique qui

b) Montrer que $\frac{X}{\exp(X) - 1} + \frac{1}{2}X = \frac{X}{2 \operatorname{th} \frac{X}{2}}$ et en déduire que $b_{2n+1} = 0$ si $n \geq 1$.

c) On définit $X \cotg X = \frac{X}{\operatorname{tg} X}$. En utilisant b) montrer que l'on a dans $\mathbb{C}[[X]]$:

$$X \cotg X = 1 + \sum_{n \geq 1} (-1)^n b_{2n} \frac{2^{2n}}{(2n)!} X^n.$$

A partir de $\operatorname{tg} X = \cotg X - 2 \cotg 2X$ (cette série a-t-elle un sens ?), montrer que

$$\operatorname{tg} X = \sum_{n \geq 1} (-1)^{n+1} \frac{2^{2n}}{(2n)!} (2^{2n} - 1) b_n X^{2n-1}.$$

Développer également $\frac{X}{\exp(X) + 1}$ et $\frac{\exp(X) - 1}{\exp(X) + 1} = 1 - \frac{2}{X} \frac{X}{\exp(X) + 1}$.

d) A partir de $X \cotg X - X \cotg 2X$ trouver également la série formelle de somme $\frac{X}{\sin X}$ en fonction des nombres de Bernoulli.

e) On note \mathcal{C} la \mathbb{C} -algèbre $\operatorname{Hom}_{\mathbb{C}}(\mathbb{C}[X])$. Avec les notations de l'exercice 14, en appelant Δ l'opérateur de différence dans $\mathbb{C}[X]$ défini par :

$$(\Delta F)(X) = F(X+1) - F(X) \quad \text{si } F \in \mathbb{C}[X],$$

démontrer que :

$$\Delta = \gamma(D) \times D = D \gamma(D), \quad \text{avec } \gamma = \frac{1}{\beta} \quad \text{dans } \mathbb{C}[[X]].$$

N.B. : Les nombres de Bernoulli jouent un rôle important en théorie des nombres et en Analyse. Le lecteur pourra calculer les 20 ou 30 premiers b_n . Il devinera que b_{2n} a le signe de

$(-1)^{n-1}$ (cf. exercice 16) et admettra que $|b_{2k}| \underset{+\infty}{\sim} 4\sqrt{\pi} \frac{k^{2k+\frac{1}{2}}}{(\pi)^{2k}}$, ce qui montre une croissance très rapide.

Exercice 16 (polynômes de Bernoulli).

Le corps de base est \mathbb{C} et on désigne par z un paramètre complexe. Soit S_z la série formelle élément de $\mathbb{C}[[X]]$ égale à $\frac{X \exp(zX)}{\exp(X) - 1}$.

a) Montrer que $S_z = \sum_{n \geq 0} \frac{1}{n!} B_n(z) X^n$ avec $B_n(z) = \sum_{k=0}^n \binom{n}{k} b_{n-k} z^k$ pour tout $n \in \mathbb{N}$, les

(b_i) désignant les nombres de Bernoulli (cf. exercice 15).

Les polynômes $B_n(z)$, à coefficients rationnels, s'appellent *polynômes de Bernoulli*.

b) Montrer, pour $n \in \mathbb{N}$: $B_n(0) = b_n$; si $n \geq 2$: $B_n(1) = B_n(0) = b_n$; si $n \geq 0$ et $z \in \mathbb{C}$: $B_n(1-z) = (-1)^n B_n(z)$; si $n \geq 1$ et $z \in \mathbb{C}$: $\sum_{k=0}^{n-1} \binom{n}{k} B_k(z) = nz^{n-1}$ et $B'_n(z) = nB_{n-1}(z)$.

c) En développant de deux manières différentes $\frac{X \left(\exp\left(\frac{X}{2}\right) + 1 \right)}{\exp(X) - 1}$ calculer $B_{2k}\left(\frac{1}{2}\right)$.

d) Montrer que, si $n \geq 0$ et $z \in \mathbb{C}$: $B_{n+1}(z) - B_{n+1}(z-1) = (n+1)(z-1)^n$. En déduire pour $N \in \mathbb{N}^*$ une expression simple de la somme $1^n + 2^n + \dots + N^n$.

e) Montrer que, si $n \in \mathbb{N}$, si $x \in \mathbb{C}$ et si $y \in \mathbb{C}$, $B_n(x+y) = \sum_{k=0}^n \binom{n}{k} B_k(x) y^{n-k}$.

f) Si $x \in \mathbb{C}$, $n \in \mathbb{N}$ et $q \in \mathbb{N}^*$, prouver : $B_n(qx) = q^{n-1} \sum_{k=0}^{q-1} B_n\left(x + \frac{k}{q}\right)$.

g) Développer en série formelle $\frac{\sin(aX)}{\sin\left(\frac{X}{2}\right)}$ pour $a \in \mathbb{C}$. Que se passe-t-il si $a + \frac{1}{2} \in \mathbb{Z}$?

Exercice 17 (polynômes de Hilbert).

On pose $H_0 = 1$ et, si $n \in \mathbb{N}^*$, $H_n(X) = \frac{1}{n!} X(X-1) \dots (X-n+1) =$

a) Montrer que $(H_n)_{n \in \mathbb{N}}$ est une base de $K[X]$ et que, pour tout $n \in \mathbb{N}$, $(H_n)_{0 \leq n \leq N}$ est une base de $K_N[X]$. Prouver : $(\forall n \in \mathbb{N}) H_n(\mathbb{Z}) \subset \mathbb{Z}$.

b) Soit $\tau \in \text{Hom}_K(K[X])$ tel que $\tau(F(X)) = F(X+1)$ pour $F \in K[X]$. Prouver :

$$\forall F \in K[X] \quad F = \sum_{n \geq 0} [(\Delta^n F)(0)] H_n, \quad \text{avec } \Delta = \tau - \text{Id}.$$

c) On donne $F \in K[X]$ et on suppose : $\exists a \in \mathbb{Z}$ tel que $F(\lfloor a, \rightarrow \rfloor) \subset \mathbb{Z}$. Prouver que les coordonnées de F dans la base $(H_n)_{n \in \mathbb{N}}$ appartiennent à \mathbb{Z} .

d) Soit $N \in \mathbb{N}$. On note \mathcal{E}_N le sous-groupe additif de $K[X]$ engendré par les $(H_n)_{0 \leq n \leq N}$ et Z_N celui engendré par les $(X^n)_{0 \leq n \leq N}$. Prouver que le groupe quotient \mathcal{E}_N/Z_N est fini, donner sa structure et calculer son cardinal.

Exercice 18 (nombres de Stirling ⁽¹⁾).

Le corps de base est \mathbb{C} . On note \mathcal{A} la \mathbb{C} -algèbre $\text{Hom}_{\mathbb{C}}(\mathbb{C}[X])$, D l'opérateur de dérivation sur $\mathbb{C}[X]$: $D(F) = F'$, et \mathcal{D} l'opérateur de dérivation dans la \mathbb{C} -algèbre $\mathcal{S} = K[[X]]$, T l'opérateur de translation dans $\mathbb{C}[X]$: $(T(F)(X)) = F(X+1)$ et enfin $\Delta = T - \text{Id}$. Si $(n, k) \in \mathbb{N}^2$, on pose $S(n, k) = \frac{1}{k!} [\Delta^k(X^n)](0)$. Montrer que

$$S(n, k) = \frac{1}{k!} \sum_{0 \leq i \leq k} (-1)^i \binom{n}{i} (k-i)^n$$

(par la formule du binôme, appliqué à $T - \text{Id}$). Les nombres $S(n, k)$ sont appelés *nombres de Stirling de deuxième espèce*.

a) Montrer que les $S(n, k)$ sont entiers (cf. exercice 17) et $S(n, k) = 0$ pour $k > n$. Prouver que les $S(n, k)$ peuvent être définis par les relations de récurrence :

$$(1) \quad S(0, 0) = 1, \quad S(n, 0) = S(0, k) = 0 \quad \text{pour } n \geq 1 \text{ et } k \geq 1$$

$$(2) \quad \forall k \geq 1, \forall n \geq 1 \quad S(n, k) = S(n-1, k-1) + kS(n-1, k).$$

(La relation (2), qu'on peut établir en partant de $X^n = X \cdot X^{n-1}$, permet de construire un triangle de Stirling analogue au triangle de Pascal.)

b) Soit $\mathcal{S}_{n,k}$ le nombre d'applications surjectives de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, k \rrbracket$. On pose : $\mathcal{S}_{0,0} = 1$; $\mathcal{S}_{n,0} = 0$ si $n \geq 1$. Vérifier que $\mathcal{S}_{0,k} = 0$ si $k \geq 1$; $\mathcal{S}_{n,k} = 0$ si $k > n$. Prouver :

$$\forall n \geq 1, \forall k \geq 1 \quad \sum_{i=0}^k \binom{k}{i} \mathcal{S}_{n, k-i} = k^n \quad (\text{nombre total d'applications de } \llbracket 1, n \rrbracket \text{ dans } \llbracket 1, k \rrbracket).$$

c) Soit $n \in \mathbb{N}$. On pose $S_n = \sum_{k \geq 0} \frac{1}{k!} \mathcal{S}_{n,k} X^k$. Montrer que $\exp(X) \times S_n = \sum_{k \geq 0} \frac{k^n}{k!} X^k$. En déduire une expression du polynôme S_n . Montrer que : $(\forall k \in \mathbb{N}) \mathcal{S}_{n,k} = k! S(n, k)$. En déduire que $S(n, k)$ est exactement le *nombre de partitions en k parties* de $\llbracket 1, n \rrbracket$.

$$d) \text{ Montrer que, pour } k \in \mathbb{N}^* : \sum_{n \geq k} \frac{1}{n!} S(n, k) X^n = \frac{1}{k!} (\exp(X) - 1)^k.$$

e) Décomposer en éléments simples sur \mathbb{C} la fraction $\varphi = \frac{X^k}{(1-X)(1-2X) \dots (1-kX)}$. En déduire $\varphi = \sum_{n \geq k} S(n, k) X^{n-k}$.

f) Soit θ l'opérateur linéaire de $\mathbb{C}[[X]]$ dans lui-même envoyant chaque série formelle S sur XS' . Démontrer que, si $\psi = \theta \circ \mathcal{D}$,

$$(\forall n \in \mathbb{N}^*) \quad \psi^n = \sum_{k=1}^n S(n, k) \theta^k \circ \mathcal{D}^k.$$

Utiliser ce résultat pour calculer, à l'aide des nombres $S(n, k)$ les séries formelles suivantes :

$$G_n = \sum_{p \geq 0} p^n X^p ; \quad \Lambda_n = \sum_{p \geq 0} \frac{p^n}{p!} X^p, \quad \text{où } n \in \mathbb{N} \text{ est donné.}$$

⁽¹⁾ James Stirling (1692-1770), mathématicien anglais dont on retient essentiellement la formule donnant un équivalent de $n!$ quand $n \rightarrow +\infty$: $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

Exercice 19 (nombres de Bell).

Si $n \in \mathbb{N}^*$, on pose $\varpi(n) = \sum_{k=1}^n S(n, k)$ où les $S(n, k)$ sont les nombres de Stirling définis dans l'exercice 18. On convient que $\varpi(0) = 1$.

a) Montrer que $\varpi(n)$ n'est autre que le *nombre total de partitions* de $\llbracket 1, n \rrbracket$.

b) On pose $\Theta(X) = \sum_{n \geq 0} \frac{1}{n!} \varpi(n) X^n$. Prouver, à partir du d) de l'exercice 18, que $\Theta(X) = \exp(\exp(X) - 1)$. Calculer alors $\mathcal{D}\Theta(X)$. En déduire la formule de récurrence : $(\forall n \in \mathbb{N}) \varpi(n+1) = \sum_{i=0}^n \binom{n}{i} \varpi(i)$.

c) (Cette question suppose connue la théorie des séries numériques). Montrer, si $n \in \mathbb{N}$, que la série numérique $\sum_p \frac{p^n}{ep!}$ converge ; soit $\gamma(n)$ sa somme. Calculer $\gamma(0)$. Montrer : $(\forall n) \gamma(n+1) = \sum_{i=0}^n \binom{n}{i} \gamma(i)$. En déduire : $(\forall n) \gamma(n) = \varpi(n)$ (formule de G. Dobinski, 1877).

Chapitre IX

ESPACES VECTORIELS ; DIMENSION DES ESPACES VECTORIELS

Dans tout ce chapitre, on considère des espaces vectoriels sur un corps de base commutatif, qui sera noté K en général.

§ IX.1 SOUS-ESPACES SUPPLÉMENTAIRES, PROJECTEURS

DÉFINITION IX.1.1

Soit F et G deux sous-espaces vectoriels d'un K -ev E ; on dit que F et G sont **supplémentaires** ssi on a à la fois :

$$(I) \quad F \cap G = \{0_E\}$$

$$(II) \quad F + G = E ,$$

c'est-à-dire : tout $x \in E$ s'écrit d'une manière sous la forme $x = u + v$, avec $u \in F$ et $v \in G$.

Si F et G sont supplémentaires, il en est de même de G et F . Par exemple $\{0_E\}$ et E sont supplémentaires.

Lorsque F et G sont supplémentaires, on écrit : $E = F \oplus G$.

Exemple 1 : Supposons $K = \mathbb{R}$ et prenons pour E le \mathbb{R} -ev $\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de \mathbb{R} dans \mathbb{R} ; soit F le sous- \mathbb{R} -ev formé des fonctions *paires* (i.e. telles que $f(-x) = f(x)$ pour tout $x \in \mathbb{R}$) et G le sous- \mathbb{R} -ev formé des fonctions *impaires* (i.e. telles que $f(-x) = -f(x)$ pour tout $x \in \mathbb{R}$).

Alors $E = F \oplus G$. En effet, si $f \in F \cap G$, on a $(\forall x \in \mathbb{R}) f(-x) = f(x) = -f(x)$ d'où $f(x) = 0$ et $f = 0_E$. Et si $f \in E$, on a $f = g + h$ avec $g : x \mapsto \frac{1}{2} [f(x) + f(-x)]$ et $h : \frac{1}{2} [f(x) - f(-x)]$, et on voit bien que $g \in F$ et $h \in G$.

Exemple 2 : Soit E un K -ev et $u \in \text{Hom}_K(E)$ tel que $u^2 = \text{Id}_E$ (on suppose

que le corps K n'est pas de caractéristique 2). Soit $F = \text{Ker}(u - \text{Id}_E)$ et $G = \text{Ker}(u + \text{Id}_E)$. Alors $E = F \oplus G$. En effet, si $x \in F \cap G$, on a : $u(x) = x = -x$, d'où $2x = 0_E$, d'où $x = 0_E$; donc $F \cap G = \{0_E\}$. Et si $x \in E$, on a : $x = y + z$, avec $y = \frac{1}{2}(x + u(x))$, $z = \frac{1}{2}(x - u(x))$, d'où $y \in F$ et $z \in G$. On notera l'analogie avec l'exemple 1. Elle n'est pas fortuite : il suffit de considérer dans $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ l'endomorphisme $u : f \mapsto \hat{f}$, avec $\hat{f}(x) = f(-x)$ pour tout $x \in \mathbb{R}$.

Avec les notations de la définition IX.1.1, supposons $E = F \oplus G$, un vecteur $x \in E$ s'écrit alors *de manière unique* $x = y + z$ avec $y \in F$ et $z \in G$. En effet, de $x = y_1 + z_1 = y_2 + z_2$ on tire

$$y_1 - y_2 = z_2 - z_1 \in F \cap G = \{0\},$$

d'où l'unicité. Réciproquement, si l'application $S : F \times G \rightarrow E$, $(y, z) \mapsto y + z$ est bijective, on a $F \cap G = \{0_E\}$ car $x \in F \cap G$ entraîne $S(x, -x) = 0_E = S(0_E, 0_E)$, d'où $(x, -x) = (0_E, 0_E)$, donc $E = F \oplus G$.

Dans ces conditions on peut définir les applications $p : E \rightarrow E$ et $q : E \rightarrow E$ par la condition

$$(1) \quad \forall x \in E \quad x = p(x) + q(x), \quad (p(x), q(x)) \in F \times G.$$

Il est aisé de voir que p et q sont *linéaires* : en effet, de (1) on déduit, pour $\lambda \in K$: $\lambda x = \lambda p(x) + \lambda q(x)$ et $(\lambda p(x), \lambda q(x)) \in F \times G$ (car F et G sont des sous-espaces), d'où $\lambda p(x) = p(\lambda x)$ et $\lambda q(x) = q(\lambda x)$. De même si $x_i \in E$ ($i \in \{1, 2\}$), de $x_i = p(x_i) + q(x_i)$, on déduit :

$$x_1 + x_2 = p(x_1) + p(x_2) + q(x_1) + q(x_2)$$

et $(p(x_1) + p(x_2), q(x_1) + q(x_2)) \in F \times G$ puisque F et G sont des sous-espaces, d'où $p(x_1) + p(x_2) = p(x_1 + x_2)$ et $q(x_1) + q(x_2) = q(x_1 + x_2)$. On remarque en outre que : $\text{Im}(p) = F$, $\text{Im}(q) = G$, $\text{Ker}(p) = G$, $\text{Ker}(q) = F$, $p + q = \text{Id}_E$, $p \circ p = p$, $q \circ q = q$, $p \circ q = q \circ p = 0$. Résumons :

THÉORÈME IX.1.1

Soit F et G deux sous- K -ev du K -ev E .

(I) Pour que F et G soient **supplémentaires**, il faut et il suffit que l'application $F \times G \rightarrow E$, $(y, z) \mapsto y + z$ soit **bijective**.

(II) Si F et G sont supplémentaires, les applications $p : E \rightarrow E$ et $q : E \rightarrow E$ définies par

$$(\forall x \in E) \quad x = p(x) + q(x), \quad (p(x), q(x)) \in F \times G$$

sont linéaires et vérifient $p^2 = p$, $q^2 = q$, $p \circ q = q \circ p = 0$, $p + q = \text{Id}_E$, $\text{Im}(p) = F = \text{Ker}(q)$, $\text{Im}(q) = G =$

On dit que p est la *projection sur F parallèlement à G* et q la projection sur G parallèlement à F .

Projecteurs

DÉFINITION IX.1.2

$\left. \begin{array}{l} \} \text{ Soit } E \text{ un } K\text{-ev ; on appelle } \mathbf{projecteur} \text{ de } E \text{ tout endomorphisme } p \\ \} \text{ de } E \text{ tel que } p^2 = p. \end{array} \right\}$

Nous venons de voir que les projections associées à deux sous-espaces supplémentaires de E sont des projecteurs. Montrons qu'il n'y a pas d'autres projecteurs que ceux-là.

THÉORÈME IX.1.2

$\left\| \begin{array}{l} \text{ Soit } p \text{ un projecteur dans le } K\text{-ev } E. \text{ Posons : } q = \text{Id}_E - p, \\ F = \text{Im } (p), G = \text{Im } (q). \text{ Alors :} \end{array} \right\|$

(I) $E = F \oplus G .$

(II) p et q sont les projections respectivement sur F parallèlement à G , et sur G parallèlement à F .

Démonstration :

Commençons par prouver que $F \cap G = \{0_E\}$. Soit $x \in F \cap G$: $\exists y \in E$ tel que $x = p(y)$; $\exists z \in E$ tel que $x = q(z)$. Alors $p(x) = p^2(y) = p(y) = x = p(z) - p^2(z) = p(z) - p(z) = 0_E$.

Montrons ensuite que $F + G = E$. Si $x \in E$ on peut l'écrire $x = p(x) + (x - p(x)) = p(x) + q(x)$ avec $p(x) \in F$ et $q(x) \in G$. Finalement $E = F \oplus G$. Mais l'écriture précédente montre bien que p et q sont les projections sur F et G associées aux sous-espaces supplémentaires F et G . ■

Remarquons que l'application nulle et l'application identique de E sont deux projecteurs particuliers.

Supplémentaires d'un sous-espace

Etant donné un sous- K -ev F d'un K -ev E on appelle **supplémentaire de F** tout sous- K -ev G de E tel que $E = F \oplus G$. Nous admettrons ici que F admet toujours *au moins un* supplémentaire (ce théorème sera prouvé plus loin dans les cas particuliers où E est de dimension finie (corollaire du théorème IX.4.3) ou encore lorsque F est de codimension finie, mais une démonstration générale fait appel à l'axiome du choix). Le seul supplémentaire de E (resp. $\{0_E\}$) est $\{0_E\}$ (resp. E), mais en dehors de ces cas triviaux, il n'y a jamais unicité d'un supplémentaire. C

différents supplémentaires de F donné présentent la particularité d'être tous isomorphes entre eux :

THÉOREME IX.1.3

|| Soit G_1 et G_2 deux supplémentaires, dans un K -ev E , d'un sous- K -ev donné F . Alors les sous- K -ev G_1 et G_2 sont **isomorphes**.

Démonstration :

Soit p_1 la projection sur G_1 parallèlement à F . Puisque $\text{Im}(p_1) = G_1$, on peut définir l'application linéaire $f = p_1|_{G_2} : G_2 \rightarrow G_1$. On a : $\text{Ker}(f) = \text{Ker}(p_1) \cap G_2 = F \cap G_2 = \{0_E\}$, donc f est injective. De plus, soit $y \in G_1$. Ce vecteur se décompose en $y = x + y_2$, avec $x \in F$ et $y_2 \in G_2$, d'où $p_1(y) = y = p_1(x) + p_1(y_2) = p_1(y_2)$. Donc $y = f(y_2)$ et f est surjective. Ainsi f est un isomorphisme de G_2 sur G_1 . ■

Supplémentaires du noyau d'une application linéaire

On peut généraliser le théorème IX.1.3 de la manière suivante :

THÉOREME IX.1.4

|| Soit E et F deux K -ev, et $u : E \rightarrow F$ une application linéaire, de noyau N et d'image I . Pour tout supplémentaire S de N dans E , l'application linéaire $f = u|_S : S \rightarrow I$ est un isomorphisme de K -ev.

Démonstration :

On a $\text{Ker}(f) = \text{Ker}(u) \cap S = N \cap S = \{0_E\}$, d'où f est injective. Soit $y \in I$; écrivons $y = u(x)$, où $x \in E$ et décomposons x sous la forme $x = z + t$, $z \in N$, $t \in S$. Alors $u(x) = u(z) + u(t) = u(t) = f(t)$, donc f est surjective. ■

On retrouve le fait que tous les supplémentaires de N sont isomorphes entre eux, puisqu'ils sont tous isomorphes à I .

Remarque 1 : Dans le cas où $E = F$, si u est un endomorphisme de E , sous prétexte que l'image $\text{Im}(u)$ est *isomorphe* à un supplémentaire du noyau N , il ne faudrait pas en conclure que l'image est elle-même supplémentaire de N , le cas où u est un projecteur étant très spécial. Pour un endomorphisme quelconque, on a généralement $I \cap N \neq \{0_E\}$.

Exercice 1 : Soit E un K -ev et $u \in \text{Hom}_K(E)$. Pour que $E = \text{Ker}(u) \oplus \text{Im}(u)$, il faut et il suffit qu'il existe deux sous- K -ev F et G de E tels que :

$$E = F \oplus G, \quad u(F) \subset F, \quad u(G) \subset G, \quad u|_G = 0 \quad \text{et} \quad u|_F \in \text{GL}$$

Exercice 2 : (On admet que dans tout K -ev, tout sous- K -ev a au moins un supplémentaire).

a) Soit E un K -ev et $u \in \text{Hom}_K(E)$. Montrer qu'il existe un projecteur p de E et un automorphisme α de E tels que $u = \alpha \circ p$. *Indication :* Soit S et J deux sous- K -ev de E tels que $E = \text{Ker}(u) \oplus S = J \oplus \text{Im}(u)$. Prendre $p =$ projection sur S parallèlement à $\text{Ker}(u)$.

b) Prouver de même qu'il existe un projecteur q de E et un automorphisme β de E tels que $u = q \circ \beta$.

Exercice 3 : Soit E un K -ev. Chercher tous les couples (p, q) de projecteurs de E tels que $pq = qp$.

Exercice 4 : On donne $n \in \mathbb{N}^*$. Soit E_n le \mathbb{C} -ev des fonctions polynomiales homogènes de degré n sur \mathbb{C}^2 (cf. § X.1). On note F le sous- \mathbb{C} -ev des polynômes *harmoniques* dans E_n (i.e. les $P \in E_n$ tels que $\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} = 0$) et G le sous- \mathbb{C} -ev des polynômes $P \in E_n$ de la forme $(x^2 + y^2)Q$, avec $Q \in E_{n-2}$. Montrer que $E_n = F \oplus G$.

Exercice 5 : Soit F un sous- K -ev d'un K -ev E tel que $F \neq \{0_E\}$ et $F \neq E$.

Démontrer la non-unicité d'un supplémentaire de F dans E .

Exercice 6 : A quelle condition la somme de deux projecteurs est-elle un projecteur ? Si c'est le cas, montrer que $\text{Im}(f) \cap \text{Im}(g) = \{0\}$ et $\text{Ker}(f) \cap \text{Ker}(g) = \text{Ker}(f + g)$.

Exercice 7 : Soit p un projecteur dans E et $q = \text{Id}_E - p$. On pose $F = \text{Im}(p)$ et $G = \text{Im}(q)$.

Montrer que pour qu'un endomorphisme u de E commute avec p , il faut et il suffit que $u(F) \subset F$ et $u(G) \subset G$.

Exercice 8 : Montrer que pour que deux endomorphismes u et v du K -ev E vérifient : $u \circ v = u$ et $v \circ u = v$, il faut et il suffit que ce soient deux projecteurs de même noyau.

Exercice 9 : Soit p et q deux projecteurs dans E . Montrer que la condition $p \circ q = q$ est nécessaire et suffisante pour que $\text{Im}(q) \subset \text{Im}(p)$.

Exercice 10 : (On suppose connue la notion d'intégrale définie).

Soit E le \mathbb{R} -ev des fonctions de classe C^∞ et 2π -périodiques $\mathbb{R} \rightarrow \mathbb{R}$, et D l'endomorphisme qui, à $f \in E$, associe sa dérivée seconde f'' . Montrer que $E = \text{Ker}(D) \oplus \text{Im}(D)$.

§ IX.2 PRODUITS ET SOMMES D'ESPACES VECTORIELS

Dans ce § IX.2, le corps commutatif de base K est fixé.

Produit

Soit $(V_i)_{i \in I}$ une famille non vide de K -ev. Au § V.1, nous avons vu qu'à partir des structures de groupe additif des V_i , on définit sur l'ensemble produit $V = \prod_{i \in I} V_i$, une structure de *groupe additif produit* : si $x = (x_i)_{i \in I}$ et $y = (y_i)_{i \in I}$ sont dans V , on a :

$$(1) \quad x + y = (x_i + y_i)_{i \in I}.$$

On définit une *loi externe de domaine* K sur ce groupe en posant, pour $\lambda \in K$ et pour $x = (x_i)_{i \in I}$ élément de V :

$$(2) \quad \lambda x = (\lambda x_i)_{i \in I}.$$

Des vérifications évidentes montrent qu'on a ainsi muni V d'une structure d'espace vectoriel.

DÉFINITION IX.2.1

Si $(V_i)_{i \in I}$ est une famille non vide de K -ev, on appelle **K -espace vectoriel produit** des V_i , et on note $\prod_{i \in I} V_i$, le K -ev construit sur l'ensemble produit des V_i à l'aide de l'addition et de la multiplication par les scalaires définies par (1) et (2).

Pour chaque $i \in I$, l'application $p_i : V \rightarrow V_i$ (où $V = \prod_{i \in I} V_i$) est K -linéaire, surjective, et son noyau est évidemment $\prod_{j \in I, j \neq i} V_j$, avec $V_i' = \{0_{V_i}\}$ et $V_j' = V_j$ si $j \neq i$. Il est clair que ce noyau N_i est canoniquement isomorphe à $\prod_{j \neq i} V_j$. Les p_i sont appelées les **projections naturelles** de V sur les V_i . De plus, pour tout K -ev E , et toute $f \in \text{Hom}_K(E, V)$, $p_i \circ f \in \text{Hom}_K(E, V_i)$ pour chaque $i \in I$. Réciproquement, si on donne, pour chaque $i \in I$, l'application $f_i \in \text{Hom}_K(E, V_i)$, alors l'application $f : E \rightarrow V$, $x \mapsto (f_i(x))_{i \in I}$ est K -linéaire, et vérifie $p_i \circ f = f_i$ pour tout $i \in I$.

Ainsi on voit qu'on obtient une *bijection* de $\text{Hom}_K(E, V)$ sur l'ensemble produit des $\text{Hom}_K(E, V_i)$ en associant, à chaque $f \in \text{Hom}_K(E, V)$, la famille $(p_i \circ f)_{i \in I}$. De plus, chaque application

$$\text{Hom}_K(E, V) \rightarrow \text{Hom}_K(E, V_i),$$

$f \mapsto p_i \circ f$ est K -linéaire (cf. propriétés (1) et (2) du § VI.2). Tenant compte de la structure de K -ev produit de $\prod_{i \in I} \text{Hom}_K(E, V_i)$, nous avons finalement :

THÉORÈME IX.2.1

Soit $(V_i)_{i \in I}$ une famille non vide de K -ev et $V = \prod_{i \in I} V_i$. Pour tout K -ev E , l'application

$$\text{Hom}_K(E, V) \rightarrow \prod_{i \in I} \text{Hom}_K(E, V_i), \quad f \mapsto (p_i \circ f)_{i \in I}$$

(où les $p_i : V \rightarrow V_i$ sont les projections naturelles) est un **isomorphisme de K -ev**.

Somme directe externe

Reprenons le K -ev produit $V = \prod_{i \in I} V_i$ de la famille $(V_i)_{i \in I}$ de K -ev.

Pour chaque $i \in I$, on a une *injection canonique* $J_i : V_i \rightarrow V$, $x \mapsto (x_j)_{j \in I}$, avec $x_i = x$ et $x_j = 0$ si $j \neq i$, qui est évidemment K -linéaire. Il est clair d'après leur définition que : $J_i \circ p_j = 0$ si $j \neq i$, $p_j \circ J_i = 0$ si $j \neq i$, $p_i \circ J_i = \text{Id}_{V_i}$. De plus, soit $q_i = J_i \circ p_i$. Alors q_i est le *projecteur* de V sur V_i parallèlement à $\text{Ker}(q_i) \cong \prod_{j \neq i} V_j$.

Le sous-espace vectoriel de V engendré par la réunion de tous les $(J_i(V_i))_{i \in I}$ est évidemment l'ensemble des $x = (x_i)_{i \in I} \in V$ tels que la famille $(x_i)_{i \in I}$ soit à *support fini*. On pose :

DÉFINITION IX.2.2

Soit V le K -ev produit de la famille $(V_i)_{i \in I}$ de K -ev. On appelle **somme directe externe** des V_i , et on note $\coprod_{i \in I} V_i$, le sous- K -ev de V formé des familles $(x_i)_{i \in I}$ à **support fini**.

Soit maintenant un K -ev quelconque F . Avec les notations ci-dessus, posons $W = \coprod_{i \in I} V_i$. Pour chaque $i \in I$, notons $\psi_i = J_i|_W$. Pour chaque

application linéaire $f : W \rightarrow F$, on a : $f \circ \psi_i \in \text{Hom}_K(V_i, F)$ ($i \in I$) ; et si i est fixé, $f \mapsto f \circ \psi_i$ est K -linéaire (cf. propriétés 1 et 2, § VI.2).

Inversement, donnons-nous, pour chaque $i \in I$, $v_i \in \text{Hom}_K(V_i, F)$. Pour $x = (x_i)_{i \in I}$ élément de W , on a : $x = \sum_{j \in I} \psi_j(x_j)$. On peut donc définir

$f(x) = \sum_{j \in I} v_j(x_j)$, et on a, pour tous $i \in I$ et $x_i \in V_i$: $f \circ \psi_i(x_i) = v_i(x_i)$,

d'où $f \circ \psi_i = v_i$. Finalement on a établi :

THÉORÈME IX.2.2

Soit $(V_i)_{i \in I}$ une famille non vide de K -ev, W leur somme directe externe, et pour chaque $i \in I$, $\psi_i : V_i \rightarrow W$ l'injection canonique. Pour tout K -ev F , l'application

$$\text{Hom}_K(W, F) \rightarrow \prod_{i \in I} \text{Hom}_K(V_i, F), f \mapsto (f \circ \psi_i)_{i \in I}$$

est un isomorphisme de K -ev.

Cas d'une famille finie

Supposons I fini (non vide) dans ce qui précède. Alors les espaces $V = \prod_{i \in I} V_i$ et $W = \coprod_{i \in I} V_i$ sont égaux. En gardant les notations antérieures,

les projecteurs $q_i = \psi_i \circ p_i$ vérifient $\sum_{i \in I} q_i = \text{Id}_V$, et $q_i \circ q_j = 0$ si $i \neq j$.

Si $I = \llbracket 1, n \rrbracket$ ($n \in \mathbb{N}^*$), l'espace V se note aussi $V_1 \times V_2 \times \cdots \times V_n$ ou en abrégé $\bigtimes_{i=1}^n V_i$. La propriété la plus remarquable de ce produit fini de K -

ev est la validité simultanée des théorèmes IX.2.1 et IX.2.2.

Somme interne de sous-espaces vectoriels

Partons maintenant d'une famille quelconque $(E_i)_{i \in I}$ de sous- K -ev d'un K -ev E . On sait que la réunion des E_i engendre un sous- K -ev S de E qui n'est autre que l'ensemble des sommes $\sum_{i \in I} x_i$, où $(x_i)_{i \in I}$ est une famille arbitraire à support fini de vecteurs de E telle que $x_i \in E_i$ pour tout i .

DÉFINITION IX.2.3

⎵ Dans les conditions ci-dessus, le sous- K -ev de E engendré par les
⎵ E_i s'appelle **somme interne** des sous- K -ev E_i , et se note : $\sum_{i \in I} E_i$.

Bien sûr, lorsque I est finie, la somme interne $\sum_{i \in I} E_i$ se confond avec l'ensemble de toutes les sommes $\sum_{i \in I} x_i$, avec $x_i \in E_i$ pour tout i .

Si $I = \llbracket 1, n \rrbracket$ ($n \in \mathbb{N}^*$) on note ce sous- K -espace vectoriel

$$E_1 + E_2 + \cdots + E_n \quad \text{ou} \quad \sum_{i=1}^n E_i.$$

En particulier la somme de deux sous- K -ev de E permet de définir sur l'ensemble des sous- K -ev de E une loi de composition interne : $(E_1, E_2) \mapsto E_1 + E_2$ qui est évidemment **commutative** et **associative**, et telle que, pour toute famille finie $(E_i)_{i \in I}$ de sous- K -ev de E , la somme interne $\sum_{i \in I} E_i$ n'est autre que le composé des E_i pour cette loi.

Dans le cas général où I est quelconque, on peut former la somme directe externe $F = \coprod_{i \in I} E_i$, et définir $s : F \rightarrow E$, $(x_i)_{i \in I} \mapsto \sum_{i \in I} x_i$ (s est l'unique application K -linéaire de F dans E telle que $s \circ \psi_i = \zeta_i$ pour tout i , où $\psi_i : E_i \rightarrow F$ et $\zeta_i : E_i \rightarrow E$ sont les injections canoniques).

Ainsi, s est K -linéaire, et la définition même du sous- K -ev $\sum_{i \in I} E_i$ montre que : $\text{Im}(s) = \sum_{i \in I} E_i$. Cette application s sera appelée l'**application linéaire canonique** de $\coprod_{i \in I} E_i$ dans E .

Exemple 1 : Soit $(e_i)_{i \in I}$ une famille quelconque de vecteurs du K -ev E . Pour chaque $i \in I$, soit $E_i = Ke_i$ le sous- K -ev engendré par e_i . Alors la somme interne $\sum_{i \in I} E_i$ est exactement le sous- K -ev engendré par les $(e_i)_{i \in I}$.

THÉORÈME IX.2.3

Soit $(E_i)_{i \in I}$ une famille de sous- K -ev du K -ev E . Les propriétés suivantes sont équivalentes :

(I) L'application linéaire canonique $s : \coprod_{i \in I} E_i \rightarrow E$ est **injective**.

(II) Pour tout $i \in I$, $E_i \cap \sum_{j \neq i} E_j = \{0_E\}$.

Démonstration :

$$(I) \Rightarrow (II).$$

Soit $i \in I$, et $x \in E_i \cap \sum_{j \neq i} E_j$; x peut s'écrire $x = \sum_{j \neq i} x_j$ (où $x_j \in E_j$, et où la famille (x_j) est à support fini). Soit y l'élément $(y_k)_{k \in I}$ de $\coprod_{k \in I} E_k$ tel que $y_i = x$ et $y_j = -x_j$ si $j \neq i$. Alors, $s(y) = 0$, donc $y = 0$, d'où $x = 0$.

$$(II) \Rightarrow (I).$$

Soit $x = (x_i)_{i \in I}$ élément de $\coprod_{i \in I} E_i$ tel que $s(x) = 0$, c'est-à-dire $\sum_{i \in I} x_i = 0$. Fixons $i \in I$; on a : $x_i = \sum_{j \in I, j \neq i} (-x_j)$, d'où $x_i = 0$. C'est vrai pour tout $i \in I$, d'où $x = 0$, et s est injective. ■

DÉFINITION IX.2.4

Soit $(E_i)_{i \in I}$ une famille de sous- K -ev du K -ev E . On dit que les (E_i) sont **linéairement indépendants** ssi les conditions équivalentes du théorème IX.2.3 sont satisfaites. S'il en est bien ainsi, la somme interne $\sum_{i \in I} E_i$ est dite **directe** et se note $\oplus_{i \in I} E_i$.

Une autre façon d'exprimer la condition (I) est la suivante : les $(E_i)_{i \in I}$ sont indépendants ssi tout $x \in \sum_{i \in I} E_i$ s'écrit **d'une manière unique** sous la forme $x = \sum_{i \in I} x_i$, avec $x_i \in E_i$ pour tout i , et les (x_i) à support fini.

Exemple 2 : Soit $(e_i)_{i \in I}$ une famille de vecteurs *non nuls* du K -ev E . Pour chaque $i \in I$, soit $E_i = Ke_i = \{\lambda e_i\}_{\lambda \in K}$ le sous- K -ev engendré par e_i . Alors les (E_i) sont **indépendants** ssi la famille (e_i) est **libre**. Et $(e_i)_{i \in I}$ est une **base** de E ssi on a : $E = \oplus_{i \in I} E_i$.

Exemple 3 : Soit $(e_i)_{i \in I}$ une famille libre de vecteurs de E , et $S = \text{Vect}((e_i)_{i \in I})$. Considérons un *partage* $(J_\lambda)_{\lambda \in L}$ de I , et pour chaque $\lambda \in L$, soit $V_\lambda = \text{Vect}(e_i)_{i \in J_\lambda}$. Alors $S = \bigoplus_{\lambda \in L} V_\lambda$.

Sommes directes internes finies

Soit une *suite finie* : E_1, E_2, \dots, E_n ($n \in \mathbb{N}^*$) de sous- K -ev de E . Dans ce cas, l'application linéaire canonique $s : E_1 \times E_2 \times \dots \times E_n \rightarrow E$ est tout simplement $(x_1, x_2, \dots, x_n) \mapsto x_1 + x_2 + \dots + x_n = \sum_{i=1}^n x_i$, et ce qui précède entraîne les résultats fondamentaux suivants :

1) Les $(E_i)_{1 \leq i \leq n}$ sont **indépendants** ssi s est **injective**, c'est-à-dire si tout $x \in E_1 + E_2 + \dots + E_n$ s'écrit **de manière unique** $x = x_1 + x_2 + \dots + x_n$ avec $x_i \in E_i$ pour tout i . Si c'est le cas, on note aussi :

$$\bigoplus_{i=1}^n E_i = E_1 \oplus E_2 \oplus \dots \oplus E_n.$$

2) Les $(E_i)_{1 \leq i \leq n}$ sont indépendants ssi $E_i \cap \left(\sum_{j \neq i} E_j \right) = \{0\}$ pour tout $i \in \llbracket 1, n \rrbracket$.

3) On a $E = \bigoplus_{i=1}^n E_i$ ssi $s : E_1 \times E_2 \times \dots \times E_n \rightarrow E$ est un **isomorphisme** de K -ev.

Il convient de remarquer qu'apparemment la condition du 2) demande n vérifications, mais pour $n = 2$ il est clair qu'une seule vérification suffit. Toujours dans le cas $n = 2$, s est un **isomorphisme** ssi E_1 et E_2 sont deux sous- K -ev **supplémentaires** de E . Remarquons également que la notation $E = E_1 \oplus E_2$ introduite ci-dessus est bien en accord avec celle du § IX.1.

Etudions enfin les endomorphismes attachés à une décomposition en somme directe interne finie. Supposons que $E = \bigoplus_{i=1}^n E_i$ ($n \in \mathbb{N}^*$). Notons

$s : P = \prod_{i=1}^n E_i \rightarrow E$ l'isomorphisme canonique, $\zeta_i : E_i \rightarrow E$ l'injection canonique de E_i dans E et π_i le projecteur de E sur E_i parallèlement à

$F_i = \sum_{j \neq i} E_j$, $\varpi_i = \pi_i|^{E_i}$, $\psi_i : E_i \rightarrow P$ l'injection canonique, $p_i : P \rightarrow E_i$ les

projections naturelles, et $q_i = \psi_i \circ p_i$. On voit immédiatement : $\zeta_i = s \circ \psi_i$, $\pi_i = s \circ q_i \circ s^{-1}$, $\sum \pi_i = \text{Id}_E$, $\pi_i \circ \pi_j = 0$ si $i \neq j$. Les π_i sont appelés

les *projecteurs associés à la décomposition* $E = \bigoplus_{i=1}^n E_i$.

Tenant compte de l'isomorphisme s , les théorèmes IX.2.1 et IX.2.2 entraînent :

• Pour tout K -ev F , l'application $\text{Hom}_K(F, E) \rightarrow \bigtimes_{i=1}^n \text{Hom}_K(F, E_i)$,
 $f \mapsto (\omega_i \circ f)_{1 \leq i \leq n}$ est un isomorphisme de K -ev.

• Pour tout K -ev G , l'application $\text{Hom}_K(E, G) \rightarrow \bigtimes_{i=1}^n \text{Hom}_K(E_i, G)$,
 $f \mapsto (f \circ \zeta_i)_{1 \leq i \leq n}$ est un isomorphisme de K -ev.

Une conséquence essentielle de ces résultats est la suivante :

THÉORÈME IX.2.4

Soit E un K -ev et E_1, E_2, \dots, E_n des sous- K -ev de E ($n \in \mathbb{N}^*$) tels que $E = \bigoplus_{i=1}^n E_i$. Pour chaque suite $(u_i)_{1 \leq i \leq n}$, où $u_i \in \text{Hom}_K(E_i)$, il existe un unique $u \in \text{Hom}_K(E)$ tel que $(\forall i \in \llbracket 1, n \rrbracket) u|_{E_i} = u_i$.
 L'application $\Phi : \bigtimes_{i=1}^n \text{Hom}_K(E_i) \rightarrow \text{Hom}_K(E)$ qui associe à chaque suite (u_i) cet endomorphisme u , est K -linéaire et injective, et vérifie

$$\Phi((u_i \circ v_i)_{1 \leq i \leq n}) = \Phi((u_i)) \circ \Phi((v_i))$$

pour tous u_i et v_i dans $\text{Hom}_K(E_i)$. L'image de Φ est l'ensemble des endomorphismes u de E tels que $u(E_i) \subset E_i$ pour tout i .

Démonstration (abrégée) :

Soit $\zeta_i : E_i \rightarrow E$ l'injection canonique ($1 \leq i \leq n$). Donnons-nous, pour chaque $i \in \llbracket 1, n \rrbracket$, $u_i \in \text{Hom}_K(E_i)$ et posons $v_i = \zeta_i \circ u_i$. Alors $v_i \in \text{Hom}_K(E_i, E)$. On a donc un unique endomorphisme u de E tel que $u \circ \zeta_i = v_i$ pour tout i . Cet endomorphisme vérifie $u(E_i) \subset E_i$ et $u|_{E_i} = u_i$ pour tout i . Réciproquement si $w \in \text{Hom}_K(E)$ laisse stable chaque E_i et vérifie $w|_{E_i} = u_i$ pour tout i , on a : $w \circ \zeta_i = v_i$ pour tout i , donc $w = u$. On a donc prouvé l'existence de l'application Φ , son injectivité et le fait que son image est bien l'ensemble des $u \in \text{Hom}_K(E)$ laissant stable chaque E_i . Enfin, si $u \in \text{Im}(\Phi)$ et $v \in \text{Im}(\Phi)$, pour tout i on a :

$$(u \circ v)|_{E_i} = (u|_{E_i}) \circ (v|_{E_i}) \quad \text{d'où la dernière assertion.} \quad \blacksquare$$

Avec les notations de ce théorème, si $(u_i)_{1 \leq i \leq n}$ est élément de $\bigtimes_{i=1}^n \text{Hom}_K(E_i)$, l'endomorphisme $\Phi(u)$ sera noté $\bigoplus_{i=1}^n u_i$, ou encore $u_1 \oplus u_2 \oplus \dots \oplus u_n$. Par définition, on a donc,

$$\text{pour } (x_i)_{1 \leq i \leq n} \in E_1 \times E_2 \times \dots \times E_n, \quad \left(\bigoplus_{i=1}^n u_i \right) \left(\sum_{i=1}^n x_i \right) = \sum_{i=1}^n u_i(x_i)$$

Exercice 1 : Soit $(E_i)_{i \in I}$ une famille de sous- K -ev d'un K -ev E . Pour chaque $i \in I$, soit $(V_{i,\lambda})_{\lambda \in L_i}$ une famille de sous- K -ev de E_i . Pour que la famille $(V_{i,\lambda})_{i \in I, \lambda \in L_i}$ soit indépendante, il faut et il suffit que la famille $(E_i)_{i \in I}$ le soit et que chaque famille $(V_{i,\lambda})_{\lambda \in L_i}$ le soit. S'il en est ainsi, on a :

$$\bigoplus_{i,\lambda} V_{i,\lambda} = \bigoplus_{i \in I} \left(\bigoplus_{\lambda \in L_i} V_{i,\lambda} \right).$$

Exercice 2 : Soit E_1, E_2, \dots, E_n des sous- K -ev de E . Pour que les E_i soient indépendants, il faut et il suffit que

$$\forall i \in \llbracket 2, n \rrbracket, \quad E_i \cap \left(\sum_{j=1}^{i-1} E_j \right) = \{0\}.$$

Généraliser.

Exercice 3 : Soit $(E_i)_{i \in I}$ une famille de sous- K -ev indépendants d'un K -ev E .

a) Pour tout $i \in I$, soit F_i un sous- K -ev de E_i . Alors les $(F_i)_{i \in I}$ sont indépendants.

b) Soit F un sous- K -ev de E . Comparer $\bigoplus_{i \in I} (E_i \cap F)$ et $\left(\bigoplus_{i \in I} E_i \right) \cap F$. Donner une CNS

pour que ces espaces soient égaux.

Exercice 4 : Soit E, F, G trois sous- K -ev d'un K -ev E . On suppose $F \subset G$, $E \cap F = E \cap G$ et $E + F = E + G$. Montrer que $F = G$.

Exercice 5 : Soit E_1, E_2, \dots, E_n des sous- K -ev de E tels que $\bigoplus_{i=1}^n E_i = E$. On donne

$u_i \in \text{Hom}_K(E_i)$ pour $1 \leq i \leq n$; soit $u = \bigoplus_{i=1}^n u_i$.

a) Montrer : $\text{Ker}(u) = \bigoplus_{i=1}^n \text{Ker}(u_i)$, $\text{Im}(u) = \bigoplus_{i=1}^n \text{Im}(u_i)$.

b) On suppose chaque $E_i \neq \{0\}$. Montrer que les automorphismes α de E tels que $\alpha(E_i) \subset E_i$ pour tout i forment un sous-groupe Γ de $\text{GL}_K(E)$. Montrer que le groupe Γ est naturellement isomorphe au groupe-produit $\bigtimes_{i=1}^n \text{GL}_K(E_i)$.

Exercice 6 : Soit E, F, G trois sous- K -cv d'un K -cv. Montrer que :

$$E \cap (F + (E \cap G)) = (E \cap F) + (E \cap G).$$

Exercice 7 : Soit E un K -ev et p_1, p_2, \dots, p_n des projecteurs deux à deux permutables de E tels que $p_1 + p_2 + \dots + p_n = \text{Id}_E$. On pose $E_i = \text{Im}(p_i)$. Montrer que $E = \bigoplus_{i=1}^n E_i$ et que les p_i sont les projecteurs associés à cette décomposition de E .

Exercice 8 : Soit \mathcal{V} le \mathbb{R} -ev des fonctions $f: \mathbb{R} \rightarrow \mathbb{R}$ continues, nulles sur \mathbb{Z} et nulles en dehors d'un intervalle borné (dépendant de f). Pour chaque $n \in \mathbb{Z}$, soit E_n le sous- \mathbb{R} -ev de \mathcal{V} formé des $f: \mathbb{R} \rightarrow \mathbb{R}$ continues et nulle hors de $[n, n+1]$. Montrer que $\mathcal{V} = \bigoplus_{n \in \mathbb{Z}} E_n$.

Exercice 9 : Soit E un K -ev, et F, G, H, F', G', H' des sous- K -ev de E .

a) Comparer : $F + (G \cap H)$ et $(F + G) \cap (F + H)$; $F \cap (G + H)$ et $(F \cap G) + (F \cap H)$.

b) Montrer : si $F \cap G = F' \cap G'$, alors $F = [F + (G \cap F')] \cap [F + (G \cap G')]$.

c) Soit F'', G'' des supplémentaires de $F \cap G$ dans F et G . On suppose $E = F + G$. Comparer E et $F' \oplus G'$, E et $F' \oplus G$, puis E et $F' \oplus G' \oplus (F \cap G)$.

Exercice 10 : Soit E un K -ev et deux projecteurs p et q tels que $p \circ q = 0$. Montrer que $r = p + q - q \circ p$ est un projecteur tel que $\text{Ker}(r) = \text{Ker}(p) \cap \text{Ker}(q)$, et $\text{Im}(r) = \text{Im}(p) \oplus \text{Im}(q)$.

§ IX.3 ESPACES DE DIMENSION FINIE

Caractérisation des bases d'un espace vectoriel

Soit x un élément d'un K -ev E . Si $x = 0_E$, la famille (x) , bien que ne contenant qu'un seul élément, n'est pas libre (car $1_K \cdot x = 0_E$ et $1_K \neq 0_K$). En revanche si $x \neq 0_E$, la famille à un seul élément (x) est libre (car $\lambda x = 0_E \Rightarrow \lambda = 0$, sinon λ aurait un inverse λ^{-1} et

$$\lambda^{-1}(\lambda x) = 1 \cdot x = x = \lambda^{-1} 0_E = 0_E$$

contrairement à l'hypothèse).

PROPOSITION IX.3.1

Soit $(e_i)_{i \in I}$ une famille d'éléments d'un K -ev E . Cette famille est libre ssi

(I) $(\forall i \in I) \quad e_i \notin \text{Vect}((e_j)_{j \in I, j \neq i})$.

Démonstration :

Il est bien évident que si l'on avait $e_i \in \text{Vect}((e_j)_{j \neq i})$ c'est-à-dire $e_i = \sum_{j \in I, j \neq i} \lambda_j e_j$, avec des $\lambda_j \in K$ à support fini, on en déduirait la relation de dépendance linéaire $\sum_{k \in I} \rho_k e_k = 0$ avec $\rho_i = -1_K \neq 0_K$ et $\rho_k = \lambda_k$ si $k \neq i$, ce qui rend impossible l'indépendance des (e_k) .

Réciproquement, supposons (I) satisfaite. Imaginons qu'il existe une famille de $(\lambda_i)_{i \in I}$ non tous nuls et à support fini tels que $\sum_{i \in I} \lambda_i e_i = 0$. A partir d'un $\lambda_i \neq 0$, on écrirait $e_i = \sum_{j \in I, j \neq i} (-\lambda_i^{-1} \lambda_j) e_j$, et donc $e_i \in \text{Vect}((e_j)_{j \neq i})$, contrairement à (I). ■

THÉORÈME IX.3.1

Soit \mathcal{B} une partie non vide d'un K -ev E . Il y a équivalence entre les propriétés suivantes :

- (I) \mathcal{B} est une base de E .
- (II) \mathcal{B} engendre le K -ev E , et aucune partie stricte de \mathcal{B} n'engendre le K -ev E .
- (III) \mathcal{B} est libre, et aucune partie de E contenant \mathcal{B} strictement n'est libre.

Démonstration :

(I) \Rightarrow (II). En effet, si \mathcal{B} est une base, \mathcal{A}

K -ev E . D'autre part, puisque \mathcal{B} est libre, si $b \in \mathcal{B}$ on a $b \notin \text{Vect}(\mathcal{B} \setminus \{b\})$ d'après la proposition IX.3.1, donc $\mathcal{B} \setminus \{b\}$ n'engendre plus le K -ev E .

(II) \Rightarrow (III). \mathcal{B} engendre E . Si $b \in \mathcal{B}$, on a $b \notin \text{Vect}(\mathcal{B} \setminus \{b\})$ à cause de l'hypothèse, donc \mathcal{B} est libre (proposition IX.3.1). Soit alors $c \in E \setminus \mathcal{B}$. Par hypothèse $c \in \text{Vect}(\mathcal{B})$, donc $\mathcal{B} \cup \{c\}$ n'est plus libre (proposition IX.3.1), d'où (III).

(III) \Rightarrow (I). Soit x un vecteur quelconque de E . Si $x \in \mathcal{B}$, alors $x \in \text{Vect}(\mathcal{B})$. Si $x \notin \mathcal{B}$, par hypothèse $\mathcal{B} \cup \{x\}$ est une partie liée de E . Il existe donc une relation de dépendance linéaire

$$\lambda x + \sum_{b \in \mathcal{B}} \lambda_b b = 0_E.$$

Nécessairement $\lambda \neq 0$, sinon \mathcal{B} serait liée contrairement à l'hypothèse, d'où $x = \sum_{b \in \mathcal{B}} (-\lambda^{-1} \lambda_b) b \in \text{Vect}(\mathcal{B})$.

En conséquence la partie libre \mathcal{B} engendre E . ■

Espaces de dimension finie

DÉFINITION IX.3.1

Un K -ev E est dit **de dimension finie** ssi il admet au moins une **famille génératrice finie**, ou ce qui revient au même, au moins une **partie génératrice finie**. Un K -ev est dit **de dimension infinie** ssi il n'est pas de dimension finie.

PROPOSITION IX.3.2

|| Soit E, F deux K -ev et $u : E \rightarrow F$ une application linéaire. Si E est de dimension finie, $\text{Im}(u)$ est de dimension finie.

Démonstration :

Si $(a_i)_{i \in I}$ est une famille génératrice finie du K -ev E , les définitions du § VI.3 montrent que $(u(a_i))_{i \in I}$ est une famille génératrice du K -ev $\text{Im}(u)$ qui est donc de dimension finie. ■

En particulier, si $u : E \rightarrow F$ est un **isomorphisme**, on voit que E est de dimension finie ssi F est de dimension finie.

PROPOSITION IX.3.3

|| Si E et F sont des K -ev de dimension finie, $E \times F$ l'est aussi.

Démonstration :

Si A (resp. B) est une partie génératrice finie de E (resp. F), alors $(A \times \{0_F\}) \cup (\{0_E\} \times B)$ est une partie génératrice finie de $E \times F$. ■

Par récurrence, on en déduit que *tout produit fini de K -ev de dimension finie est un K -ev de dimension finie*. Compte tenu des résultats du § IX.2, il s'ensuit que si $E = \bigoplus_{i=1}^n E_i$ et si chaque E_i est de dimension finie, il en est de même de E .

PROPOSITION IX.3.4

|| Soit E un K -ev **de dimension finie**, et soit A une partie génératrice de E . Il existe une **partie finie** G de A qui engendre le K -ev E .

Démonstration :

On sait par hypothèse que E possède une partie génératrice finie Γ . Pour chaque $x \in \Gamma$, soit S_x une partie finie de A telle que $x = \sum_{a \in S_x} \lambda_{a,x} \cdot a$ avec des $\lambda_{a,x}$ dans K . L'ensemble $G = \bigcup_{x \in \Gamma} S_x$ est fini, et on a : $\Gamma \subset \text{Vect}(G)$, donc $E = \text{Vect}(\Gamma) \subset \text{Vect}(G)$, donc G engendre le K -ev E . ■

COROLLAIRE

|| Dans un K -ev de dimension finie E , toute base est finie.

Démonstration :

Soit \mathcal{B} une partie-base de E et G une partie génératrice finie de E incluse dans \mathcal{B} (proposition IX.3.4). Alors $G = \mathcal{B}$ d'après le théorème IX.3.1, donc \mathcal{B} est finie. ■

THÉOREME IX.3.2

|| Soit E un K -ev de dimension finie, \mathcal{L} une partie libre de E et \mathcal{G} une partie génératrice de E . Il existe une partie-base \mathcal{B} de E telle que $\mathcal{L} \subset \mathcal{B} \subset \mathcal{L} \cup \mathcal{G}$.

Démonstration :

Pour $E = \{0_E\}$ il n'y a rien à prouver. Supposons donc $E \neq \{0_E\}$. Soit G une partie génératrice de E incluse dans \mathcal{G} (cf. proposition IX.3.4). Nécessairement $G \neq \emptyset$ car $E \neq \{0_E\}$. Soit

$$p = \text{card}(G) \quad (p \in \mathbb{N}^*).$$

Notons \mathcal{E} l'ensemble des parties \mathcal{M} de G telles que $\mathcal{L} \cup \mathcal{M}$ soit une partie libre de E . \mathcal{E} n'est pas vide ($\emptyset \in \mathcal{E}$). L'entier $q = \text{Max}\{\text{card}(\mathcal{M}), \mathcal{M} \in \mathcal{E}\}$ est donc bien défini. Choisissons alors $\mathcal{M}_0 \in \mathcal{E}$ tel que $\text{card}(\mathcal{M}_0) = q$. La partie $\mathcal{B} = \mathcal{L} \cup \mathcal{M}_0$ de E est libre. Il reste à prouver qu'elle engendre E . Pour cela il suffit, bien sûr, de voir que $G \subset \text{Vect}(\mathcal{B})$. Soit donc $g \in G$. Si $g \in \mathcal{M}_0$, alors $g \in \text{Vect}(\mathcal{B})$. Si $g \notin \mathcal{M}_0$, alors $\mathcal{M}_0 \cup \{g\} \subset G$ et $\mathcal{M}_0 \cup \{g\} \notin \mathcal{E}$ puisque so

$q + 1$; donc $\mathcal{B} \cup \{g\}$ est liée ; écrivant une relation de dépendance linéaire et tenant compte que \mathcal{B} est libre, on en déduit que $g \in \text{Vect}(\mathcal{B})$, par « division » par le coefficient de g . ■

COROLLAIRE 1

|| Dans un K -ev de dimension finie, toute famille libre est finie.

En effet, toute partie libre est incluse dans une partie-base d'après le théorème IX.3.2, et toute partie-base est finie (corollaire de la proposition IX.3.4). ■

COROLLAIRE 2

|| Tout K -ev de dimension finie E admet au moins une base.

Si $E = \{0_E\}$, par convention, \emptyset est une base de E . Si $E \neq \{0_E\}$, en prenant $\mathcal{L} = \emptyset$ et $\mathcal{G} = E$ dans le théorème IX.3.2, on conclut à l'existence d'au moins une partie-base dans E . ■

Remarque 1 : On peut démontrer, en utilisant l'axiome du choix, que tout espace vectoriel, même de dimension infinie, admet une base et généraliser également le théorème IX.3.2, mais ces résultats dépassent le cadre de cet ouvrage (voir [27]).

Le théorème de la dimension finie

LEMME 1

|| Soit E un K -ev non nul, $p \in \mathbb{N}^*$, e_1, e_2, \dots, e_p des vecteurs de E et $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ des vecteurs de $\text{Vect}(e_1, e_2, \dots, e_p)$. Si les $(\varepsilon_i)_{1 \leq i \leq p}$ sont libres, alors e_1, e_2, \dots, e_p appartiennent à $\text{Vect}(\varepsilon_1, \dots, \varepsilon_p)$.

Démonstration :

On procède par récurrence sur p . Si $p = 1$, nécessairement $e_1 \neq 0_E$ et $\varepsilon_1 = \lambda_1 e_1$ avec $\lambda_1 \in K^*$ et alors $e_1 = \lambda_1^{-1} \varepsilon_1$.

Supposons le lemme vrai à l'ordre $p \geq 1$. Considérons $e_1, e_2, \dots, e_{p+1} \in E$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{p+1} \in \text{Vect}(e_1, \dots, e_{p+1})$ avec $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{p+1})$ libre. Écrivons $\varepsilon_{p+1} = \sum_{i=1}^{p+1} \alpha_i e_i$ ($\alpha_i \in K$). Puisque $\varepsilon_{p+1} \neq 0_E$ l'un des α_i au moins est $\neq 0$. Quitte à renuméroter les e_i , on peut supposer $\alpha_{p+1} \neq 0$. Alors $e_{p+1} \in \text{Vect}(e_1, e_2, \dots, e_p, \varepsilon_{p+1})$. Il existe donc $\lambda_1, \lambda_2, \dots, \lambda_p$ dans K tels que $\varepsilon_i + \lambda_i \varepsilon_{p+1} \in \text{Vect}(e_1, \dots, e_p)$ pour $1 \leq i \leq p$. Posons

$$\varepsilon'_i = \varepsilon_i + \lambda_i \varepsilon_{p+1} \quad (1 \leq i \leq p).$$

Les $(\varepsilon'_i)_{1 \leq i \leq p}$ sont libres car les $(\varepsilon_i)_{1 \leq i \leq p+1}$ le sont (il suffi

relation de dépendance linéaire entre les ε'_i pour s'en convaincre). L'hypothèse de récurrence s'applique, d'où

$$e_i \in \text{Vect}(\varepsilon'_1, \dots, \varepsilon'_p) \subset \text{Vect}(\varepsilon_1, \dots, \varepsilon_{p+1})$$

pour $1 \leq i \leq p$. De plus, $e_{p+1} \in \text{Vect}(e_1, \dots, e_p, \varepsilon_{p+1})$, d'où $e_{p+1} \in \text{Vect}(\varepsilon_1, \dots, \varepsilon_{p+1})$. ■

THÉORÈME IX.3.3

|| Soit E un K -ev non nul et de dimension finie. Toutes les bases de E ont le même nombre d'éléments.

Démonstration :

Soit \mathcal{B} et \mathcal{C} deux parties-base de E . On sait qu'il en existe (corollaire 2 du théorème IX.3.2). On sait de plus que \mathcal{B} et \mathcal{C} sont finies, et non vides. Posons $m = \text{card}(\mathcal{B})$, $n = \text{card}(\mathcal{C})$, avec par exemple $n \geq m$. Écrivons $\mathcal{B} = \{e_1, e_2, \dots, e_m\}$, $\mathcal{C} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$. La suite $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m)$ est libre, et $\varepsilon_i \in \text{Vect}(e_1, \dots, e_m)$ pour $1 \leq i \leq m$. Donc, d'après le lemme 1, $e_i \in \text{Vect}(\varepsilon_1, \dots, \varepsilon_m)$ pour $1 \leq i \leq m$. Donc $\mathcal{B} \subset \text{Vect}(\varepsilon_1, \dots, \varepsilon_m)$, ce qui entraîne que $\{\varepsilon_1, \dots, \varepsilon_m\}$ engendre le K -ev E . Donc, d'après la propriété (II) du théorème IX.3.1, c'est que $m = n$. ■

DÉFINITION IX.3.2

⎵ Soit E un K -ev non nul de dimension finie. On appelle **dimension** de E , et on note $\dim_K(E)$ (ou $\dim(E)$ s'il n'y a aucune ambiguïté sur K), l'entier $d \in \mathbb{N}^*$ tel que pour toute base $(e_i)_{i \in I}$ de E , on ait $\text{card}(I) = d$. Si $E = \{0_E\}$, on convient que $\dim_K(E) = 0$.

Exemple 1 : Soit $d \in \mathbb{N}^*$. On a vu que le K -ev K^d admet pour base la suite (e_1, \dots, e_d) telle que $e_i = (\delta_{ij})_{1 \leq j \leq d}$, avec $\delta_{ii} = 1_K$ et $\delta_{ij} = 0$ si $i \neq j$ (cf. § VI.3). C'est même la base canonique de K^d . Donc K^d est un K -ev de dimension finie, et sa dimension est d . Cet exemple montre qu'il existe des K -ev de toute dimension (finie). Nous avons vu qu'en particulier \mathbb{C} est un \mathbb{R} -ev de dimension 2, puisque le \mathbb{R} -ev \mathbb{C} n'est autre que $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Exemple 2 : Soit $d \in \mathbb{N}$. On a vu au § VII.1 que $(1, X, \dots, X^d)$ est une base du K -ev $K_d[X]$. Ce K -ev est donc de dimension finie, et sa dimension est $d + 1$.

En revanche, le K -ev $K[X]$ n'est pas de dimension finie puisqu'il admet la base $(1, X, \dots, X^d, \dots)$ qui est dénombrable mais pas finie.

Exemple 3 : Soit E un K -ev de dimension 1, et (b) une base de E . Toute autre base (c) est du type (λb) avec $\lambda \in K^*$. Pour une telle base (c) l'application $K \rightarrow E$, $\rho \mapsto \rho c$ est un isomorphisme de K -ev. Il n'y a dans E que deux sous- K -ev : $\{0_E\}$ et E . Réciproquement, un

n'admettant comme sous- K -ev que $\{0_E\}$ et E est de dimension 1. Par exemple \mathbb{C} est un \mathbb{C} -ev de dimension 1.

La notion de dimension d'un K -ev permet de classer les K -ev de dimension finie à isomorphisme près :

THÉORÈME IX.3.4

|| Pour que deux K -ev de dimension finie soient **isomorphes**, il faut et il suffit qu'ils aient **même dimension**.

Démonstration :

Soit E et F deux K -ev de dimension finie, $m = \dim_K(E)$ et $n = \dim_K(F)$. Nous supposons $m \geq 1$ et $n \geq 1$ pour que la question se pose vraiment. Soit donc $f: E \rightarrow F$ un isomorphisme. Une base (e_1, \dots, e_m) de E est transformée en une base $(f(e_i))_{1 \leq i \leq m}$ de F (cf. théorème VI.3.2), donc $\dim_K(F) = m$. Réciproquement, si $m = n$, soit $(\varepsilon_1, \dots, \varepsilon_m)$ une base de F . L'unique application K -linéaire $f: E \rightarrow F$ telle que $f(e_i) = \varepsilon_i$ pour $1 \leq i \leq m$ est un isomorphisme de K -ev (cf. théorème VI.3.2). ■

En conséquence, tout K -ev E de dimension finie $d \geq 1$ est isomorphe au K -ev K^d . Cependant en général il n'y a lieu de privilégier aucun isomorphisme particulier de E sur K^d .

Donnons pour terminer les définitions de quelques termes usuels. Dans un K -ev *quelconque* E (qu'il soit de dimension finie ou infinie, que K soit fini ou non), les sous- K -ev de dimension finie égale à 1 s'appellent les *droites vectorielles* de E (ce sont les Ka , où $a \in E \setminus \{0\}$) ; les sous- K -ev de dimension 2 s'appellent les *plans vectoriels*.

Exercice 1 : Soit \mathcal{L} une partie libre d'un K -ev E et b, c deux éléments de E . On suppose : $c \notin \text{Vect}(\mathcal{L})$ et $c \in \text{Vect}(\mathcal{L} \cup \{b\})$. Montrer que $b \in \text{Vect}(\mathcal{L} \cup \{c\})$ (« théorème d'échange »). En déduire une autre démonstration du théorème IX.3.3, en supposant acquis ce qui précède le théorème IX.3.2.

Exercice 2 : Soit E un K -ev muni d'une base $(e_i)_{i \in I}$. On suppose I totalement ordonné par une relation \leq . Soit $(\varepsilon_i)_{i \in I}$ une famille dans E telle que $(\forall i \in I) \varepsilon_i = \sum_{j \leq i} \lambda_{ij} e_j$, avec $\lambda_{ij} \in K^*$. Montrer que $(\varepsilon_i)_{i \in I}$ est une base de E .

Exercice 3 : a) Soit E un K -ev de dimension finie $d \geq 1$. On suppose le corps K fini, de cardinal q . Montrer que : $\text{card}(E) = q^d$.

b) Soit K un corps fini, de caractéristique p . Montrer que $\text{card}(K) = p^n$ pour un certain $n \in \mathbb{N}^*$.

c) Montrer que, pour chaque entier $k \in \llbracket 1, d \rrbracket$, le nombre de suites libres (x_1, \dots, x_k) à k termes dans E est : $(q^d - 1)(q^d - q) \dots (q^d - q^{k-1})$. En particulier quel est le nombre de bases de E , le nombre d'automorphismes du K -ev E ?

Exercice 4 : Soit $n \in \mathbb{N}^*$ et E un K -ev de dimension n , muni d'une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$. A toute permutation $\sigma \in \mathfrak{S}_n$ on associe l'unique $f_\sigma \in \text{Hom}_K(E)$ telle que $f_\sigma(e_i) = e_{\sigma(i)}$ ($1 \leq i \leq n$).

a) Montrer que $\sigma \mapsto f_\sigma$ définit un homomorphisme *injectif* du groupe \mathfrak{S}_n dans le groupe $\text{GL}_K(E)$ [utiliser le théorème VI.3.2].

b) Soit \mathcal{D} le sous- K -ev de E engendré par le vecteur $u = \sum_{i=1}^n e_i$, et \mathcal{H} celui formé des $x = \sum_{i=1}^n x_i e_i$ tels que $\sum_{i=1}^n x_i = 0$. Montrer que $E = \mathcal{H} \oplus \mathcal{D}$ [on pourra utiliser la base $(e_1 - e_i)_{2 \leq i \leq n}$ de \mathcal{H}].

c) Montrer que les seuls sous- K -ev V de E tels que $(\forall \sigma \in \mathfrak{S}_n, \forall x \in V) f_{\sigma(x)} \in V$ sont $\{0_E\}$, \mathcal{D} , \mathcal{H} et E [si $V \not\subset \mathcal{D}$, alors $e_1 - e_i \in V$].

Exercice 5 : Soit K un corps de caractéristique $\neq 3$ et E le K -ev $K_2[X]$.

a) Montrer que $(1, X+1, X^2-X+1)$ est une base de E .

b) A chaque $P \in E$ on fait correspondre l'ensemble \mathcal{P} obtenu en multipliant P par un polynôme arbitraire M de $K[X]$ et en prenant le reste mod (X^3+1) de tous les PM . Montrer qu'on obtient ainsi pour \mathcal{P} un sous- K -ev de E .

c) Montrer que les sous- K -ev ainsi associés à $S = X+1$ et à $T = X^2-X+1$ sont supplémentaires. Quelles sont leurs dimensions respectives ?

Exercice 6 : Soit u un endomorphisme d'un K -ev de dimension finie. Montrer que pour que $\text{Ker}(u)$ et $\text{Im}(u)$ soient supplémentaires, il faut et il suffit que $\text{Im}(u^2) = \text{Im}(u)$.

Exercice 7 : Soit E le \mathbb{R} -espace vectoriel des fonctions définies sur $] -1, +1[$ et à valeurs dans \mathbb{R} engendré par

$$g_1 : x \mapsto \sqrt{\frac{1+x}{1-x}}, \quad g_2 : x \mapsto \sqrt{\frac{1-x}{1+x}}, \quad g_3 : x \mapsto \frac{1}{\sqrt{1-x^2}} \quad \text{et} \quad g_4 : x \mapsto \frac{x}{\sqrt{1-x^2}}.$$

Quelle est la dimension du \mathbb{R} -ev E ? En donner une base.

Exercice 8 : Soit E un K -ev de dimension $n \in \mathbb{N}^*$ et $\mathcal{C}(E)$ l'ensemble des *applications constantes* de E dans E . Montrer que $\mathcal{C}(E)$ est un K -ev. Déterminer sa dimension.

Exercice 9 : Soit E un K -ev de dimension finie $n \geq 2$. Montrer que le K -ev $\text{Hom}_K(E)$ est engendré par $\text{Hom}_K(E) \setminus \text{GL}_K(E)$ et qu'il est aussi engendré par $\text{GL}_K(E)$.

Exercice 10 : Soit dans \mathbb{R}^4 les vecteurs $(1, 0, 0, -1)$, $(2, 1, 1, 0)$, $(1, 1, 1, 1)$, $(1, 2, 3, 4)$, $(0, 1, 2, 3)$. Extraire de ces cinq vecteurs une base du sous-espace qu'ils engendrent.

§ IX.4 PROPRIÉTÉS DES ESPACES DE DIMENSION FINIE

THÉORÈME IX.4.1

Soit E un K -ev de dimension finie $n \geq 1$. Toute famille **libre** de E a **au plus n éléments**, et elle en a n ssi c'est une base de E . Toute famille **génératrice** du K -ev E a **au moins n éléments**, et elle en a n ssi c'est une base de E .

Démonstration :

Il suffit de prouver le théorème avec des *parties libres* (resp. *génératrices*). Soit donc \mathcal{L} une partie libre (resp. \mathcal{G} une partie génératrice) du K -ev E . Le théorème IX.3.2 fournit une partie-base \mathcal{B} de E telle que $\mathcal{L} \subset \mathcal{B}$ (resp. $\mathcal{B} \subset \mathcal{G}$). On a $\text{card}(\mathcal{B}) = n$ d'après

de la dimension finie. Or deux ensembles finis dont l'un est inclus dans l'autre sont égaux ssi ils ont même cardinal, d'où le théorème IX.4.1. ■

THÉOREME IX.4.2

|| Soit E un K -ev de dimension finie $n \geq 1$ et F un sous- K -ev non nul de E . Alors F est de dimension finie ; on a $\dim_K(F) \leq \dim_K(E)$, et l'égalité a lieu ssi $F = E$.

Démonstration :

Une famille de vecteurs de F est libre dans le K -ev F ssi elle l'est dans le K -ev E . On vient de voir qu'une famille libre de E a au plus n éléments (théorème IX.4.1). Il en va de même d'une famille libre de F , qui a *au moins un* élément car $F \neq \{0_E\}$. On peut donc définir l'entier $p = \text{maximum du nombre d'éléments d'une famille libre dans } F$, et on a : $p \geq 1$. Soit alors (e_1, e_2, \dots, e_p) une suite libre à p termes dans F . C'est une base de F , en vertu du théorème IX.3.1 (III). Donc F est bien de dimension finie ; sa dimension est $p \leq n$. Le théorème IX.4.1 montre enfin qu'on a : $p = n$ ssi (e_1, \dots, e_p) est une base de E , c'est-à-dire si $F = E$. ■

THÉOREME IX.4.3 (théorème de la base incomplète)

|| Soit E un K -ev de dimension finie $n \geq 2$, et soit $p \in \llbracket 1, n-1 \rrbracket$. Si (e_1, e_2, \dots, e_p) est une suite libre dans E , on peut trouver des vecteurs e_{p+1}, \dots, e_n de E tels que (e_1, \dots, e_n) soit une base de E .

Démonstration :

Par le théorème IX.3.2, on obtient une partie-base \mathcal{B} de E telle que $\{e_1, \dots, e_p\} \subset \mathcal{B}$. On a $\text{card}(\mathcal{B}) = n$ et il suffit d'ordonner les éléments de $\mathcal{B} \setminus \{e_1, \dots, e_p\}$ en une suite (e_{p+1}, \dots, e_n) pour conclure. ■

Remarque 1 : Malgré sa simplicité, le théorème IX.4.3 est sans doute le plus profond des théorèmes de la théorie des espaces vectoriels : contrairement au théorème IX.3.3 (de la dimension) qui n'utilise que la structure d'anneau commutatif du corps de base K , le théorème IX.4.3 utilise à plein la structure de corps de K .

COROLLAIRE

|| Soit F un sous- K -ev d'un K -ev de dimension finie ; alors F admet au moins un supplémentaire dans E .

Démonstration :

Bornons-nous au cas où $\dim_K(E) = n \geq 2$ et où $p = \dim_K(F) \in \llbracket 1, n-1 \rrbracket$. Soit (e_1, \dots, e_p) une base du K -ev F . On la complète en une base (e_1, \dots, e_n) de E et on constate que $G = \text{Vect}(e_{p+1}, \dots, e_n)$ est un supplémentaire de F dans E (cf. § IX.2, exemple 3). ■

THÉORÈME IX.4.4

|| Soit E et F deux K -ev de dimension finie. Alors $E \times F$ est de dimension finie, et on a : $\dim_K (E \times F) = \dim_K (E) + \dim_K (F)$.

Démonstration :

Si E (resp. F) = $\{0_E\}$ alors $E \times F \cong F$ (resp. $\cong E$) et il n'y a rien à prouver. Si E et F sont non nuls, soit (e_1, e_2, \dots, e_p) et $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q)$ des bases respectives de E et F ($p \geq 1, q \geq 1$). En posant $v_1 = (e_1, 0_F), \dots, v_p = (e_p, 0_F); v_{p+1} = (0_E, \varepsilon_1), \dots, v_{p+q} = (0_E, \varepsilon_q)$, on vérifie que (v_1, \dots, v_{p+q}) est une base de $E \times F$. ■

COROLLAIRE 1

|| Si E_1, \dots, E_n sont des K -ev de dimension finie, alors $E_1 \times E_2 \times \dots \times E_n$ l'est aussi, et sa dimension est $\sum_{i=1}^n \dim_K (E_i)$. Si V_1, \dots, V_n sont des sous- K -ev d'un K -ev E , tous de dimension finie et si $E = \bigoplus_{i=1}^n V_i$, alors E est aussi de dimension finie et sa dimension est $\sum_{i=1}^n \dim_K (V_i)$.

La première partie se déduit par récurrence à partir du théorème IX.4.4.

La seconde partie en découle, puisque si $E = \bigoplus_{i=1}^n V_i$, E est isomorphe au

K -ev $\bigtimes_{i=1}^n V_i$. ■

En particulier, on notera que si F et G sont deux sous- K -ev supplémentaires du K -ev E , E est de dimension finie ssi F et G le sont ; et lorsqu'il en est ainsi, on a : $\dim_K (E) = \dim_K (F) + \dim_K (G)$.

COROLLAIRE 2 (formule du rang)

|| Soit E un K -ev de dimension finie, F un K -ev quelconque, et $f \in \text{Hom}_K (E, F)$. Alors :

$$\dim_K (E) = \dim_K (\text{Ker} (f)) + \dim_K (\text{Im} (f))$$

Démonstration :

On sait déjà que $\text{Im} (f)$ et $\text{Ker} (f)$ sont de dimension finie. Soit S un supplémentaire de $\text{Ker} (f)$ dans E (corollaire du théorème IX.4.3).

On sait que S est isomorphe à $\text{Im} (f)$ (théorème I)

$\dim_K (\text{Im} (f)) = \dim_K (S)$. Mais $\dim_K (\text{Ker} (f)) + \dim_K (S) = \dim_K (E)$ résulte du corollaire 1, d'où le résultat encadré. ■

Ce corollaire 2 s'appelle *formule du rang* parce que, par définition, le **rang** d'une application linéaire (que l'on note $\text{rg} (f)$) est la dimension (finie ou non) de son image.

COROLLAIRE 3

|| Soit F et G deux sous- K -ev de dimension finie d'un K -ev E . Alors $F + G$ est de dimension finie, et on a :

$$\dim_K (F + G) = \dim_K (F) + \dim_K (G) - \dim_K (F \cap G).$$

Démonstration :

Soit $s : F \times G \rightarrow E$, $(x, y) \mapsto x + y$ l'application canonique. Son image est $F + G$; son noyau N est l'ensemble des couples $(x, -x)$, $x \in F \cap G$. L'application $F \cap G \rightarrow N$, $x \mapsto (x, -x)$ est un isomorphisme de K -ev, d'où $\dim_K (N) = \dim_K (F \cap G)$. La formule du rang s'applique à s (car $F \times G$ est de dimension finie) et fournit : $\dim_K (F \times G) = \dim_K (N) + \dim_K (F + G)$, d'où le résultat, puisque $\dim_K (F \times G) = \dim_K (F) + \dim_K (G)$ et $\dim_K (N) = \dim_K (F \cap G)$. ■

THÉORÈME IX.4.5

|| Soit L une extension du corps K , soit E un L -ev et $E_{(K)}$ le K -ev déduit de E par restriction des scalaires à K ⁽¹⁾. Pour que $E_{(K)}$ soit de dimension finie, il faut et il suffit que E soit de dimension finie sur L et que L soit un K -ev de dimension finie. Si c'est le cas, on a :

$$\dim_K (E_{(K)}) = \dim_L (E) \times \dim_K (L).$$

Démonstration :

Le lecteur vérifiera sans difficulté que si $E_{(K)}$ est de dimension finie, alors E est un L -ev de dimension finie, et L un K -ev de dimension finie. Supposons inversement que (a_1, \dots, a_p) est une base du K -ev L et (v_1, \dots, v_q) une base du L -ev E . Nécessairement $p \geq 1$ (car $K \subset L$) ; on peut également supposer $q \geq 1$, seul cas non trivial. Montrons qu'alors la famille $(a_i v_j)_{(i,j) \in \llbracket 1,p \rrbracket \times \llbracket 1,q \rrbracket}$ est une base de $E_{(K)}$, ce qui achèvera la démonstration.

Cette famille est libre, car si $(\rho_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$ est une famille dans K telle que $\sum_{i,j} \rho_{ij} a_i v_j = 0_E$, on en déduit $\sum_{j=1}^q \left(\sum_{i=1}^p \rho_{ij} a_i \right) v_j = 0_E$, d'où d'abord $\sum_{i=1}^p \rho_{ij} a_i = 0$ pour tout j par L -indépendance linéaire des v_j , et

⁽¹⁾ On pourra se reporter au § VI.1, exemple 3.

ensuite $\rho_{ij} = 0$ pour tous i et j par K -indépendance linéaire des a_i . Enfin, cette famille engendre le K -ev $E_{(K)}$, car si $x \in E$, on a $x = \sum_{j=1}^q \lambda_j v_j$ pour des $\lambda_j \in L$ appropriés, puis, pour chaque j , $\lambda_j = \sum_{i=1}^p \rho_{ij} a_i$ pour des $\rho_{ij} \in K$ appropriés, d'où $x = \sum_{(i,j) \in \llbracket 1,p \rrbracket \times \llbracket 1,q \rrbracket} \rho_{ij} a_i v_j$. ■

Exemple 1 : Par restriction des scalaires à \mathbb{R} , un \mathbb{C} -ev E de dimension finie n fournit un \mathbb{R} -ev de dimension finie $2n$.

Espaces d'applications linéaires

THÉORÈME IX.4.6

|| Soit E et F deux K -ev de dimension finie. Le K -ev $\text{Hom}_K(E, F)$ est aussi de dimension finie, et sa dimension est $\dim_K(E) \times \dim_K(F)$.

Démonstration :

Bornons-nous au cas où E et F , non nuls, ont pour bases respectives (e_1, \dots, e_p) et $(\varepsilon_1, \dots, \varepsilon_q)$, avec $p = \dim_K(E) \geq 1$ et $q = \dim_K(F) \geq 1$. Pour chaque $(i, j) \in \llbracket 1, q \rrbracket \times \llbracket 1, p \rrbracket$, le théorème VI.3.2 fournit $u_{ij} \in \text{Hom}(E, F)$ tel que $u_{ij}(e_j) = \varepsilon_i$, $u_{ij}(e_k) = 0$ si $k \neq j$. Il suffit de prouver que (u_{ij}) est une base du K -ev $\text{Hom}_K(E, F)$. La famille (u_{ij}) est libre, car si (λ_{ij}) est une famille de scalaires telle que $\sum_{i,j} \lambda_{ij} u_{ij} = 0$, on a, pour chaque k ($1 \leq k \leq p$) $\sum_{i,j} \lambda_{ij} u_{ij}(e_k) = 0$, d'où $\sum_{i=1}^q \lambda_{ik} \varepsilon_i = 0$, d'où $\lambda_{ik} = 0$ pour $1 \leq i \leq q$ puisque les (ε_i) sont indépendants, et finalement tous les λ_{ij} sont nuls.

La famille (u_{ij}) engendre le K -ev $\text{Hom}_K(E, F)$ car si $u \in \text{Hom}_K(E, F)$, en posant $u(e_j) = \sum_{i=1}^q a_{ij} \varepsilon_i$ pour $1 \leq j \leq p$ (les a_{ij} dans K), on vérifie que $u - \sum_{i,j} a_{ij} u_{ij} = v$ prend la valeur 0 sur chaque e_j ($1 \leq j \leq p$), donc $v = 0$, d'où $u = \sum_{i,j} a_{ij} u_{ij}$. ■

Application au dual

DÉFINITION IX.4.1

⎧ Si E est un K -ev, on appelle **dual algébrique** de E (ou **dual de E**), et
 ⎧ on note E^* , le K -ev $\text{Hom}_K(E, K)$. Les éléments de E^* s'appellent
 ⎧ les **formes linéaires** sur E .

Le dual d'un K -ev sera systématiquement étudié au (

Bornons-nous ici à quelques conséquences simples de la théorie de la dimension finie.

THÉORÈME IX.4.7

|| Soit E un K -ev de dimension finie $n \geq 1$. Alors E^* est aussi de dimension finie, et sa dimension est n .

C'est une conséquence immédiate du théorème IX.4.6, compte tenu du fait que $\dim_K(K) = 1$. ■

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base du K -ev E de dimension finie $n \geq 1$. D'après la démonstration du théorème IX.4.6, une base du K -ev E^* est la suite $(\varphi_1, \varphi_2, \dots, \varphi_n)$ telle que $\varphi_i(e_j) = \delta_{ij}$ pour tous $(i, j) \in \llbracket 1, n \rrbracket^2$, où $\delta_{ii} = 1_K$ et $\delta_{ij} = 0$ si $i \neq j$. Cette base particulière $(\varphi_1, \dots, \varphi_n)$ de E^* s'appelle **la base duale de \mathcal{B}** et se note souvent $\mathcal{B}^* = (e_1^*, e_2^*, \dots, e_n^*)$, notation qu'il faut utiliser avec prudence, car elle risque de faire oublier que chaque e_i^* dépend de toute la base \mathcal{B} . Soit $x = \sum_{i=1}^n x_i e_i \in E$ ($(x_1, \dots, x_n) \in K^n$). Pour chaque $i \in \llbracket 1, n \rrbracket$, on a

$$\varphi_i(x) = \sum_{j=1}^n x_j \varphi_i(e_j) = x_i.$$

Ainsi, φ_i fait correspondre, à chaque vecteur $x \in E$, sa i -ième coordonnée dans la base \mathcal{B} , et on peut écrire la relation fondamentale :

$$(\forall x \in E) \quad x = \sum_{i=1}^n \varphi_i(x) e_i.$$

Rang ; caractérisation des bijections linéaires

Par définition, le **rang** d'une famille $(a_i)_{i \in I}$ de vecteurs d'un K -ev E est la dimension, finie ou non, du sous- K -ev $\text{Vect}((a_i)_{i \in I})$ de E . Cette notion de rang se relie de manière évidente à celle de *rang d'une application linéaire*. En effet supposons que la famille $(a_i)_{i \in I}$ engendre le K -ev E ; si F est une autre K -ev et si $f \in \text{Hom}_K(E, F)$, le rang de f est, par définition,

$$\dim(\text{Im}(f)), \text{ c'est-à-dire : } \dim[\text{Vect}((f(a_i))_{i \in I})].$$

Donc le rang de f n'est autre que le rang de la famille $(f(a_i))_{i \in I}$ pour toute famille génératrice $(a_i)_{i \in I}$ de E .

THÉORÈME IX.4.8

|| Soit E et F deux K -ev de **même** dimension finie $n \geq 1$, et soit $u \in \text{Hom}_K(E, F)$.

Les conditions suivantes sont équivalentes :

- (I) u est bijectif ;
- (II) u est injectif ;
- (III) u est surjectif ;
- (IV) $\text{rg}(u) = n$;
- (V) u est inversible à droite ;
- (VI) u est inversible à gauche.

Démonstration :

On applique la formule du rang : $\dim(E) = n = \text{rg}(u) + \dim(\text{Ker}(u))$. Immédiatement (II) \Leftrightarrow (IV). Ensuite le théorème IX.4.2 montre que (I) \Leftrightarrow (II). Il est clair que (I) \Rightarrow (III) et que (III) \Rightarrow (IV). Voilà pour les quatre premières propriétés. De façon évidente (I) \Rightarrow (V) et (I) \Rightarrow (VI). Supposons u inversible à droite et soit $v \in \text{Hom}_K(F, E)$ tel que $u \circ v = \text{Id}_F$. Alors u est surjectif, puisque $u \circ v$ l'est, d'où (V) \Rightarrow (III).

Supposons enfin u inversible à gauche et soit $w \in \text{Hom}(F, E)$ tel que $w \circ u = \text{Id}_E$. Alors u est injectif, puisque $w \circ u$ l'est, d'où (VI) \Rightarrow (II). ■

Exercice 1 : Soit E et F deux K -ev de dimensions finies respectives $n \geq 1$ et $p \geq 1$, et soit G un sous- K -ev de E . On pose $\mathcal{L}_G = \{u \in \text{Hom}_K(E, F) \mid G \subset \text{Ker}(u)\}$, et $r = \dim(G)$.

- a) \mathcal{L}_G est un sous- K -ev de $\text{Hom}_K(E, F)$.
- b) Calculer $\dim_K(\mathcal{L}_G)$ en fonction de n , p et r .

Exercice 2 : Soit A, B, C, D des K -ev de dimension finie. On considère des applications K -linéaires $\alpha, \beta, \gamma: A \xrightarrow{\alpha} B \xrightarrow{\beta} C \xrightarrow{\gamma} D$. Montrer : $\text{rg}(\beta\alpha) + \text{rg}(\gamma\beta) \leq \text{rg}(\beta) + \text{rg}(\gamma\beta\alpha)$.

Exercice 3 : Soit E un K -ev de dimension finie $n \geq 1$, et u, v dans $\text{Hom}_K(E)$. Prouver : $\text{rg}(u) + \text{rg}(v) \leq n + \text{rg}(u \circ v) \leq n + \text{Min}(\text{rg}(u), \text{rg}(v))$.

Exercice 4 : Soit x_1, \dots, x_n dans \mathbb{R} avec $x_1 < x_2 < \dots < x_n$. On note E le \mathbb{R} -ev des fonctions de classe C^1 : $f: \mathbb{R} \rightarrow \mathbb{R}$ telles que :

$$f|_{]-\infty, x_1]}, f|_{[x_n, +\infty[} \text{ et } f|_{[x_i, x_{i+1}]} \quad (1 \leq i \leq n-1)$$

soient des fonctions polynomiales de degré ≤ 2 . Montrer que E est de dimension finie ; préciser cette dimension ; trouver une base de E .

Exercice 5 : Soit E un K -ev de dimension finie $n \geq 1$, et u, v éléments de $\text{Hom}_K(E)$. Montrer : $\text{rg}(u \circ v) \geq \text{rg}(u) + \text{rg}(v) - n$.

Exercice 6 : Soit E_1, \dots, E_n des K -ev de dimensions finies respectives d_1, d_2, \dots, d_n reliés par des applications K -linéaires $E_1 \xrightarrow{u_1} E_2 \rightarrow \dots \xrightarrow{u_{n-1}} E_n$. On suppose u_1 injective, u_{n-1} surjective, et $\text{Im}(u_i) = \text{Ker}(u_{i+1})$ pour $i \leq n-2$. Démontrer que $\sum_{i=1}^n (-1)^i d_i = 0$.

Exercice 7 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. On pose, pour $k \in \mathbb{N}$, $J_k = u^k(E)$, $N_k = \text{Ker}(u^k)$. On rappelle que $u^0 = \text{Id}_E$.

- a) Les suites (J_k) et (N_k) sont stationnaires.
- b) Si ν est le plus petit entier p tel que $J_p = J_{p+1}$, on a : $N_\nu = N_{\nu+1}$ et $N_k \neq N_{k+1}$ pour $k \leq \nu$.

c) On a : $E = J_v \oplus N_v$; $u(J_v) \subset J_v$, $u(N_v) \subset N_v$, et $u|_{J_v}$ est un automorphisme, et $u|_{N_v}$ est nilpotent (c'est-à-dire, $v = u|_{N_v}$ vérifie $v^k = 0$ pour $k \geq 1$ convenable).

Exercice 8 : Soit u et v deux éléments de $\text{Hom}_K(E, F)$ avec E et F , K -ev de dimension finie. Montrer :

$$(*) \quad |\text{rg}(u) - \text{rg}(v)| \leq \text{rg}(u+v) \leq \max(\dim(E), \dim(F), \text{rg}(u) + \text{rg}(v)).$$

Montrer que, pour $\text{rg}(u)$ et $\text{rg}(v)$ donnés, $\text{rg}(u+v)$ peut prendre toutes les valeurs possibles autorisées par (*).

Exercice 9 : Soit A et B deux sous- K -ev de même dimension p d'un K -ev E de dimension finie n , le corps de base K étant infini. Prouver qu'il existe un sous- K -ev de E tel que $A \oplus X = B \oplus X = E$.

Généraliser avec N sous- K -ev A_1, A_2, \dots, A_N de E , de même dimension.

Exercice 10 : Soit E un K -ev de dimension finie $n \geq 1$, et u, v des éléments de $\text{Hom}_K(E)$ tels que $uv = 0$ et $u+v \in \text{GL}_K(E)$. Montrer $\text{rg}(u) + \text{rg}(v) = n$.

Exercice 11 : Soit E un K -ev de dimension finie, et u, v dans $\text{Hom}_K(E)$. On suppose : $\text{Im}(u) + \text{Im}(v) = \text{Ker}(u) + \text{Ker}(v)$. Montrer que ces sommes sont directes. Cela reste-t-il vrai si E n'est pas de dimension finie ?

Exercice 12 : Soit E un K -ev et (x_1, \dots, x_n) une suite de n vecteurs de rang s . Soit $r < n$; on suppose la suite (x_1, \dots, x_r) de rang s' . Montrer $s' \geq r + s - n$.

Exercice 13 : Soit \mathcal{P} l'ensemble des projecteurs d'un K -ev E . Si $p \in \mathcal{P}$ et $q \in \mathcal{P}$ on note $p \leq q$ ssi $pq = qp = p$.

a) Montrer que (\mathcal{P}, \leq) est un ensemble ordonné.

b) Soit p et q dans \mathcal{P} avec $pq = qp$. On pose $p \wedge q = pq$ et $p \vee q = p + q - pq$.

Montrer : $p \wedge q \in \mathcal{P}$, et $p \wedge q$ est la borne inférieure de $\{p, q\}$ dans \mathcal{P} , et $\text{Im}(p \wedge q) = \text{Im}(p) \cap \text{Im}(q)$.

Montrer : $p \vee q \in \mathcal{P}$, $p \vee q$ est la borne supérieure de $\{p, q\}$ dans \mathcal{P} , et $\text{Im}(p \vee q) = \text{Im}(p) + \text{Im}(q)$.

Exercice 14 : Soit E et F deux K -ev de dimension finie, $f \in \text{Hom}_K(E, F)$, $g \in \text{Hom}_K(F, E)$. On fait les hypothèses $g \circ f \circ g = g$ et $f \circ g \circ f = f$.

a) Montrer que : $E = \text{Im}(g) \oplus \text{Ker}(f)$.

b) Comparer les rangs de f et de g .

Exercice 15 (Tout corps fini est commutatif) :

On considère un corps fini K , de cardinal q , et on note Γ l'ensemble $\{x \in K \mid \forall y \in K, xy = yx\}$, et $r = \text{card}(\Gamma)$.

a) Vérifier que Γ est un sous-corps de K et un corps commutatif, dont on notera p la caractéristique. Le sous-corps premier de Γ (isomorphe à $\mathbb{Z}/p\mathbb{Z}$) sera noté F . On considérera K comme une F -algèbre, et on posera $e = \dim_F(K)$.

Vérifier que $r = p^d$ pour $d \in \mathbb{N}^*$, et que $q = p^{de} = r^e$, $de = \dim_F(K)$.

b) Soit L un sous-corps quelconque de K , distinct de K ; on considère ici la structure de F -ev de K ; démontrer qu'il existe $\alpha \in \mathbb{N}$ ($\alpha \geq 2$) et des éléments $x_1 = 1_K, x_2, \dots, x_\alpha$ de K tels que $K = \bigoplus_{i=1}^{\alpha} x_i L$, et en déduire : $q = (\text{card}(L))^\alpha$.

On se propose de montrer par l'absurde que $e = 1$, et on fait donc l'hypothèse : $(\mathcal{H}) e > 1$.

c) On fait opérer le groupe multiplicatif K^* sur lui-même à gauche par automorphismes intérieurs (voir exemple 3, § V.6), c'est-à-dire par $(g, x) \mapsto gxg^{-1}$. (Les orbites singleton sont donc les $\{x\}_{x \in \Gamma^*}$). On désigne par \mathcal{C} une partie de K^* contenant un élément et un seul de chaque K^* -orbite. Pour $a \in \mathcal{C}$, on note G_a le stabilisateur de a . Démontrer :

$$(1) \quad \text{card}(K^*) = \sum_{a \in \mathcal{C}} (\text{card}(K^*) / \text{card}(G_a)).$$

d) Soit $a \in \mathcal{C}$; montrer que $L_a = G_a \cup \{0\}$ est un sous-corps de K contenant Γ . En déduire l'existence d'un diviseur ν_a de e tel que $\text{card}(L_a) = r^{\nu_a}$. (Utiliser b)). A l'aide de (1), en déduire :

$$(2) \quad r^e - 1 = r - 1 + \sum_{a \in \mathcal{C}, \nu_a \neq e} \frac{r^e - 1}{r^{\nu_a} - 1}.$$

e) Soit $\Phi_e(X)$ le polynôme cyclotomique d'ordre e dans $\mathbb{C}[X]$. (Voir exercices 12-13, § IX.7 ci-dessous : mais seule la définition et les propriétés évidentes vues dans l'exercice 16, § VII.4 sont ici utiles.) Montrer que si $\nu_a < e$, alors $\Phi_e(r)$ divise $\frac{r^e - 1}{r^{\nu_a} - 1}$ dans \mathbb{N}^* .

f) En déduire, avec (2), que $\Phi_e(r)$ divise $r - 1$. Aboutir à une contradiction en étudiant $|\Phi_e(r)|$, en déduire que (\mathcal{H}) ne tient pas, et donc que K est commutatif.

§ IX.5 HYPERPLANS

DÉFINITION IX.5.1

Soit E un K -ev non nul ; on appelle **hyperplan** de E tout sous- K -ev de E **distinct de E** tel que les seuls sous- K -ev de E contenant H soient H et E .

Si l'on ordonne par inclusion l'ensemble $\mathcal{G}(E)$ des sous- K -ev de E autres que E , cela revient à dire que les hyperplans sont les **éléments maximaux** de cet ensemble ordonné.

THÉORÈME IX.5.1

Soit E un K -ev non nul et H un sous- K -ev de E autre que E .

(I) Si H admet un supplémentaire de dimension 1, alors c'est un hyperplan.

(II) Si H est un hyperplan, pour tout $a \in E \setminus H$, on a $E = H \oplus Ka$ et donc les supplémentaires de H sont de dimension 1, (cf. théorème IX.1.3).

Démonstration :

(I) Supposons $E = H \oplus D$, où D est une droite vectorielle. On a : $D = Ka$, avec $a \in E \setminus H$. Soit F un sous- K -ev de E contenant strictement H . Prenons $b \in F \setminus H$ et écrivons $b = \lambda a + c$, avec $c \in H$ et $\lambda \in K$. Nécessairement $\lambda \neq 0$ puisque $b \notin H$, d'où $a = \lambda^{-1}(b - c) \in F$, et par suite $H \oplus D \subset F$, d'où $F = E$. Donc H est un hyperplan.

(II) Supposons que H est un hyperplan, et soit $a \in E \setminus H$. La droite vectorielle $D = Ka$ vérifie $D \cap H = \{0_E\}$ puisque $a \notin H$. Donc le sous- K -ev $H \oplus D$ de E contient H strictement et par suite $H \oplus D = E$. ■

Il résulte de ce théorème que, pour un hyperplan H de E , l'existence de supplémentaires de H dans E se trouve démontrée. On savait déjà que tous ces supplémentaires sont isomorphes entre eux. Dans le cas d'

tous les supplémentaires sont de dimension 1 et on sait comment les former.

COROLLAIRE 1

|| Si E est un K -ev **de dimension finie** $n \geq 1$, les hyperplans de E sont les sous- K -ev de dimension $n - 1$.

Démonstration :

Soit H un sous- K -ev de E autre que E . On sait que H admet au moins un sous- K -ev supplémentaire dans E (corollaire du théorème de la base incomplète IX.4.3). Choisissons un tel sous- K -ev S . Si $p = \dim_K(H)$, on a $p \leq n - 1$ et $\dim(S) = n - p$. D'après le théorème IX.5.1, pour que H soit un hyperplan, il faut et il suffit que $\dim(S) = 1$, c'est-à-dire que $p = n - 1$. ■

COROLLAIRE 2

|| Soit E un K -ev de dimension finie $n \geq 1$; si F est un sous- K -ev de E distinct de E , il est contenu dans au moins un hyperplan de E ; et F est l'intersection des hyperplans qui le contiennent.

Démonstration :

Soit $p = \dim(F)$; si $1 \leq p \leq n - 1$, choisissons une base $\mathcal{B} = (e_1, \dots, e_n)$ de E obtenue en complétant une base (e_1, \dots, e_p) de F ; alors d'après le corollaire 1 ci-dessus, le sous- K -ev H engendré par (e_1, \dots, e_{n-1}) est un hyperplan de E , et il est clair que $F \subset H$; soit d'autre part un vecteur $x \in E \setminus F$; les vecteurs (e_1, \dots, e_p, x) sont linéairement indépendants ; le théorème de la base incomplète permet donc de construire une base (f_1, \dots, f_n) de E telle que $f_i = e_i$ pour $1 \leq i \leq p$ et $f_n = x$; alors l'hyperplan $L = \text{Vect}(f_1, \dots, f_{n-1})$ contient F , et ne contient pas x . Si $F = \{0_E\}$, raisonnement analogue. ■

Hyperplans et formes linéaires

THÉORÈME IX.5.2

|| Soit E un K -ev non nul.

(I) Si φ est une forme linéaire non nulle sur E , alors $H = \text{Ker}(\varphi)$ est un hyperplan de E .

(II) Réciproquement, soit H un hyperplan de E . Il existe des formes linéaires φ sur E telles que $\text{Ker}(\varphi) = H$. Si φ_0 est l'une d'elles, l'ensemble des formes linéaires φ telles que $H \subset \text{Ker}(\varphi)$ est $\{\lambda \varphi_0\}_{\lambda \in K} = K\varphi_0$ (droite vectorielle engendrée par φ_0 dans E^*).

Démonstration :

(I) Soit $\varphi \in E^* \setminus \{0\}$. Alors $H = \text{Ker}(\varphi)$ est un sous- K -ev de E distinct de E , et $\text{Im}(\varphi)$ est un sous- K -ev non nul

donc $\text{Im}(\varphi) = K$. Choisissons $a \in E \setminus H$, c'est-à-dire $a \in E$ tel que $\varphi(a) \neq 0$, d'où, si $D = Ka$, $D \cap H = \{0\}$. Soit $b \in E$; le vecteur $c = b - \rho a$, avec $\rho = (\varphi(a))^{-1} \varphi(b)$, vérifie : $\varphi(c) = \varphi(b) - \rho \varphi(a) = 0$, donc $c \in H$; or $b = c + \rho a$; on a donc $E \subset H \oplus D$ et finalement $E = H \oplus D$, donc H est un hyperplan.

(II) Soit $a \in E \setminus H$, d'où $E = H \oplus D$ avec $D = Ka$. L'application $\zeta : K \rightarrow D$, $\lambda \mapsto \lambda a$, est un isomorphisme de K -ev. Soit p la projection de E sur D parallèlement à H et $f : E \rightarrow D$, $x \mapsto p(x)$. Alors $\varphi = \zeta^{-1} \circ f : E \rightarrow K$ est K -linéaire, surjective, et de noyau H . Fixons enfin $\varphi_0 \in E^* \setminus \{0\}$ telle que $\text{Ker}(\varphi_0) = H$. Il est clair que, pour $\lambda \in K$, $H \subset \text{Ker}(\lambda \varphi_0)$. Réciproquement, soit $\varphi \in E^*$ telle que $H \subset \text{Ker}(\varphi)$. La forme linéaire $\psi = \varphi_0(a) \varphi - \varphi(a) \varphi_0$ est nulle sur H ainsi qu'en a ; elle est donc nulle sur D , et donc sur $H \oplus D = E$, donc $\psi = 0$. De plus $\varphi_0(a) \neq 0$ car $a \notin H$, d'où $\varphi = (\varphi_0(a))^{-1} \varphi_0 \in K \varphi_0$. ■

Exercice 1 : Le corps de base K est supposé fini, de cardinal q , et le K -ev E est supposé de dimension finie $n \geq 2$. On considère une famille $(H_\lambda)_{\lambda \in \Lambda}$ d'hyperplans de E telle que $\bigcup_{\lambda \in \Lambda} H_\lambda = E$. Démontrer : $\text{card}(\Lambda) \geq q + 1$.

Exercice 2 : Avec les hypothèses de l'exercice 1, donner en fonction de n et q le nombre d'hyperplans de E , et le nombre de supplémentaires d'un hyperplan donné de E .

Exercice 3 : Soit E un \mathbb{Q} -ev admettant une base dénombrable $(e_i)_{i \in I}$ (I équipotent à \mathbb{N}). Démontrer que l'ensemble E est dénombrable.

On considère alors \mathbb{R} comme \mathbb{Q} -ev; on admet l'existence d'une base $(x_i)_{i \in I}$ de ce \mathbb{Q} -ev (une telle base est dite *de Hamel*, mais on ne cherchera pas à en exhiber une !). Démontrer que I est équipotent à \mathbb{R} . Soit $i_0 \in I$ fixé, et soit φ la \mathbb{Q} -forme linéaire sur \mathbb{R} associant, à tout $\xi \in \mathbb{R}$, sa coordonnée sur (x_{i_0}) dans la base (x_i) . Démontrer que φ n'est pas continue sur \mathbb{R} , que $\text{Ker}(\varphi)$ est partout dense dans \mathbb{R} , et que pour tout intervalle ouvert non vide U de \mathbb{R} , l'ensemble $\varphi(U)$ est partout dense dans \mathbb{R} .

§ IX.6 ENDOMORPHISMES. GROUPE LINÉAIRE

Soit E un K -ev. Rappelons (cf. théorème VI.2.4, et exemple 2 du § VI.4) que l'ensemble $\text{Hom}_K(E)$ des endomorphismes de E est muni de manière naturelle d'une structure de K -algèbre, dont l'ensemble des éléments inversibles, pour $E \neq \{0\}$, s'appelle *groupe linéaire* de E et se note $\text{GL}_K(E)$. Il est commode de convenir que si $E = \{0\}$, $\text{GL}_K(E) = \{\text{Id}_E\}$.

Soit E et F deux K -ev isomorphes, et $f : E \rightarrow F$ un isomorphisme de K -ev. L'application $\widehat{f} : \text{Hom}_K(E) \rightarrow \text{Hom}_K(F)$, $u \mapsto f \circ u \circ f^{-1}$, est alors un isomorphisme de K -algèbres, dont la réciproque est $\widehat{f^{-1}}$. Il en résulte que lorsque E est de dimension finie, $\text{Hom}_K(E)$ ne dépend, à isomorphisme près, que de la dimension de E .

Nous donnons ici quelques propriétés élémentaires de

$\text{Hom}_K(E)$ et du groupe linéaire $\text{GL}_K(E)$ lorsque E est un K -ev de dimension finie $n \geq 1$. Une conséquence immédiate du théorème IX.4.6 donne la

PROPOSITION IX.6.1

|| La K -algèbre $\text{Hom}_K(E)$ est un K -ev de dimension finie, et on a :
|| $\dim_K(\text{Hom}_K(E)) = n^2$.

On remarque en particulier que la dimension d'un K -ev du type $\text{Hom}_K(E)$ n'est pas arbitraire, et que si E est de dimension 1, alors $\text{Hom}_K(E)$ l'est aussi.

Le théorème IX.4.8 peut ici être précisé :

THÉORÈME IX.6.1

|| Si $u \in \text{Hom}_K(E)$, les conditions suivantes sont équivalentes :
|| (I) $u \in \text{GL}_K(E)$; (V) u est inversible à gauche ;
|| (II) u est injectif ; (VI) u est inversible à droite ;
|| (III) u est surjectif ; (VII) u est régulier à gauche ;
|| (IV) $\text{rg}(u) = n$; (VIII) u est régulier à droite.

Démonstration :

Les conditions (I) à (VI) sont équivalentes d'après le théorème IX.4.8. (I) \Rightarrow (VII) et (I) \Rightarrow (VIII) sont évidentes. Montrons que (VII) \Rightarrow (VI). Soit u régulier à gauche, élément de $\text{Hom}_K(E)$ dont on sait que c'est un K -ev de dimension finie. Par hypothèse l'application $v \mapsto u \circ v$, $\text{Hom}_K(E) \rightarrow \text{Hom}_K(E)$ est injective. D'autre part elle est K -linéaire ; donc elle est bijective ($\text{Hom}_K(E)$ de dimension finie). Il existe donc un $w \in \text{Hom}_K(E)$ tel que $u \circ w = \text{Id}_E$, ce qui prouve que u est inversible à droite. On prouve de la même manière que (VII) \Rightarrow (V). ■

Appliquons maintenant le théorème VI.3.2 pour étudier les *automorphismes* de E à l'aide des bases de E indexées par $\llbracket 1, n \rrbracket$, qu'on appelle aussi bases ordonnées de E .

THÉORÈME IX.6.2

|| Soit (e_1, e_2, \dots, e_n) et $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ deux bases de E . Il existe un, et un seul, élément $u \in \text{Hom}_K(E)$ tel que $u(e_i) = \varepsilon_i$ pour $1 \leq i \leq n$, et cet élément u appartient au groupe linéaire $\text{GL}_K(E)$.

Démonstration :

C'est une conséquence immédiate du théorème VI.3.2. ■

COROLLAIRE

|| Soit V et W deux sous- K -ev de même dimension de E . Il existe au moins un élément $u \in \text{GL}_K(E)$ tel que $u(V) = W$.

Démonstration :

Soit (e_1, \dots, e_p) et $(\varepsilon_1, \dots, \varepsilon_p)$ des bases respectives de V et W . Complétons-les en des bases (e_1, \dots, e_n) et $(\varepsilon_1, \dots, \varepsilon_n)$ de E et notons u l'automorphisme du K -ev E tel que $u(e_i) = \varepsilon_i$ pour $1 \leq i \leq n$. Alors u transforme la base (e_1, \dots, e_p) de V en la base $(\varepsilon_1, \dots, \varepsilon_p)$ de W , d'où $u(V) = W$ (Si $V = W = \{0\}$, le théorème est trivial). ■

Dans le langage de la théorie des groupes, le théorème IX.6.2 et son corollaire peuvent être traduits de la manière suivante : désignons par \mathfrak{B} l'ensemble de toutes les bases de E indexées par $\llbracket 1, n \rrbracket$. L'application de $\mathrm{GL}_K(E) \times \mathfrak{B}$ dans \mathfrak{B} , $(u, (e_1, \dots, e_n)) \mapsto (u(e_1), \dots, u(e_n))$ est une action à gauche de $\mathrm{GL}_K(E)$ sur \mathfrak{B} qui est *fidèle* et *transitive*. Désignons de même par $\mathcal{G}_p(E)$ l'ensemble de tous les sous- K -ev de E qui ont une dimension $p \in \llbracket 0, n \rrbracket$ fixée. L'application :

$$(u, V) \mapsto u(V), \quad \mathrm{GL}_K(E) \times \mathcal{G}_p(E) \rightarrow \mathcal{G}_p(E)$$

est une action à gauche de $\mathrm{GL}_K(E)$ sur l'ensemble $\mathcal{G}_p(E)$ qui est *transitive*, mais elle n'est évidemment pas fidèle car il y a en général beaucoup d'automorphismes de E qui laissent invariant un sous- K -ev donné.

Si E est de dimension 1, on a vu que $\dim_K(\mathrm{Hom}_K(E)) = 1$. Dans ce cas, l'application $K \rightarrow \mathrm{Hom}_K(E)$, $\lambda \mapsto \lambda \mathrm{Id}_E$ est un *isomorphisme* de K -algèbres, $\mathrm{Hom}_K(E) = K \mathrm{Id}_E$ est donc un corps commutatif isomorphe à K , et le groupe linéaire $\mathrm{GL}_K(E)$ s'identifie au groupe multiplicatif K^* , il est abélien.

Centre du groupe $\mathrm{GL}_K(E)$ pour $n \geq 2$

Ci-après, on suppose $n = \dim_K(E) \geq 2$.

Rappelons qu'on appelle **centre** d'un groupe G le sous-groupe de G constitué des éléments du groupe *permutables avec tous les autres* (cf § V.6, exemple 3).

THÉORÈME IX.6.3

|| Le centre du groupe linéaire $\mathrm{GL}_K(E)$, où E est un K -ev de dimension finie ≥ 2 , est le groupe des homothéties de E .

Démonstration :

Il est évident qu'une homothétie de E est permutable avec tout endomorphisme $u \in \mathrm{Hom}_K(E)$. Étudions la réciproque : soit $u \in \mathrm{GL}_K(E)$ tel que, déjà pour tout $v \in \mathrm{GL}_K(E)$, on ait $u \circ v = v \circ u$.

Montrons d'abord que u laisse globalement invariante toute droite vectorielle de E : si ce n'était pas le cas, il existerait $e_1 \in E$ et $e_2 \in E$ linéairement indépendants tels que $u(e_1) = e_2$. Une fois e_1 et e_2 ainsi choisis, on complète (e_1, e_2) en une base (e_1, e_2, \dots, e_n) de E . Considérons l'automorphisme v de E tel que $v(e_1) = e_1$, $v(e_2) = e_1 + e_2$, $v(e_i) = e_i$ si $i \geq 3$. Il s'agit bien d'un automorphisme car $(e_1, e_1 + e_2, e_3, \dots$

base de E . Alors on aurait $u \circ v(e_1) = e_2$ et $v \circ u(e_1) = e_1 + e_2 \neq e_2$ contrairement à l'hypothèse $u \circ v = v \circ u$, d'où l'assertion.

On déduit facilement de là que u est une homothétie. En effet, pour chaque $x \in E \setminus \{0\}$, ce qui précède montre qu'il existe un unique $\lambda_x \in K^*$ tel que $u(x) = \lambda_x \cdot x$. Si $x \in E \setminus \{0\}$ et $\rho \in K^*$, il est clair que $\lambda_{\rho x} = \lambda_x$. Si x et y sont linéairement indépendants dans E (ce qui entraîne $x + y \neq 0$) on a :

$$\begin{aligned} u(x + y) &= \lambda_{x+y}(x + y) = \lambda_{x+y} \cdot x + \lambda_{x+y} \cdot y = \\ &= u(x) + u(y) = \lambda_x \cdot x + \lambda_y \cdot y, \end{aligned}$$

d'où $\lambda_x = \lambda_y = \lambda_{x+y}$. Finalement tous les λ_x sont égaux à un même $\lambda \in K$, d'où $u = \lambda \text{Id}_E$. ■

Exercice 1 : Le corps de base K est fini, de cardinal q ; E est un K -ev de dimension $n \geq 1$.

a) Quel est le cardinal de $\text{GL}_K(E)$?

b) Soit $p \in \llbracket 0, n \rrbracket$ et $\mathcal{G}_p(E)$ l'ensemble des sous- K -ev de dimension p de E . En considérant l'opération à gauche de $\text{GL}_K(E)$ sur l'ensemble $\mathcal{G}_p(E)$, trouver $\text{card}(\mathcal{G}_p(E))$.

Exercice 2 : Soit E un K -ev et $u \in \text{Hom}_K(E)$. On note Φ_u (resp. Ψ_u) l'élément de $\text{Hom}_K(\text{Hom}_K(E))$ défini par $\Phi_u(v) = u \circ v$ (resp. $\Psi_u(v) = v \circ u$) pour $v \in \text{Hom}_K(E)$. Si E est de dimension finie $n \geq 1$, et si $\text{rg}(u) = r$, calculer $\text{rg}(\Phi_u)$ et $\text{rg}(\Psi_u)$ en fonction de n et r .

Exercice 3 : Soit E un K -ev de dimension finie $n \geq 1$. Trouver les parties G de $\text{Hom}_K(E) \setminus \text{GL}_K(E)$ qui sont des groupes pour la loi $(u, v) \mapsto u \circ v$.

Exercice 4 : Soit $K = \mathbb{Z}/2\mathbb{Z}$ et un K -ev E de dimension 2. En utilisant le théorème IX.6.2, montrer que $\text{GL}_K(E)$ est isomorphe à \mathfrak{S}_3 .

Exercice 5 : Soit E un K -ev de dimension finie $n \geq 2$. On note \mathcal{D} l'ensemble $\{u - v\}_{u \in \text{GL}_K(E), v \in \text{GL}_K(E)}$. Montrer que $\text{Hom}_K(E) = \text{GL}_K(E) \cup \mathcal{D}$.

Exercice 6 : Soit E un K -ev de dimension finie $n \geq 2$. On donne des entiers $d_1, d_2, \dots, d_p \geq 1$ tels que $d_1 + d_2 + \dots + d_p = n$. Soit $\mathcal{S}(d_1, \dots, d_p)$ l'ensemble des suites (V_1, \dots, V_p) de sous- K -ev de E telles que $\bigoplus_{i=1}^p V_i = E$. Etudier l'action à gauche naturelle du groupe $\text{GL}_K(E)$ sur $\mathcal{S}(d_1, \dots, d_p)$. Si de plus K est fini, de cardinal q , trouver le cardinal d'une orbite et d'un stabilisateur.

Exercice 7 : On donne un K -ev E de dimension finie $n \geq 2$ et des entiers d_0, d_1, \dots, d_p tels que $0 = d_0 < d_1 < \dots < d_p = n$. Soit $\mathcal{D}(d_0, \dots, d_p)$ l'ensemble des suites (W_0, \dots, W_p) de sous- K -ev de E telles que $\{0\} = W_0 \subset W_1 \subset \dots \subset W_p = E$.

a) Mêmes questions qu'à l'exercice 6, mais avec l'ensemble $\mathcal{D}(d_0, \dots, d_p)$.

b) On fixe (W_0, W_1, \dots, W_p) dans $\mathcal{D}(d_0, \dots, d_p)$. Soit \mathcal{S}_W l'ensemble des suites (V_0, \dots, V_p) de sous- K -ev de E telles que $\bigoplus_{i=0}^j V_i = W_j$ pour tout j . Mêmes questions qu'à l'exercice 6, mais avec l'ensemble \mathcal{S}_W .

Exercice 8 : (Idéaux de l'algèbre des endomorphismes de E en dimension finie).

On considère un K -ev E de dimension finie $n \geq 1$, et on note \mathcal{A} la K -algèbre $\text{Hom}_K(E)$.

a) Montrer que les seuls idéaux bilatères de \mathcal{A} sont $\{0\}$ et \mathcal{A} . Indication : si $b \neq \{0\}$ est un idéal bilatère, montrer qu'il contient un projecteur de rang 1, puis tous les projecteurs de rang 1.

b) Soit H un hyperplan de E . Prouver que l'ensemble des $u \in \mathcal{A}$ tels que $H \subset \text{Ker}(u)$ est un idéal à gauche de \mathcal{A} minimal pour l'inclusion dans l'ensemble des idéaux à gauche non nuls de \mathcal{A} . Réciproquement montrer que tout idéal à gauche non nul minimal (pour l'inclusion) de \mathcal{A} peut être défini de cette manière, donc qu'on a une bijection naturelle entre l'ensemble $\mathcal{G}_{n-1}(E)$ des hyperplans de E et l'ensemble $\mathcal{L}(E)$ des idéaux à gauche non nuls, minimaux pour l'inclusion, de E .

c) On fixe $\mathfrak{M} \in \mathcal{L}(E)$, associé à un hyperplan H .

c₁) Soit $\alpha \in \mathfrak{M} \setminus \{0\}$ et soit $x_0 \in E$ tel que $\alpha(x_0) \neq 0$. On note \mathfrak{N} l'ensemble des $u \in \mathcal{A}$ tels que $u(x_0) = 0$. Montrer que \mathfrak{N} est un idéal à gauche de \mathcal{A} , maximal pour l'inclusion dans l'ensemble des idéaux à gauche de \mathcal{A} distincts de \mathcal{A} (on dit en abrégé idéal à gauche maximal).

Etablir $\mathfrak{M} + \mathfrak{N} = \mathcal{A}$ et $\mathfrak{M} \cap \mathfrak{N} = \{0\}$.

c₂) Soit $\varphi : \mathfrak{M} \rightarrow E$, $u \mapsto u(x_0)$. Montrer que φ est un isomorphisme de groupes abéliens, et que : $(\forall u \in \mathcal{A}, \forall v \in \mathfrak{M}) \quad \varphi(uv) = u[\varphi(v)]$.

Exercice 9 : On conserve les notations et hypothèses de l'exercice 8.

Soit θ un automorphisme de l'anneau \mathcal{A} . On fixe $y_0 \in E$ tel que $[\theta(\alpha)](y_0) \neq 0$.

a) Soit $\mathfrak{N}_\theta = \{u \in \mathcal{A} \mid [\theta(u)](y_0) = 0\}$. Montrer que \mathfrak{N}_θ est un idéal à gauche maximal de \mathcal{A} , et que : $\mathfrak{M} \cap \mathfrak{N}_\theta = \{0\}$, $\mathfrak{M} + \mathfrak{N}_\theta = \mathcal{A}$.

b) Soit $\Phi_\theta : \mathcal{A} \rightarrow E$, $u \mapsto [\theta(u)](y_0)$. Montrer que $\varphi_\theta = \Phi_\theta|_{\mathfrak{M}}$ est un isomorphisme de groupes abéliens de \mathfrak{M} sur E et que : $\forall u \in \mathcal{A}, \forall v \in \mathfrak{M}, \varphi_\theta(uv) = \theta(u) \cdot [\varphi_\theta(v)]$.

c) En déduire qu'il existe un automorphisme γ du groupe additif E tel que, $\forall u \in \mathcal{A}$, $\theta(u) = \gamma \circ u \circ \gamma^{-1}$ (utiliser l'exercice 8c).

d) Soit f un endomorphisme du groupe additif de E . On suppose $f \circ u = u \circ f$ pour tout $u \in \mathcal{A}$. Montrer que f est une homothétie de E (f laisse stable toute droite vectorielle de E). En déduire que l'application $\sigma = \theta \Big|_{\mathcal{H}(E)}^{\mathcal{H}(E)}$ (où $\mathcal{H}(E)$ est le corps des homothéties de E) s'identifie à un automorphisme du corps K si $\mathcal{H}(E)$ est identifié à K . Puis montrer que l'application γ trouvée en c) est σ -linéaire, i.e. $(\forall x \in E, \forall \lambda \in K) \quad \gamma(\lambda x) = \sigma(\lambda) \gamma(x)$. Qu'en déduit-on si $K = \mathbb{R}$?

Exercice 10 : Soit E un K -ev de dimension finie $n \geq 1$ et u, v deux endomorphismes de E ayant exactement le même noyau H et la même image G de dimension 1. Cela entraîne-t-il qu'il existe $\alpha \in K^*$ tel que $u = \alpha v$?

§ IX.7 ÉLÉMENTS ALGÈBRIQUES D'UNE EXTENSION D'UN CORPS

Soit K un corps commutatif. Rappelons qu'on appelle *extension* de K tout corps L admettant K comme sous-corps (par exemple \mathbb{C} est une extension de \mathbb{R} , qui est une extension de \mathbb{Q}). Nous renvoyons aux § VII.4 et VII.7 pour les définitions et propriétés de base concernant la notion d'élément algébrique. Complétons cette étude avec :

THÉORÈME IX.7.1

|| Soit $x \in L$ un élément algébrique de degré d sur K .
 || (I) Le polynôme minimal F de x sur K est irréductible de

- (II) La K -algèbre $K[x]$ engendrée par x dans L est un sous-corps de L , et une base du K -ev $K[x]$ est $(1, x, x^2, \dots, x^{d-1})$.
- (III) Il existe un et un seul homomorphisme de K -algèbres entre les corps $M = K[X]/\alpha$ et $K[x]$ qui envoie \bar{X} sur x (α désignant l'idéal $FK[X]$ de $K[X]$) : c'est l'isomorphisme de K -ev ζ tel que $\zeta(\bar{X}^i) = x^i$ pour $0 \leq i \leq d-1$ (\bar{X} désignant l'image de X dans M) et c'est un isomorphisme de K -algèbres (donc aussi de corps).

Démonstration :

(I) Soit G_1 et G_2 dans $K[X]$ tels que $F = G_1 G_2$. On a :

$$0 = F(x) = G_1(x) G_2(x), \quad \text{d'où} \quad G_1(x) = 0 \quad \text{ou} \quad G_2(x) = 0$$

puisque L est un corps. Mais $G_i(x) = 0$ signifie que F divise G_i d'après la définition du polynôme minimal F ; donc G_1 ou G_2 est multiple de F , et par suite F est irréductible dans $K[X]$.

(II) La suite $(1, x, \dots, x^{d-1})$ engendre le K -ev $K[x]$, car si $G \in K[X]$, la division euclidienne de G par F donne : $G = FQ + R$ avec $R \in K[X]$ et $\deg(R) < d$. Prenant les valeurs en $x \in L$, il vient $G(x) = R(x)$, car $F(x) = 0$. Donc, si $R = r_0 + r_1 X + \dots + r_{d-1} X^{d-1}$, on a :

$$G(x) = r_0 + r_1 x + \dots + r_{d-1} x^{d-1} \in \text{Vect}_K(1, x, \dots, x^{d-1}).$$

Or, par définition, $K[x] = \{G(x)\}_{G \in K[X]}$, et on a donc prouvé que $K[x] \subset \text{Vect}(1, x, \dots, x^{d-1})$. L'inclusion réciproque étant évidente, on a bien $K[x] = \text{Vect}_K(1, x, \dots, x^{d-1})$. Le fait que la suite $(1, x, \dots, x^{d-1})$ est K -libre résulte du fait que F est minimal : si $\lambda_0, \dots, \lambda_{d-1}$ sont des éléments de K tels que $\lambda_0 + \lambda_1 x + \dots + \lambda_{d-1} x^{d-1} = 0$, le polynôme $H = \lambda_0 + \lambda_1 X + \dots + \lambda_{d-1} X^{d-1}$ vérifie $H(x) = 0$, d'où $H \equiv 0 \pmod{(F)}$, d'où $H = 0$ puisque $\deg(H) < d$, d'où la nullité des λ_i . Finalement $(1, x, \dots, x^{d-1})$ est une base du K -ev $K[x]$. Soit maintenant $y \in K[x] \setminus \{0\}$. L'application $K[x] \rightarrow K[x]$, $z \mapsto yz$ est K -linéaire et injective car l'anneau $K[x]$ est intègre en tant que sous-anneau d'un corps. Puisque $K[x]$ est un K -ev de dimension finie, il s'ensuit qu'elle est bijective (cf. théorème IX.6.1). En particulier, on a donc un $z \in K[x]$ tel que $yz = 1$, et y est donc inversible dans $K[x]$. Par suite $K[x]$ est un corps.

(III) Rappelons que $(1, \bar{X}, \bar{X}^2, \dots, \bar{X}^{d-1})$ est une base du K -ev M (cf. théorème VII.7.2). Soit $u : M \rightarrow K[x]$ un homomorphisme de K -algèbres tel que $u(\bar{X}) = x$. Alors $u(\bar{X}^i) = (u(\bar{X}))^i = x^i$ pour $0 \leq i \leq d-1$, donc u coïncide avec l'unique application K -linéaire $\zeta : M \rightarrow K[x]$ telle que $\zeta(\bar{X}^i) = x^i$ pour $0 \leq i \leq d-1$. Il reste à prouver que l'application ζ ainsi obtenue est un homomorphisme de K -algèbres, ce qui achèvera bien la démonstration puisqu'on sait déjà que ζ est bijective. On a d'abord $\zeta(1) = 1$. Soit G_1, G_2 dans $K[X]$, $G = G_1 G_2$, R_1 (resp. R_2) le reste de G_1 (resp. G_2) dans la division euclidienne par F , et R celui de G . On sait que : $F(\bar{X}) = 0_M$, $F(x) = 0$, et $R \equiv R_1 R_2 \pmod{(F)}$. On en déduit

$$G_1(\bar{X}) G_2(\bar{X}) = R_1(\bar{X}) R_2(\bar{X}) = R(\bar{X}),$$

$$\text{et aussi :} \quad G_1(x) G_2(x) = R_1(x) R_2(x) = R(x).$$

Mais $R_i(x) = \zeta[R_i(\bar{X})]$ et $R(x) = \zeta[R(\bar{X})]$, donc

$$\zeta(G_1(\bar{X}) G_2(\bar{X})) = \zeta(G_1(\bar{X})) \zeta(G_2(\bar{X})). \quad \blacksquare$$

Remarque 1 : Pour montrer que $K[x]$ est un corps, on aurait pu utiliser directement le *théorème de Bezout*. En effet, reprenons $y \in K[x] \setminus \{0\}$ et écrivons $y = \lambda_0 + \lambda_1 x + \dots + \lambda_{d-1} x^{d-1}$ avec les λ_i non tous nuls. Introduisons le polynôme $G = \lambda_0 + \lambda_1 X + \dots + \lambda_{d-1} X^{d-1} \in K[X]$. On a $G \in K_{d-1}[X] \setminus \{0\}$, et comme F est irréductible dans $K[X]$, G et F sont premiers entre eux. Soit U, V dans $K[X]$ tels que $UF + VG = 1$, en prenant les valeurs en $x \in L$, on en déduit : $V(x)G(x) = 1$ et comme $G(x) = y$ et que $V(x) \in K[x]$, on a bien trouvé un inverse à y dans $K[x]$ qui est bien un corps. Cette méthode offre l'avantage de fournir un calcul effectif de y^{-1} dans $K[x]$, par exemple en calculant le pgcd de F et G avec l'algorithme d'Euclide.

Extensions de degré fini

Soit L une extension du corps commutatif K . On peut alors regarder L comme un espace vectoriel sur K . Si ce K -ev est de dimension finie, c'est-à-dire s'il existe des éléments en nombre fini a_1, a_2, \dots, a_r dans L tels que tout élément de L puisse s'écrire sous la forme $x_1 a_1 + x_2 a_2 + \dots + x_r a_r$, avec des $x_i \in K$, on dit que L est une **extension de degré fini** de K . La dimension de L comme K -ev s'appelle alors le **degré** de L sur K , et se note $[L : K]$, et on appelle **base** de L sur K toute base de L considéré comme K -ev. En particulier lorsque $K = \mathbb{Q}$, les extensions de degré fini de K , sont, par définition, les **corps de nombres algébriques**.

Exemple 1 : Soit L une extension du corps commutatif K . Nous savons déjà que si $x \in L$ est algébrique sur K , alors le K -ev $K[x]$ est de dimension finie d . Nous allons voir que réciproquement si $x \in L$ et si le K -ev $K[x]$ est de dimension finie, alors x est algébrique sur K . En effet, soit n la dimension de ce K -ev. Il est impossible de trouver dans ce K -ev une famille libre de plus de n vecteurs (cf. théorème IX.4.1). Si donc on considère les éléments $1, x, x^2, \dots, x^n$ de $K[x]$, ils sont en nombre $n+1$ et donc liés par une relation de dépendance linéaire du type $\lambda_0 + \lambda_1 x + \dots + \lambda_n x^n = 0$, $\lambda_i \in K$ ($0 \leq i \leq n$) et x est donc algébrique sur K . On dit dans ce cas que $K[x]$ est une **extension algébrique simple** de K . Il est facile de vérifier en effet que, non seulement x , mais tout $y \in K[x]$ est encore algébrique sur K , car les éléments $1, y, y^2, \dots, y^n$ de $K[x]$ sont nécessairement liés à cause de la dimension finie du K -ev $K[x]$.

Exercice 1 : Soit \mathcal{A} une K -algèbre sans diviseur de zéro telle que le K -ev \mathcal{A} soit de dimension finie. Montrer que \mathcal{A} est une algèbre à division ; en particulier, si \mathcal{A} est de plus commutative, c'est un corps commutatif.

Exercice 2 : Soit L une extension du corps commutatif K , et soit $x \in L, y \in L$. On suppose x algébrique sur K , et y algébrique sur $K[x]$. Montrer que y est algébrique sur K et que le degré de y sur K est le produit du degré de y sur $K[x]$ par celui de x sur K (cf. théorème IX.4.5).

Exercice 3 : Soit L une extension du corps commutatif K , et x, y deux éléments de L algébriques sur K . Montrer que la K -algèbre $K[x, y]$ engendrée par x et y dans L est un sous-corps de L , et que c'est un K -ev de dimension finie. *Indication :* utiliser l'exercice 2.

Etendre par récurrence à un nombre fini x_1, x_2, \dots, x_n d'éléments de L algébriques sur K .

Exercice 4 : Soit L une extension du corps commutatif K . Prouver que l'ensemble \hat{K} de tous les éléments de L algébriques sur K est un sous-corps de L . Montrer que si L est algébriquement clos, \hat{K} l'est aussi.

Application. On prend $K = \mathbb{Q}$ et $L = \mathbb{C}$. Le corps $\widehat{\mathbb{Q}}$ n'est autre que l'ensemble des nombres algébriques. C'est donc un corps algébriquement clos. Vérifier que c'est un ensemble dénombrable.

Exercice 5 : Soit K un corps commutatif, L une extension finie commutative de K et M une extension finie de L . Montrer que M est extension finie de K et que

$$[M : K] = [M : L] \times [L : K].$$

Exercice 6 : Soit $\theta = \sqrt[3]{2} + \sqrt{3}$: θ est un nombre algébrique.

- Trouver le polynôme minimal de θ sur \mathbb{Q} .
- Démontrer que $\mathbb{Q}[\theta] = \mathbb{Q}[\sqrt[3]{2}, \sqrt{3}]$.
- Trouver les \mathbb{Q} -automorphismes (automorphismes de \mathbb{Q} -algèbre) du corps $\mathbb{Q}[\theta]$.
- Trouver tous les sous-corps de $\mathbb{Q}[\theta]$.

Exercice 7 : Soit K le corps $\mathbb{Q}[\sqrt[3]{2}, j]$ (où $j = \exp\left(\frac{2i\pi}{3}\right)$).

- Si $\theta = \sqrt[3]{2} + j\sqrt[3]{4}$, montrer que $K = \mathbb{Q}(\theta)$ et calculer le polynôme minimal de θ sur \mathbb{Q} .
- Prouver que le groupe des \mathbb{Q} -automorphismes de K est isomorphe à \mathfrak{S}_3 .
- Trouver tous les sous-corps de K .

Exercice 8 : (adjonction de racines)

Au lieu de se donner à l'avance une extension du corps commutatif K , on peut partir de la seule donnée de K et d'un polynôme irréductible $F \in K[X]$ et construire un corps $K[x]$ engendré par une racine « imaginaire » du polynôme F . C'est de cette façon qu'on a construit le corps \mathbb{C} des complexes à partir de \mathbb{R} en adjoignant une racine « imaginaire » i de $X^2 + 1$. Montrer que ce procédé est général. Montrer que si $K[x]$ et $K[y]$ sont des extensions algébriques simples du même corps K engendrées respectivement par des racines x et y du même polynôme F irréductible dans $K[X]$, alors il existe un seul isomorphisme de corps de $K[x]$ sur $K[y]$ tel que $x \mapsto y$ et qui conserve chaque élément de K .

Exemple : le polynôme $X^2 - X - 1$ est irréductible dans $F_3[X]$, où $F_3 = \mathbb{Z}/3\mathbb{Z}$. En déduire l'existence d'un corps à 9 éléments, où chaque élément est de la forme $a + bu$ où $a \in F_3$, $b \in F_3$ et $u^2 = u + 1$. Préciser les règles d'addition et de multiplication de ce corps noté F_9 ainsi que l'inverse de chaque élément. Montrer que tout élément d'un corps fini à 9 éléments est de degré ≤ 2 sur F_3 et en déduire que deux corps quelconques à 9 éléments sont isomorphes.

Exercice 9 : (corps finis)

a) Démontrer que le nombre d'éléments d'un corps fini de caractéristique p est une puissance de p .

b) Démontrer qu'il y a $\frac{p^2 - p}{2}$ polynômes normalisés irréductibles du second degré dans $\mathbb{Z}/p\mathbb{Z}[X]$ et en déduire que, pour tout p , il existe un corps à p^2 éléments.

c) Chercher le nombre de polynômes normalisés irréductibles du 3^e degré dans $\mathbb{Z}/p\mathbb{Z}[X]$.

Exercice 10 : Soit L un corps commutatif de caractéristique 0, et K un sous-corps de L . On suppose que $L = K(\theta)$, où θ est un élément algébrique de degré n sur K . Pour tout corps M intermédiaire entre K et L (c'est-à-dire tout sous-corps M de L tel que $K \subset M \subset L$) on note $[L : M] = \dim_M(L)$, et $\text{Gal}(L, M)$ l'ensemble des automorphismes de M -algèbre du corps L (c'est un groupe pour la composition des applications). On suppose que le polynôme minimal F_θ de θ sur K est dissocié dans $L[X]$.

a) Soit \mathcal{C} l'ensemble des racines de F_θ dans L . Montrer : pour tout $\xi \in \mathcal{C}$, il existe un unique $s \in \text{Gal}(L, K)$ tel que $s(\theta) = \xi$. En déduire : $\text{card}(\text{Gal}(L, K)) = n$.

b) Pour chaque sous-groupe G de $\text{Gal}(L, K)$, soit $\mathcal{F}_G = \{x \in L \mid (\forall s \in G) s(x) = x\}$. Vérifier que \mathcal{F}_G est un corps intermédiaire entre K et L , et qu'on a : $\mathcal{F}_{\text{Gal}(L, K)} = K$ (si $z \in \mathcal{F}_{\text{Gal}(L, K)}$, considérer $\varphi \in K[X]$ tel que $z = \varphi(\theta)$ et éti

c) Soit M un corps intermédiaire entre K et L . Montrer que le polynôme minimal P de θ sur M est dissocié dans $L[X]$. En déduire $\text{card}(\text{Gal}(L, M)) = [L : M]$.

d) Soit \mathcal{F} l'ensemble des corps intermédiaires entre K et L , soit \mathcal{G} l'ensemble des sous-groupes de $\text{Gal}(L, K)$, Φ l'application $\mathcal{F} \rightarrow \mathcal{G}$, $M \mapsto \text{Gal}(L, M)$ et $\Psi : \mathcal{G} \rightarrow \mathcal{F}$, $G \mapsto \mathcal{F}_G$. Montrer que Φ et Ψ sont des bijections réciproques l'une de l'autre [indication : si $G \in \mathcal{G}$, étudier le polynôme $\prod_{s \in G} (X - s(\theta))$].

e) Si $M \in \mathcal{F}$, montrer que les conditions (I) et (II) sont équivalentes :

(I) $\text{Gal}(L, M) \triangleleft \text{Gal}(L, K)$;

(II) Pour tout $s \in \text{Gal}(L, K)$ on a $s(M) = M$.

En supposant ces conditions remplies, prouver successivement :

e_1) il existe $\zeta \in M$ tel que $M = K[\zeta]$ et tel que le polynôme minimal de ζ sur K soit dissocié dans $M[X]$. Indication : prendre $\sigma_1, \sigma_2, \dots, \sigma_e \in \text{Gal}(L, K)$ représentant les e éléments du groupe quotient $\text{Gal}(L, K)/\text{Gal}(L, M)$; prouver que la fonction $h : M \rightarrow M$, $x \mapsto \prod_{i < j} (\sigma_i(x) - \sigma_j(x))$ est non nulle. Puis prouver que tout $\zeta \in M$ tel que $h(\zeta) \neq 0$ convient ;

e_2) les groupes $\text{Gal}(M, K)$ et $\text{Gal}(L, K)/\text{Gal}(L, M)$ sont isomorphes.

Exercice 11 : On appelle *quadratifrei* tout naturel non nul qui n'est divisible par le carré d'aucun nombre premier (i.e. du type $p_1 p_2 \dots p_r$ ou égal à 1). On donne $n \in \mathbb{N}^*$ et des quadratifrei a_1, a_2, \dots, a_n tous ≥ 2 et deux à deux premiers entre eux. Si $H \subset \llbracket 1, n \rrbracket$, on note $z_H = \prod_{i \in H} \sqrt{a_i}$, $K_H = \mathbb{Q}[(\sqrt{a_i})_{i \in H}]$. On convient que $z_\emptyset = 1$ et $K_{\llbracket 1, n \rrbracket} = L$.

a) Vérifier que chaque K_H est un sous-corps de \mathbb{R} , et que si $H \subset H' \subset \llbracket 1, n \rrbracket$, alors $K_H \subset K_{H'}$.

b) Montrer que la famille $(z_H)_{H \subset \llbracket 1, n \rrbracket}$ est \mathbb{Q} -linéairement indépendante (faire une récurrence sur n). En déduire $\dim_{\mathbb{Q}}(L) = 2^n$.

c) On munit l'ensemble $\mu_2 = \{-1, +1\}$ de sa structure de groupe multiplicatif. Soit E le groupe multiplicatif $\mathcal{F}(\llbracket 1, n \rrbracket, \mu_2)$. Pour $f \in E$, on note σ_f l'unique élément de $\text{GL}_{\mathbb{Q}}(L)$ tel que, pour tout $H \subset \llbracket 1, n \rrbracket$, on ait : $\sigma_f(z_H) = \left(\prod_{i \in H} f(i) \right) z_H$. Montrer que $f \mapsto \sigma_f$ définit un isomorphisme du groupe E sur le groupe $\text{Gal}(L, \mathbb{Q})$ (cf. exercice 10).

d) Pour $f \in E$, soit $\theta_f = \sum_{i=1}^n f(i) \sqrt{a_i}$. Montrer que le polynôme $\Phi(X) = \prod_{f \in E} (X - \theta_f)$ appartient à $\mathbb{Q}[X]$, et qu'il est irréductible dans $\mathbb{Q}[X]$. Quel est le polynôme minimal de chaque θ_f sur \mathbb{Q} ?

Dans la suite, on pose $\theta = \theta_I$, où I est le neutre de E . Calculer Φ pour $n = 3$ et pour $n = 4$.

e) Quels sont les sous-groupes de E ? En déduire le nombre des sous-corps de L .

Application : Expliquer pourquoi le nombre $1 + \sqrt{2} + \sqrt{3} + \sqrt{4} + \dots + \sqrt{N}$ est irrationnel.

Exercice 12 (Polynômes cyclotomiques ; leur irréductibilité sur \mathbb{Q}).

On définit la *fonction de Möbius* $\mu : \mathbb{N}^* \rightarrow \mathbb{Z}$ de la façon suivante : si $n \in \mathbb{N}$ est de la forme $n = p_1 p_2 \dots p_r$ avec les p_i premiers et distincts, alors $\mu(n) = (-1)^r$; $\mu(1) = 1$ et dans tous les autres cas $\mu(n) = 0$.

a) Comparer cette définition avec celle donnée dans l'exercice 13 du § VIII.5. Montrer que :

$$\sum_{d|n} \mu(d) = \begin{cases} 0 & \text{si } n > 1 \\ 1 & \text{si } n = 1 \end{cases}.$$

b) Soit G un groupe abélien noté multiplicativement et soit $f : \mathbb{N}^* \rightarrow G$ une application. On définit $g : \mathbb{N}^* \rightarrow G$ par

$$(1) \quad (\forall n \in \mathbb{N}^*) \quad g(n) = \prod_{d|n} \left[f\left(\frac{n}{d}\right) \right]^{\mu(d)}.$$

Montrer :

$$(2) \quad (\forall n \in \mathbb{N}^*) \quad f(n) = \prod_{d|n} g(d).$$

Inversement, on donne $g: \mathbb{N}^* \rightarrow G$ et on définit $f: \mathbb{N}^* \rightarrow G$ par (2). Montrer (1).

c) On reprend, pour $n \in \mathbb{N}^*$, le polynôme cyclotomique $\Phi_n(X) \in \mathbb{C}[X] \subset \mathbb{C}(X)$ défini dans l'exercice 16, § VII.4. Montrer, si $n > 1$:

$$\Phi_n(X) = \prod_{d|n} \left(X^{\frac{n}{d}} - 1 \right)^{\mu(d)} = \prod_{d|n} \left(\frac{X^{\frac{n}{d}} - 1}{X - 1} \right)^{\mu(d)}.$$

d) Soit ζ une racine primitive n -ième de 1 dans \mathbb{C} , avec $n > 1$. On note f le polynôme minimal de ζ sur \mathbb{Q} , qui divise donc Φ_n puisque $\Phi_n \in \mathbb{Z}[X]$.

d₁) Utiliser l'exercice 3 du § VII.3 pour prouver que $f \in \mathbb{Z}[X]$. On notera g le polynôme tel que $X^n - 1 = fg$. Vérifier que $g \in \mathbb{Z}[X]$.

d₂) Soit p un nombre premier qui ne divise pas n . Si $F \in \mathbb{Z}[X]$, soit $\bar{F} \in \mathbb{Z}/p\mathbb{Z}[X]$ le polynôme dont les coefficients sont les classes mod (p) de ceux de F . Montrer : si $g(\zeta^p) = 0$, alors on a : $g(X^p) = f(X)h(X)$ avec $h \in \mathbb{Z}[X]$.

En déduire : $\bar{f}(X)\bar{h}(X) = (\bar{g}(X))^p$, et \bar{f} et \bar{g} ne sont pas premiers entre eux dans $\mathbb{Z}/p\mathbb{Z}[X]$.

d₃) Montrer à partir de $\bar{f}(X)\bar{g}(X) = X^n - 1$ que cependant \bar{f} et \bar{g} devraient être premiers entre eux.

d₄) En déduire : toute racine primitive n -ième de 1 dans \mathbb{C} est racine de f , et par suite $f = \Phi_n(X)$.

En conclure que $\Phi_n(X)$ est irréductible dans $\mathbb{Q}[X]$ et aussi dans $\mathbb{Z}[X]$ (on pourra utiliser le théorème IX.7.1 (I)). (Méthode de Weber ⁽¹⁾, 1885.)

Application : Soit ζ une racine n -ième primitive de 1 dans \mathbb{C} . On pose $\gamma = \zeta + \frac{1}{\zeta}$. Quel est le degré de γ sur \mathbb{Q} ? (réponse : $\frac{1}{2} \varphi(n)$). Quel est son polynôme minimal sur \mathbb{Q} ? (on remarquera si θ est racine de Φ_n , $\frac{1}{\theta}$ aussi).

Exercice 13 (discriminant du polynôme cyclotomique).

On fixe $n \in \mathbb{N}^*$, $n > 2$. On note \mathcal{P}_n l'ensemble des racines primitives n -ièmes de 1 dans \mathbb{C} et par S_n le support premier de l'entier n . On se propose de calculer $\Delta_n = \prod_{1 \leq i < j \leq \varphi(n)} (t_i - t_j)^2$, où

$(t_k)_{1 \leq k \leq \varphi(n)}$ est un numérotage de \mathcal{P}_n . On note $\Phi_n(X)$ le polynôme cyclotomique $\prod_{k=1}^{\varphi(n)} (X - t_k)$.

a) Soit $D_n = \prod_{\zeta \in \mathcal{P}_n} \Phi_n'(\zeta)$. Calculer Δ_n en fonction de D_n .

b) Pour $I \subset S_n$, soit $P_I = \prod_{p \in I} p$, $N_I = \frac{n}{P_I}$ et $[I] = \text{card}(I)$.

Exprimer $\Phi_n'(X)$ à l'aide des $[I]$, des N_I et des P_I (cf. exercice 12b)). Montrer, si $\zeta \in \mathcal{P}_n$:

$$\Phi_n'(\zeta) = n\zeta^{n-1} \prod_{J \subset S_n, J \neq \emptyset} (\zeta^{N_J} - 1)^{((-1)^{|J|})}.$$

Montrer :

$$D_n = s_n n^{\varphi(n)} \prod_{J \subset S_n, J \neq \emptyset} \left\{ \prod_{\zeta \in \mathcal{P}_n} (\zeta^{N_J} - 1) \right\}^{((-1)^{|J|})}, \quad \text{avec } s_n = \prod_{\zeta \in \mathcal{P}_n} \zeta.$$

⁽¹⁾ Wilhelm Weber, mathématicien allemand (1804-1891), ami de Gauss.

c) Montrer : $\varphi(n) = \prod_{d|n} \frac{n}{d} \mu(d)$, où μ est la fonction de Möbius (cf. exercice 12). Calculer $\prod_{d|n} s_d$ et en déduire $s_n = (-1)^{\varphi(n)}$.

d) Soit $J \subset S_n$, $J \neq \emptyset$. On pose $H_J = \prod_{\zeta \in \mathcal{P}_n} (\zeta^{N_J} - 1)$ et $E_J = \frac{\varphi(n)}{\varphi(P_J)}$.

d₁) Montrer $H_J = [\Phi_{P_J}(1)]^{E_J} \times (-1)^{\varphi(n)}$ (observer que ζ^{N_J} est une racine primitive P_J -ième de 1 et calculer combien de fois on trouve ainsi ces racines).

d₂) Pour $d|n$, soit $g(d) = \Phi_d(1)$. Montrer : $g(1) = \prod_{m|d} \left(\frac{d}{m}\right)^{\mu(n)}$ (cf. exercice 12). En déduire :

$$g(1) = 0, \quad g(d) = \begin{cases} p & \text{si } d > 1 \text{ et } d \text{ de la forme } p^\alpha, \alpha \in \mathbb{N}^* \\ 1 & \text{si } d > 1 \text{ et } d \text{ n'est pas une puissance de nombre premier.} \end{cases}$$

e) Montrer, si $J \subset S_n$, $J \neq \emptyset$: $H_J = (-1)^{\varphi(n)}$ si $\text{card}(J) \geq 2$, et

$$H_J = p^{\varphi(n)/p-1} \times (-1)^{\varphi(n)} \quad \text{si } J = \{p\}.$$

f) Etablir la « formule du discriminant » : $\Delta_n = (-1)^{\frac{1}{2}\varphi(n)} \times \frac{n^{\varphi(n)}}{\prod_{p \in S_n} p^{p-1}}$.

Exercice 14 : Soit $\zeta = e^{2i\pi/5}$ et K le corps $\mathbb{Q}[\zeta]$.

On note A l'ensemble des nombres $a_0 + a_1 \zeta + a_2 \zeta^2 + a_3 \zeta^3$ avec $a_i \in \mathbb{Z}$ ($i \in \llbracket 0, 3 \rrbracket$).

a) Montrer que $\Phi(X) = X^4 + X^3 + X^2 + X + 1$ est irréductible dans $\mathbb{Q}[X]$. C'est donc le polynôme minimal de ζ sur \mathbb{Q} , et $(1, \zeta, \zeta^2, \zeta^3)$ est une base du \mathbb{Q} -ev K .

b) Soit Γ le groupe des automorphismes du corps K . Montrer : $\forall k \in \llbracket 1, 4 \rrbracket$, il existe un unique $\sigma \in \Gamma$ tel que $\sigma(\zeta) = \zeta^k$. En déduire : Γ est un groupe cyclique à 4 éléments, engendré par s tel que $s(\zeta) = \zeta^2$. Prouver : $\mathbb{Q} = \{x \in K \mid \forall \sigma \in \Gamma, \sigma(x) = x\}$.

c) Si $z \in K$ on pose : $N(z) = \prod_{\sigma \in \Gamma} \sigma(z)$ et $\text{Tr}(z) = \sum_{\sigma \in \Gamma} \sigma(z)$. Prouver que N est à valeurs

dans \mathbb{Q}_+ et que Tr est à valeurs dans \mathbb{Q} . Prouver que Tr est une forme linéaire sur le \mathbb{Q} -ev K , que $N(z_1 z_2) = N(z_1) N(z_2)$ pour $z_i \in K$ [si $z \in K$, on pourra prouver aussi : $N(z) = \det(z^*)$ et $\text{Tr}(z) = \text{trace de } z^*$, où $z^* \in \text{Hom}_{\mathbb{Q}}(K)$ est défini par $z^*(x) = zx$ si $x \in K$, dès qu'on connaîtra les notions de déterminant et de trace d'un endomorphisme].

d) Montrer que A est un sous-anneau de K ; prouver que si $x \in A$, le polynôme minimal de x sur \mathbb{Q} appartient à $\mathbb{Z}[X]$; prouver $A \cap \mathbb{Z} = \mathbb{Z}$.

e) Soit \mathfrak{p} l'idéal $(1 - \zeta)A$ de A . Prouver : $5\mathbb{Z} \subset \mathfrak{p} \cap \mathbb{Z}$. A partir de $\frac{1}{5} \notin A$ (à vérifier !) montrer que $\mathfrak{p} \cap \mathbb{Z} = 5\mathbb{Z}$.

Soit $x = a_0 + a_1 \zeta + a_2 \zeta^2 + a_3 \zeta^3 \in K$ ($a_i \in \mathbb{Q}$). On suppose que le polynôme minimal de x sur \mathbb{Q} appartient à $\mathbb{Z}[X]$. Calculer $\text{Tr}(x(1 - \zeta))$ en fonction des a_i . En déduire : $a_0 \in \mathbb{Z}$. Prouver de proche en proche que tous les a_i sont dans \mathbb{Z} . En déduire : $x \in A$.

Exercice 15 (suite du précédent).

On se propose de montrer que l'ensemble A des nombres complexes $a_0 + a_1 \zeta + a_2 \zeta^2 + a_3 \zeta^3$ (où $\zeta = e^{2i\pi/5}$ et les a_i sont dans \mathbb{Z}) constitue un *anneau principal*. De manière précise on va prouver que la fonction N définie dans l'exercice 14 définit un stathme euclidien sur A (méthode de Landau ⁽¹⁾). Prouver qu'il suffit pour cela d'établir :

$$(\mathcal{P}_1) \quad (\forall \alpha \in K, \exists \beta \in A \mid N(\alpha - \beta) < 1).$$

On fixe donc $\alpha \in K$, $\alpha = a_0 + a_1 \zeta + a_2 \zeta^2 + a_3 \zeta^3$ ($a_i \in \mathbb{Q}$ pour tout i). On note b_i la partie entière de a_i et $c_i = a_i - b_i$ pour $i \in \llbracket 0, 3 \rrbracket$.

⁽¹⁾ Edmund Georg Hermann Landau, mathématicien allemand (1877-1938)

a) Prouver que l'une au moins des propriétés suivantes est vraie :

$$(1) \quad \exists i \in \llbracket 0, 3 \rrbracket, \quad c_i \leq \frac{1}{5};$$

$$(2) \quad \exists i \in \llbracket 0, 3 \rrbracket, \quad 1 - c_i \leq \frac{1}{5};$$

$$(3) \quad \exists (i, j) \in \llbracket 0, 3 \rrbracket^2, \quad i \neq j \quad \text{et} \quad 0 \leq c_j - c_i \leq \frac{1}{5}.$$

b) Ecrire $\alpha \zeta$, $\alpha \zeta^2$, $\alpha \zeta^3$ et $\alpha \zeta^4$ dans la base $\mathcal{B} = (1, \zeta, \zeta^2, \zeta^3)$ du \mathbb{Q} -ev K . Montrer : $\exists k \in \llbracket 1, 4 \rrbracket$ tel que $\alpha \zeta^k$ vérifie la condition suivante : si $\alpha \zeta^k = A_0 + A_1 \zeta + A_2 \zeta^2 + A_3 \zeta^3$ ($\forall_i \in \mathbb{Q}$), il existe B_0, B_1, B_2, B_3 dans \mathbb{Z} et $i \in \llbracket 0, 3 \rrbracket$ tels que $|A_i - B_i| \leq \frac{1}{5}$ et $|A_j - B_j| \leq \frac{1}{2}$ pour $j \neq i$. Dans la suite on fixe un tel k et de tels B_i , et on pose

$$\gamma = \sum_{\lambda=0}^3 B_\lambda \zeta^\lambda, \quad \text{et} \quad \gamma_1 = \alpha \zeta^k - \gamma = m_0 + m_1 \zeta + m_2 \zeta^2 + m_3 \zeta^3 \quad (m_i \in \mathbb{Q}).$$

c) Pour $j \in \llbracket 1, 4 \rrbracket$, soit $\gamma_j = m_0 + m_1 \zeta^j + m_2 \zeta^{2j} + m_3 \zeta^{3j}$. Démontrer :

$$\gamma_1 \gamma_4 = S + T \sqrt{5}, \quad \text{où } T \in \mathbb{Q} \quad \text{et} \quad S = \left(\sum_{i=0}^3 m_i^2 \right) - \frac{1}{2} \left(\sum_{0 \leq j < l \leq 3} m_j m_l \right).$$

Exprimer $\gamma_2 \gamma_3$ à l'aide de S et T , puis exprimer $N(\gamma_1)$ en fonction de S et T . En déduire : $N(\gamma_1) \leq S^2$.

d) Majorer $\sum_{i=0}^3 m_i^2$. Montrer : $-1 < S < 1$, et achever de prouver (\mathcal{P}_1) .

Chapitre X

FONCTIONS POLYNOMIALES SUR K^n ; ÉQUATIONS ALGÈBRIQUES

§ X.1 POLYNÔMES A n LETTRES

Dans les § X.1 et X.2, K désigne un corps commutatif.

Monômes

Soit \mathcal{A} une K -algèbre. Nous désignerons par a une famille finie non vide $(a_i)_{i \in I}$ d'éléments de \mathcal{A} . Nous avons vu au § VI.4 que la K -algèbre \mathcal{B} engendrée par les (a_i) est le K -ev engendré dans \mathcal{A} par l'élément neutre e et les produits $a_{i_1} a_{i_2} \dots a_{i_n}$, où n parcourt \mathbb{N}^* et où (i_1, i_2, \dots, i_n) est une suite d'éléments de I . En particulier, supposons \mathcal{A} **commutative** : \mathcal{B} est le K -ev engendré par tous les éléments $\prod_{i \in I} a_i^{\alpha_i}$, où $\alpha_i \in \mathbb{N}$ pour tout $i \in I$.

Ainsi s'introduit tout naturellement le **monoïde** \mathbb{N}^I (voir § II.4). Soit $\alpha = (\alpha_i)_{i \in I}$ un élément de \mathbb{N}^I ; associons-lui l'élément $\mathcal{M}_\alpha(a) = \prod_{i \in I} a_i^{\alpha_i}$ de

\mathcal{A} , appelé **monôme de multi-indice α** (en abrégé : d'indice α) **des (a_i)** . Les règles de calcul dans une K -algèbre commutative montrent immédiatement la loi fondamentale suivante :

$$(1) (\forall \alpha = (\alpha_i) \in \mathbb{N}^I, \forall \beta = (\beta_i) \in \mathbb{N}^I) \quad \mathcal{M}_{\alpha + \beta}(a) = \mathcal{M}_\alpha(a) \mathcal{M}_\beta(a).$$

Remarque 1 : Ce qui précède reste valable sous la seule hypothèse que les (a_i) sont *deux à deux permutable* dans la K -algèbre \mathcal{A} . Cette remarque présente un grand intérêt lorsqu'on travaille dans l'algèbre des endomorphismes d'un K -ev.

DÉFINITION X.1.1

⎧ Soit n un entier ≥ 1 ; notons $\varphi_1, \varphi_2, \dots, \varphi_n$ les projections naturelles
⎧ de K^n dans K . La sous- K -algèbre engendrée par $(\varphi_1, \varphi_2, \dots, \varphi_n)$
⎧ dans l'algèbre de fonctions $\mathcal{F}(K^n, K)$ s'appelle **K -algèbre des**
⎧ **fonctions polynomiales sur K^n** .

Cette K -algèbre, que nous noterons $K[\varphi_1, \varphi_2, \dots, \varphi_n]$ (ce qui sous-entend que les φ_i sont les projections naturelles de K^n dans K), est donc le K -ev engendré par les monômes $\mathcal{M}_\alpha(\varphi_1, \varphi_2, \dots, \varphi_n) = \varphi_1^{\alpha_1} \varphi_2^{\alpha_2} \dots \varphi_n^{\alpha_n}$ lorsque $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ décrit \mathbb{N}^n . Pour $n = 1$, on retrouve la K -algèbre des fonctions polynomiales sur K (cf. § VII.4).

Exemple 1 : Une *forme linéaire* sur le K -ev K^n est une fonction polynomiale ; car dans la base canonique, c'est une fonction du type $\lambda_1 \varphi_1 + \dots + \lambda_n \varphi_n$, $\lambda_i \in K$. Le K -ev engendré par $\varphi_1, \dots, \varphi_n$ n'est autre que le *dual* $(K^n)^*$.

Indépendance algébrique

DÉFINITION X.1.2

Soit \mathcal{A} une K -algèbre commutative et $(a_i)_{i \in I} = a$ une famille finie non vide dans \mathcal{A} . On dit que les (a_i) sont **algébriquement libres sur K** ssi la famille des monômes $(\mathcal{M}_\alpha(a))_{\alpha \in \mathbb{N}^I}$ est **linéairement indépendante sur K** .

Cela signifie que les relations : $(\forall \alpha \in \mathbb{N}^I) \lambda_\alpha \in K$, la famille (λ_α) est à support fini et $\sum_{\alpha \in \mathbb{N}^I} \lambda_\alpha \mathcal{M}_\alpha(a) = 0$, entraînent : $(\forall \alpha \in \mathbb{N}^I) \lambda_\alpha = 0$. Si les

(a_i) sont *algébriquement libres* et engendrent la K -algèbre \mathcal{A} , nous savons que par définition $(\mathcal{M}_\alpha(a))_{\alpha \in \mathbb{N}^I}$ est une *base* du K -ev \mathcal{A} .

THÉORÈME X.1.1

Soit une K -algèbre commutative \mathcal{A} engendrée par une famille finie non vide $a = (a_i)_{i \in I}$ **algébriquement libre** sur K . Pour toute K -algèbre \mathcal{B} , pour toute famille $b = (b_i)_{i \in I}$ d'éléments deux à deux permutables de \mathcal{B} , il existe un, et un seul, homomorphisme de K -algèbres $f : \mathcal{A} \rightarrow \mathcal{B}$ tel que $f(a_i) = b_i$ pour tout i .

Démonstration :

Si f existe, pour tout multi-indice $\alpha = (\alpha_i)_{i \in I}$ dans \mathbb{N}^I , on a : $f(\mathcal{M}_\alpha(a)) = \prod_{i \in I} (f(a_i))^{\alpha_i} = \mathcal{M}_\alpha(b)$ puisque les b_i sont deux à deux permutables. Or les $(\mathcal{M}_\alpha(a))_{\alpha \in \mathbb{N}^I}$ forment, par hypothèse une *base* du K -ev \mathcal{A} . Donc f ne peut être que l'unique application K -linéaire $g : \mathcal{A} \rightarrow \mathcal{B}$ telle que $g(\mathcal{M}_\alpha(a)) = \mathcal{M}_\alpha(b)$ pour tout $\alpha \in \mathbb{N}^I$ (cf. théorème VI.3.2).

Réciproquement, soit f cette unique application K -linéaire. Pour prouver que f est un homomorphisme de K -algèbres, il suffit de prouver que $f(xy) = f(x)f(y)$ si $x \in \mathcal{A}$ et $y \in \mathcal{A}$, puisque $f(e_{\mathcal{A}}) = \prod_{i \in I} b_i^0 = e_{\mathcal{B}}$. Or, si

x et y sont des monômes en a : $x = \mathcal{M}_\alpha(a)$, $y = \mathcal{M}_\beta(a)$, on a d'après (1) :

$$f(x) f(y) = \mathcal{M}_\alpha(b) \mathcal{M}_\beta(b) = \mathcal{M}_{\alpha+\beta}(b) = f(\mathcal{M}_{\alpha+\beta}(a)) = f(xy)$$

et on étend cette propriété à $f(x) f(y) = f(xy)$ pour $x \in \mathcal{A}$ et $y \in \mathcal{A}$ par K -linéarité. ■

DÉFINITION X.1.3

⎧ Avec les notations et hypothèses du théorème X.1.1, l'homomor-
 ⎧ phisme de K -algèbres $f: \mathcal{A} \rightarrow \mathcal{B}$ tel que $f(a_i) = b_i$ pour tout
 ⎧ $i \in I$ est dit obtenu **par substitution** des b_i aux a_i dans f .

Cette définition se justifie ainsi : un élément $x \in \mathcal{A}$ s'écrit de manière unique $x = \sum_{\alpha \in \mathbb{N}^I} \lambda_\alpha \mathcal{M}_\alpha(a)$, où (λ_α) est une famille à support fini dans K ,

puisque $(\mathcal{M}_\alpha(a))$ est une base du K -ev \mathcal{A} . Alors $f(x) = \sum_{\alpha \in \mathbb{N}^I} \lambda_\alpha \mathcal{M}_\alpha(b)$,

c'est-à-dire que $f(x)$ s'obtient à partir de l'expression de x en remplaçant partout les lettres (a_i) par les lettres (b_i) de même indice.

Notons qu'en application du théorème VI.3.2, f est *injectif* ssi les (b_i) sont algébriquement libres sur K ; f est *surjectif* ssi les (b_i) engendrent la K -algèbre \mathcal{B} (qui est donc alors commutative) ; et f est un *isomorphisme de K -algèbres* ssi les (b_i) sont algébriquement libres sur K et engendrent la K -algèbre \mathcal{B} . En particulier, *deux K -algèbres commutatives engendrées par un même nombre fini n d'éléments algébriquement libres sur K sont toujours isomorphes*. A cause de cela, on peut n'étudier de telles algèbres que lorsque les éléments en question sont *ordonnés en une suite finie* (a_1, a_2, \dots, a_n) .

DÉFINITION X.1.4

⎧ Soit $n \in \mathbb{N}^*$. On appelle **K -algèbre de polynômes à n lettres**
 ⎧ **a_1, a_2, \dots, a_n** , toute K -algèbre \mathcal{A} **commutative**, engendrée par n
 ⎧ **éléments (a_1, a_2, \dots, a_n) algébriquement libres sur K** . (On dit aussi
 ⎧ **que \mathcal{A} est une K -algèbre commutative libre en les lettres (a_i)**)).

Comme on l'a vu, une telle K -algèbre notée $K[a_1, a_2, \dots, a_n]$ se caractérise par son « pouvoir de substitution » qui résulte du théorème X.1.1. Dans l'expression $x = \sum_{\alpha \in \mathbb{N}^n} \lambda_\alpha \mathcal{M}_\alpha(a) = \sum_{\substack{\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \\ \alpha \in \mathbb{N}^n}} \lambda_\alpha a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}$ d'un élé-

ment $x \in K[a_1, a_2, \dots, a_n]$, les lettres (a_i) se comportent donc comme l'indéterminée de la K -algèbre des polynômes à une indéterminée sur K , d'où le vocabulaire « algèbre de polynômes en les lettres a_i ».

Soit $\mathcal{A} = K[a_1, a_2, \dots, a_n]$ une K -algèbre de polynômes en

(a_i) , et soit $x \in \mathcal{A}$, $y \in \mathcal{A}$:

$$x = \sum_{\alpha \in \mathbb{N}^n} \lambda_{\alpha} \mathcal{M}_{\alpha}(a), \quad y = \sum_{\beta \in \mathbb{N}^n} \lambda_{\beta} \mathcal{M}_{\beta}(a).$$

Alors, tenant compte de (1) et du principe de Fubini (cf. § III.1 et § III.2), on a :

$$(2) \quad xy = \sum_{\gamma \in \mathbb{N}^n} c_{\gamma} \mathcal{M}_{\gamma}(a), \quad \text{où} \quad c_{\gamma} = \sum_{\alpha + \beta = \gamma} a_{\alpha} b_{\beta} \quad \text{pour tout } \gamma \in \mathbb{N}^n.$$

La relation (2) est l'expression du produit dans $K[a_1, a_2, \dots, a_n]$.

Indépendance algébrique des projections de K^n dans K si K est un corps infini

THÉORÈME X.1.2

Soit $\varphi_1, \varphi_2, \dots, \varphi_n$ les projections canoniques de K^n dans K , et soit J une partie finie non vide de \mathbb{N}^n . Si $(a_{\alpha})_{\alpha \in J}$ est une famille de scalaires telle que $f = \sum_{\alpha \in J} a_{\alpha} \varphi_1^{\alpha_1} \varphi_2^{\alpha_2} \dots \varphi_n^{\alpha_n}$ prenne la valeur 0 sur $E_1 \times E_2 \times \dots \times E_n$, où pour chaque $i \in \llbracket 1, n \rrbracket$, E_i est une partie de K de cardinal $> \max_{\alpha \in J, \alpha = (\alpha_1, \dots, \alpha_n)} (\alpha_i)$, alors tous les a_{α} sont nuls.

Démonstration :

Si $n = 1$ on reconnaît le théorème VII.4.2 (dans un K -espace vectoriel de dimension d , le nombre d'éléments d'une partie libre est $\leq d$). Supposons le théorème vrai à l'ordre $n - 1 \geq 1$. Soit J_n l'ensemble $\{\alpha_n, \alpha = (\alpha_i) \in J\}$; alors

$$f = \sum_{k \in J_n} A_k \varphi_n^k, \quad \text{avec} \quad A_k = \sum_{\alpha \in J \mid \alpha_n = k} a_{\alpha} \varphi_1^{\alpha_1} \dots \varphi_{n-1}^{\alpha_{n-1}}.$$

Fixons $(x_1, \dots, x_{n-1}) \in E_1 \times E_2 \times \dots \times E_{n-1}$ et posons :

$$\tilde{A}_k = \sum_{\alpha \in J \mid \alpha_n = k} a_{\alpha} \varphi_1^{\alpha_1} \dots \varphi_{n-1}^{\alpha_{n-1}}(x_1, \dots, x_{n-1}) = \sum_{\alpha \in J \mid \alpha_n = k} a_{\alpha} x_1^{\alpha_1} \dots x_{n-1}^{\alpha_{n-1}}.$$

La fonction $z \mapsto g(z) = \sum_{k \in J_n} \tilde{A}_k z^k$, $K \rightarrow K$ s'annule pour $z \in E_n$, car

$g(z) = f(x_1, \dots, x_{n-1}, z)$ pour tout z . Mais $\text{card}(E_n) > \max(J_n)$ par hypothèse, donc d'après le théorème VII.4.2, tous les \tilde{A}_k sont nuls. C'est vrai pour tout $(x_1, \dots, x_{n-1}) \in E_1 \times E_2 \times \dots \times E_{n-1}$. L'hypothèse de récurrence montre alors que, pour chaque $k \in J_n$, les coefficients $(a_{\alpha})_{\alpha \in J \mid \alpha_n = k}$ sont tous nuls. Finalement tous les a_{α} sont nuls. ■

COROLLAIRE 1

|| Si le corps K est **infini**, les projections $\varphi_1, \dots, \varphi_n$ de K^n dans K sont **algébriquement libres** sur K ; autrement dit la famille des monômes $\mathcal{M}_\alpha(\varphi) = \varphi_1^{\alpha_1} \varphi_2^{\alpha_2} \dots \varphi_n^{\alpha_n}$ ($\alpha \in \mathbb{N}^n$) forme alors une **base du K -ev** des fonctions polynomiales sur K^n .

En conséquence, cette K -algèbre $K[\varphi_1, \varphi_2, \dots, \varphi_n]$ est une K -algèbre de polynômes en les n lettres $\varphi_1, \varphi_2, \dots, \varphi_n$.

Considérons alors une partie E de K^n ; une fonction $f : E \rightarrow K$ sera dite **polynomiale** ssi il existe $g \in K[\varphi_1, \varphi_2, \dots, \varphi_n]$ telle que $g|_E = f$. Les fonctions polynomiales sur E forment une sous- K -algèbre de l'algèbre des fonctions de E dans K , qui est l'image par l'homomorphisme de K -algèbres $g \mapsto g|_E$ de $K[\varphi_1, \varphi_2, \dots, \varphi_n]$. On a alors, en revenant au théorème X.1.2 :

COROLLAIRE 2

|| Si le corps K est **infini**, et si la partie E de K^n contient un ensemble $E_1 \times E_2 \times \dots \times E_n$ avec **chaque E_i infini**, l'application de restriction $g \mapsto g|_E$ définit un **isomorphisme** de l'algèbre $K[\varphi_1, \varphi_2, \dots, \varphi_n]$ sur l'algèbre des fonctions polynomiales sur E .

COROLLAIRE 3

|| Pour chaque $n \in \mathbb{N}^*$, il existe au moins une K -algèbre de polynômes à n lettres.

Démonstration :

Pour un corps K infini, c'est une conséquence du corollaire 1. Supposons K fini. Alors K admet au moins une extension infinie, ne serait-ce que le corps $L = K(X)$ des fractions rationnelles à une indéterminée X sur K . Soit $\varphi_1, \dots, \varphi_n$ les projections naturelles de L^n dans L . La L -algèbre $L[\varphi_1, \varphi_2, \dots, \varphi_n]$ devient, par restriction des scalaires à K , une K -algèbre ; et puisque $(\varphi_1, \varphi_2, \dots, \varphi_n)$ sont algébriquement libres sur L , ils le sont *a fortiori* sur K . Donc la sous- K -algèbre $K[\varphi_1, \varphi_2, \dots, \varphi_n]$ de $L[\varphi_1, \varphi_2, \dots, \varphi_n]$ est une algèbre de polynômes en les (φ_i) . ■

Degré et valuation

On fixe $n \in \mathbb{N}^*$. Pour $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$ posons $\|\alpha\| = \alpha_1 + \alpha_2 + \dots + \alpha_n$. Il est clair que $\|\alpha + \beta\| = \|\alpha\| + \|\beta\|$ pour tous α et β dans \mathbb{N}^n .

Considérons alors une K -algèbre de polynômes à n lettres : $K[X_1, X_2, \dots, X_n]$. Comme au § VII.1 nous utiliserons l'ensemble $\overline{\mathbb{N}} = \mathbb{N} \cup \{-\infty, +\infty\}$.

DÉFINITION X.1.5

Soit $A = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n} \in K[X_1, X_2, \dots, X_n]$. Si $A = 0$, on pose $\text{val}(A) = +\infty$ et $\text{deg}(A) = -\infty$. Si $A \neq 0$, on appelle **degré de A** (resp. **valuation de A**) et on note $\text{deg}(A)$ (resp. $\text{val}(A)$) l'entier $\text{Max} \{ \|\alpha\| \mid \alpha \in \mathbb{N}^n \text{ et } a_\alpha \neq 0 \}$ (resp. $\text{Min} \{ \|\alpha\| \mid \alpha \in \mathbb{N}^n \text{ et } a_\alpha \neq 0 \}$).

On vérifie aisément, quels que soient A et B dans $K[X_1, \dots, X_n]$ les propriétés :

- $\text{deg}(A + B) \leq \text{Max}(\text{deg}(A), \text{deg}(B))$, et si $\text{deg}(A) \neq \text{deg}(B)$,

il y a égalité

- $\text{val}(A + B) \geq \text{Min}(\text{val}(A), \text{val}(B))$, et si $\text{val}(A) \neq \text{val}(B)$,

il y a égalité.

DÉFINITION X.1.6

Un polynôme $A \in K[X_1, X_2, \dots, X_n]$ est dit **homogène de degré d** (où $d \in \mathbb{N}$ est donné) ssi $A = 0$ ou $\text{deg}(A) = \text{val}(A) = d$. (On considère ainsi le polynôme nul comme homogène de degré arbitraire.)

Soit \mathcal{H}_d l'ensemble des polynômes homogènes de degré donné d . Dire que $A = \sum a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}$ appartient à \mathcal{H}_d signifie :

$$(\forall \alpha \in \mathbb{N}^n) \quad a_\alpha \neq 0 \Rightarrow \|\alpha\| = d.$$

En d'autres termes les seuls monômes à coefficients non nuls présents dans A sont de degré d . Il en résulte que \mathcal{H}_d est le sous- K -ev de $K[X_1, \dots, X_n]$ engendré par les $(X_1^{\alpha_1} \dots X_n^{\alpha_n})_{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n, \|\alpha\| = d}$. Ces monômes, étant linéairement indépendants, forment une **base du K -ev \mathcal{H}_d** . Leur nombre est $\binom{n+d-1}{d}$ d'après le théorème III.4.5. Ainsi \mathcal{H}_d est un sous- K -ev de dimension finie de $K[X_1, \dots, X_n]$, et $\dim_K(\mathcal{H}_d) = \binom{n+d-1}{d}$. Si nous désignons par $\mathcal{S}_{n,d}$ l'ensemble

$$\{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \mid \alpha_1 + \dots + \alpha_n = d\},$$

alors la famille $(\mathcal{S}_{n,d})_{d \in \mathbb{N}}$ est un partage de \mathbb{N}^n . Il en résulte aussitôt (cf. § IX.2, exemple 2) que : $K[X_1, \dots, X_n] = \bigoplus_{d \in \mathbb{N}} \mathcal{H}_d$.

Le nom de polynômes homogènes donné aux éléments de \mathcal{H}_d est à rapprocher de celui des fonctions homogènes étudiées en Analyse. En effet, si $A \in \mathcal{H}_d$ et si les x_i sont éléments de K ainsi que λ , on a bien $A(\lambda x_1, \dots, \lambda x_n) = \lambda^d A(x_1, \dots, x_n)$.

THÉORÈME X.1.3

Soit $K[X_1, \dots, X_n]$ une K -algèbre de polynômes à n lettres X_1, \dots, X_n . Si A et B sont non nuls dans cette algèbre, on a :

(I) $\deg(AB) = \deg(A) + \deg(B)$
 (II) $\text{val}(AB) = \text{val}(A) + \text{val}(B)$.

En particulier, l'anneau $K[X_1, \dots, X_n]$ est *intègre*.

Démonstration :

Les deux propriétés se prouvant par des méthodes analogues, faisons le raisonnement seulement pour (I).

Posons

$$d = \deg(A), \quad e = \deg(B), \quad A = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n},$$

$$B = \sum_{\alpha \in \mathbb{N}^n} c_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n} \quad \text{et} \quad AB = \sum_{\alpha} c_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n}.$$

Si $a_\alpha \neq 0$ (resp. $b_\beta \neq 0$), nécessairement $\|\alpha\| \leq d$ (resp. $\|\beta\| \leq e$). Donc si $\|\gamma\| > d + e$, on est sûr que

$$c_\gamma = \sum_{\alpha + \beta = \gamma} a_\alpha b_\beta = 0,$$

d'où $\deg(AB) \leq d + e$. Le théorème sera démontré si l'on trouve un $\gamma \in \mathbb{N}^n$ tel que $\|\gamma\| = d + e$ et $c_\gamma \neq 0$. Pour cela ordonnons \mathbb{N}^n *lexicographiquement*. Soit \mathcal{E} (resp. \mathcal{F}) l'ensemble (non vide) des $\alpha \in \mathbb{N}^n$ tels que $a_\alpha \neq 0$ et $\|\alpha\| = d$ (resp. des $\beta \in \mathbb{N}^n$ tels que $b_\beta \neq 0$ et $\|\beta\| = e$). Les coefficients c_γ pour $\|\gamma\| = d + e$ sont ceux du produit

$$\left(\sum_{\alpha \in \mathcal{E}} a_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n} \right) \left(\sum_{\beta \in \mathcal{F}} b_\beta X_1^{\beta_1} \dots X_n^{\beta_n} \right).$$

Notons $u = \text{Min}(\mathcal{E})$, où $u = (u_i)$ et $v = \text{Min}(\mathcal{F}) = (v_i)$ et considérons $(\alpha, \beta) \in \mathcal{E} \times \mathcal{F}$, $(\alpha, \beta) \neq (u, v)$, par exemple $\alpha \neq u$. On a donc : soit $\beta = v$, et dans ce cas il existe un indice $p \in \llbracket 1, n \rrbracket$ tel que $\alpha_i = u_i$ pour $i < p$ et $\alpha_p > u_p$, à moins bien sûr que \mathcal{E} et \mathcal{F} se réduisent tous deux à des singletons, mais dans ce cas on connaît l'unique terme de plus haut degré dans le produit AB et il a un coefficient non nul ; soit $\beta \neq v$, et alors il existe deux indices p et q dans $\llbracket 1, n \rrbracket$ tels que $\alpha_i = u_i$ pour $i < p$,

$\beta_i = v_i$ pour $i < q$, $\beta_q > v_q$ (on peut d'ailleurs supposer $p \leq q$ car α et β jouent des rôles identiques).

Alors, de toute façon, $\alpha_p + \beta_p > u_p + v_p$ et $\alpha_i + \beta_i = u_i + v_i$ pour $i < p$, d'où $\alpha + \beta > u + v$. Donc le seul couple $(\alpha, \beta) \in \mathcal{E} \times \mathcal{F}$ tel que $\alpha + \beta = u + v$ est (u, v) , d'où

$$(3) \quad c_{u+v} = a_u b_v.$$

En particulier, $c_{u+v} \neq 0$. Finalement on a bien $\deg(AB) = d + e$. ■

Remarque 2 : La relation (3) qui explicite un monôme particulier parmi les termes de plus haut degré du produit AB est en elle-même bien plus précise que le théorème X.1.3. On aurait pu raisonner de manière analogue avec $u = \text{Max}(\mathcal{E})$ et $v = \text{Max}(\mathcal{F})$.

THÉOREME X.1.4

On suppose le corps de base K infini. Soit f et g deux fonctions polynomiales sur K^n , avec $g \neq 0$. Si f est nulle en tout point de $K^n \setminus g^{-1}(0)$, alors $f = 0$ (« principe du prolongement des identités algébriques »).

Démonstration :

Les fonctions polynomiales sur K^n forment une algèbre de polynômes (cf. corollaire 1 du théorème X.1.2). Cette algèbre est donc intègre (théorème X.1.3). Or $fg = 0$ à cause de l'hypothèse, et $g \neq 0$. Donc $f = 0$. ■

Exemple 2 (E. Galois). On donne a_1, a_2, \dots, a_n tous distincts dans le corps de base K supposé infini. Montrer qu'il existe $(x_1, x_2, \dots, x_n) \in K^n$ tel que les $n!$ éléments $\left(\sum_{i=1}^n a_i x_{\sigma(i)} \right)_{\sigma \in \mathfrak{S}_n}$ soient distincts dans K .

Solution : Soit $\sigma \in \mathfrak{S}_n$, $\tau \in \mathfrak{S}_n$, $\tau \neq \sigma$. La fonction polynomiale $G_{\sigma, \tau} : K^n \rightarrow K$, $(x_1, x_2, \dots, x_n) \mapsto \sum_{i=1}^n (a_{\sigma(i)} - a_{\tau(i)}) x_i$ est non nulle, car les $a_{\sigma(i)} - a_{\tau(i)}$ sont non tous nuls. Puisque K est infini, la fonction

$$G = \prod_{\substack{(\sigma, \tau) \in \mathfrak{S}_n \times \mathfrak{S}_n \\ \sigma \neq \tau}} G_{\sigma, \tau}$$

est donc non nulle. Tout élément $(x_1, x_2, \dots, x_n) \in K^n$ tel que $G(x_1, x_2, \dots, x_n) \neq 0$ répond à la question.

Exemple 3 : Le corps de base est infini. On donne $n \in \mathbb{N}^*$, $n \geq 2$. Soit B une partie finie de $K^n \setminus \{0\}$. Montrer qu'il existe une forme linéaire φ sur K^n telle que $\varphi(x) \neq 0$ pour tout $x \in B$.

Solution : Soit $\varphi_1, \varphi_2, \dots, \varphi_n$ les projections naturelles de K^n dans K ; elles forment une base du K -ev $(K^n)^*$, dual de K^n . Pour chaque $b = (b_1, b_2, \dots, b_n) \in B$, soit $f_b : K^n \rightarrow K$,

$$(u_1, u_2, \dots, u_n) \mapsto b_1 u_1 + b_2 u_2 + \dots + b_n u_n.$$

f_b est polynomiale non nulle sur K^n (car $b \neq 0$). Puisque K est infini, la fonction $f = \prod_{b \in B} f_b$ est non nulle. Choisissons $(u_1, u_2, \dots, u_n) \in K^n$ tel que $f(u_1, \dots, u_n) \neq 0$. Alors pour tout $b \in B$, on a : $u_1 b_1 + \dots + u_n b_n \neq 0$, ce qui peut s'écrire, en introduisant la forme linéaire $\varphi = \sum_{i=1}^n u_i \varphi_i$ sur K^n : $\varphi(b) \neq 0$. Donc φ convient. Dans un langage géométrique, cet exemple prouve qu'il existe toujours un *hyperplan* H ($H = \text{Ker}(\varphi)$) tel que $H \cap B = \emptyset$, ce qui n'est pas aussi évident que cela paraît.

Exercice 1 : Le corps de base est $K = \mathbb{R}$ (ou $K = \mathbb{C}$). On donne $n \in \mathbb{N}^*$. Une fonction polynomiale $f : K^n \rightarrow K$ est supposée nulle au voisinage d'un point $a \in K^n$. Montrer que $f = 0$; donc, si $f \neq 0$, $\{x \in K^n \mid f(x) \neq 0\}$ est un ouvert dense de K^n .

Exercice 2 : Le corps de base est $K = \mathbb{R}$ ou $K = \mathbb{C}$ et $n \in \mathbb{N}^*$. Soit une fonction $f : K^n \rightarrow K$ telle que tout point $a \in K^n$ possède un voisinage V_a pour lequel $f|_{V_a}$ est polynomiale. Montrer que f est polynomiale.

Exercice 3 : Soit $f : \mathbb{C}^n \rightarrow \mathbb{C}$ une fonction polynomiale, avec $n \in \mathbb{N}^*$. Que peut-on dire de f si $|f|$ est une fonction constante ?

Que peut-on dire de f si $\text{Im}(f) \subset \mathbb{R}$?

Exercice 4 : Soit $n \in \mathbb{N}$, $n \geq 2$. On donne une partie dénombrable B de $K^n \setminus \{0\}$ avec $K = \mathbb{R}$ ou $K = \mathbb{C}$. Montrer qu'il existe un hyperplan H tel que $H \cap B = \emptyset$.

Exercice 5 : Soit K un corps commutatif. On donne $f : K^2 \rightarrow K$ possédant les propriétés suivantes : pour tout $x \in K$, la fonction $f_x : K \rightarrow K$, $y \mapsto f(x, y)$ est polynomiale pour tout $y \in K$, la fonction $f^y : K \rightarrow K$, $x \mapsto f(x, y)$ est polynomiale.

a) Si $K = \mathbb{R}$ ou \mathbb{C} , montrer que f est polynomiale.

b) Si $K = \mathbb{Q}$, montrer que f n'est pas forcément polynomiale.

Indication : Soit $n \mapsto r_n$ une bijection de \mathbb{N}^* sur \mathbb{Q} . Utiliser les fonctions $(x, y) \mapsto \prod_{k=1}^n (x - r_k)(y - r_k)$.

Exercice 6 : Soit $n \in \mathbb{N}^*$, un corps de base K infini et une partie finie non vide B de K^n , de cardinal $p \geq 2$. Trouver des fonctions polynomiales $(f_b)_{b \in B}$ sur K^n telles que $f_b(b) = 1$ et $f_b(b') = 0$ pour tous b et b' dans B , $b \neq b'$.

Exercice 7 : Soit K un corps commutatif infini et $K[x, y]$ une algèbre de polynômes en 2 lettres x et y . On considère $X = x^2 - y$ et $Y = x^2 + x - y$. Montrer que l'homomorphisme de substitution $\zeta : K[x, y] \rightarrow K[x, y]$ tel que $\zeta(x) = X$ et $\zeta(y) = Y$ est un *automorphisme* de la K -algèbre $K[x, y]$.

Exercice 8 : Le corps de base K est infini et $n \in \mathbb{N}^*$. Soit A et $B \in K[X_1, \dots, X_n]$ K -algèbre de polynômes à n lettres (X_i) . Si AB est homogène, alors A et B le sont

Exercice 9 : On donne une algèbre de polynômes à n lettres $K[X_1, \dots, X_n]$ ($n \geq 1$, K infini) et des polynômes F_1, \dots, F_p dans $K[X_1, \dots, X_n]$ avec $\deg(F_i) = d_i \geq 1$ pour $1 \leq i \leq p$. Montrer qu'il existe des polynômes Φ_1, \dots, Φ_n homogènes de degré 1 et linéairement indépendants tels que, pour tout $i \in \llbracket 1, p \rrbracket$, $F_i(\Phi_1, \Phi_2, \dots, \Phi_n)$ soit de la forme $X_1^{d_i} + A_{i,1} X_1^{d_i-1} + \dots + A_{i,d_i}$ avec $A_{i,j} \in K[X_2, \dots, X_n]$.

Exercice 10 : On donne une algèbre de polynômes à n lettres ($n \geq 1$, K infini) et des polynômes $\Phi_1, \Phi_2, \dots, \Phi_n$ homogènes de degré 1 et linéairement indépendants dans $K[X_1, X_2, \dots, X_n]$ ainsi que des scalaires c_1, c_2, \dots, c_n dans K . On pose $Y_i = \Phi_i + c_i$. Montrer que l'homomorphisme de substitution $S: K[X_1, \dots, X_n] \rightarrow K[X_1, \dots, X_n]$ tel que $S(X_i) = Y_i$ pour tout i est un *automorphisme* de la K -algèbre $K[X_1, \dots, X_n]$ et que $\deg(S(F)) = \deg(F)$ pour tout F . Réciproque. Cette réciproque tient-elle encore si l'on renonce à la conservation du degré ? (cf. exercice 7).

Exercice 11 : Le corps de base K est infini et l'entier $n \geq 2$. Pour $i \in \llbracket 1, n \rrbracket$, $j \in \llbracket 1, n \rrbracket$, $i \neq j$ on pose $H_{i,j} = \{(x_1, x_2, \dots, x_n) \in K^n \mid x_i = x_j\}$ et on donne une fonction polynomiale $f: K^n \rightarrow K$. En n'utilisant que les résultats sur les polynômes à une indéterminée (cf. Chap. VII), démontrer :

a) Pour que f s'annule sur $H_{i,j}$ (i et j donnés, $i \neq j$), il faut et il suffit que f soit divisible par $\varphi_i - \varphi_j$ dans la K -algèbre $K[\varphi_1, \dots, \varphi_n]$ des fonctions polynomiales sur K^n .

b) Pour que f s'annule sur $\bigcup_{(i,j) \in \llbracket 1, n \rrbracket^2, i \neq j} H_{i,j}$, il faut et il suffit que f soit divisible par

$$\prod_{1 \leq i < j \leq n} (\varphi_i - \varphi_j) \text{ dans } K[\varphi_1, \varphi_2, \dots, \varphi_n].$$

Exercice 12 : Soit une K -algèbre de polynômes à n lettres $K[X_1, X_2, \dots, X_n] = \mathcal{P}$ ($n \geq 1$, K infini). On considère L_1, L_2, \dots, L_n et G_1, G_2, \dots, G_n dans \mathcal{P} tels que : pour chaque $i \in \llbracket 1, n \rrbracket$, $\text{val}(G_i) \geq 2$; L_1, \dots, L_n sont homogènes de degré 1 et linéairement indépendants. Soit $Y_i = L_i + G_i$ ($1 \leq i \leq n$). Montrer que (Y_1, \dots, Y_n) sont algébriquement libres sur K .

Exercice 13 : Montrer à l'aide d'un exemple que l'anneau intègre $K[X_1, \dots, X_n]$ n'est pas principal pour $n \geq 2$.

§ X.2 DÉRIVÉES PARTIELLES. DEGRÉ PARTIEL

Soit K un corps commutatif, et $K[X_1, X_2, \dots, X_n]$ une K -algèbre de polynômes à n lettres (X_i) . Si $P = \sum_{\alpha \in \mathbb{N}^n} \lambda_\alpha X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$ est un élément

de $K[X_1, X_2, \dots, X_n]$, et si $i \in \llbracket 1, n \rrbracket$, P s'écrit de manière unique :

$$(1) \quad P = \sum_{k \in \mathbb{N}} A_k X_i^k, \text{ avec } A_k \in K[(X_j)_{j \neq i}] \text{ pour tout } k, \text{ la famille } (A_k)$$

étant à support fini. On dit qu'on a *ordonné P par rapport à X_i* . L'existence de la décomposition (1) est conséquence immédiate du *principe de Fubini* (cf. § III.1), et son unicité résulte de l'indépendance linéaire sur K des monômes des (X_j) .

Cette écriture (1) nous permet de définir deux importantes notions :

- Tout d'abord, la *dérivée partielle de P par rapport à X_i* notée $\frac{\partial P}{\partial X_i}$.

c'est, par définition, le polynôme $\sum_{k \in \mathbb{N}} (k+1) A_{k+1} X_i^k$. On vérifie facilement que la dérivation ainsi définie obéit aux lois suivantes :

(2) Pour i fixé, l'application $D_i : K[X_1, \dots, X_n] \rightarrow K[X_1, \dots, X_n]$, $P \mapsto \frac{\partial P}{\partial X_i}$ est K -linéaire.

(3) Pour i fixé, quels que soient F et G dans $K[X_1, \dots, X_n]$, on a :

$$D_i(FG) = D_i(F) \cdot G + F \cdot D_i(G).$$

(4) Pour tous i et j dans $\llbracket 1, n \rrbracket$ et

$$P \in K[X_1, \dots, X_n] \quad D_i D_j(P) = D_j D_i(P).$$

Les endomorphismes D_1, D_2, \dots, D_n du K -ev $K[X_1, \dots, X_n]$ sont donc deux à deux permutables. Ils engendrent dans $\text{Hom}_K(K[X_1, \dots, X_n])$ une sous- K -algèbre commutative appelée **algèbre des polynômes de dérivation**. Le résultat d'un monôme $D_1^{\alpha_1} \dots D_n^{\alpha_n}$ sur $P \in K[X_1, \dots, X_n]$ se note aussi

$$\frac{\partial^{\alpha} P}{\partial X_1^{\alpha_1} \dots \partial X_n^{\alpha_n}}.$$

(5) Si P est *homogène* de degré d dans $K[X_1, \dots, X_n]$, alors

$$\sum_{i=1}^n X_i \frac{\partial P}{\partial X_i} = d \cdot P \quad (\text{formule d'Euler}).$$

● Ensuite, le **degré partiel de P par rapport à X_i** : c'est, par définition, $-\infty$ si $P = 0$, et $\text{Max} \{k \mid A_k \neq 0\}$ si $P \neq 0$. On le note $\deg_{X_i}(P)$. Ce degré vérifie les propriétés habituelles :

$$(6) \quad \deg_{X_i}(A + B) \leq \text{Max}(\deg_{X_i}(A), \deg_{X_i}(B))$$

et si $\deg_{X_i}(A) \neq \deg_{X_i}(B)$ il y a égalité.

$$(7) \quad \deg_{X_i}(AB) = \deg_{X_i}(A) + \deg_{X_i}(B).$$

Ces propriétés se démontrent comme les propriétés analogues des polynômes d'une variable (cf. § VII.1).

Pour le distinguer du degré partiel, le degré de P introduit au § X.1 prend parfois le nom de **degré total** de P .

Remarque 1 : On peut aussi définir le degré partiel de $P \in K[X_1, X_2, \dots, X_n]$ par rapport à certaines des indéterminées $(X_i)_{i \in J}$, où $J \subset \llbracket 1, n \rrbracket$ est donné. Supposons par exemple que $J = \llbracket 1, p \rrbracket$. On justifie comme ci-dessus l'existence

la décomposition de P sous la forme

$$(8) \quad P = \sum_{\beta \in \mathbb{N}^p} B_\beta X_1^{\beta_1} \dots X_p^{\beta_p}, \quad \text{avec } B_\beta \in K[X_{p+1}, \dots, X_n] \text{ pour tout } \beta,$$

la famille (B_β) étant à support fini. Le degré partiel de P par rapport à (X_1, X_2, \dots, X_p) est alors, par définition, $-\infty$ si $P = 0$, et $\text{Max} \{ \|\beta\| \mid \beta \in \mathbb{N}^p \text{ et } B_\beta \neq 0 \}$ si $P \neq 0$. On le note $\deg_{(X_1, \dots, X_p)}(P)$. Il vérifie encore les propriétés analogues à (6) et (7).

Formule de Taylor

THÉORÈME X.2.1

$$\left\| \begin{array}{l} \text{Supposons le corps } K \text{ de caractéristique } 0. \\ \text{Alors, si } n \in \mathbb{N}^*, \text{ si } (h_1, h_2, \dots, h_n) \in K^n, \text{ et si } P \in K[X_1, \dots, X_n] \\ \text{est un polynôme à } n \text{ lettres } (X_i) \text{ sur } K, \text{ on a la formule de Taylor :} \\ \\ (9) \quad P(X_1 + h_1, X_2 + h_2, \dots, X_n + h_n) = \\ \qquad \qquad \qquad = \sum_{p=0}^{+\infty} \frac{1}{p!} (h_1 D_1 + \dots + h_n D_n)^p \cdot P \\ \\ \text{en notant } D_i \text{ l'opérateur } Q \mapsto \frac{\partial Q}{\partial X_i} \quad (1 \leq i \leq n, Q \in K[X_1, \dots, X_n]). \end{array} \right.$$

Démonstration :

Notons d'abord que (9) a un sens car la somme du second membre est finie, puisque

$$\begin{aligned} (h_1 D_1 + \dots + h_n D_n)^p \cdot P &= \\ &= \sum_{\beta_1 + \beta_2 + \dots + \beta_n = p} \frac{p!}{\beta_1! \dots \beta_n!} h_1^{\beta_1} \dots h_n^{\beta_n} \frac{\partial^p P}{\partial X_1^{\beta_1} \dots \partial X_n^{\beta_n}} \end{aligned}$$

est certainement nul dès que $p > \deg(P)$.

Par K -linéarité, il suffit de prouver (9) lorsque P est un monôme, par exemple $A = X_1^{\alpha_1} \dots X_n^{\alpha_n}$. Dans ce cas, on a :

$$\begin{aligned} A(X_1 + h_1, X_2 + h_2, \dots, X_n + h_n) &= (X_1 + h_1)^{\alpha_1} \dots (X_n + h_n)^{\alpha_n} = \\ &= \left(\sum_{0 \leq \beta_1 \leq \alpha_1} \binom{\alpha_1}{\beta_1} h_1^{\beta_1} X_1^{\alpha_1 - \beta_1} \right) \dots \left(\sum_{0 \leq \beta_n \leq \alpha_n} \binom{\alpha_n}{\beta_n} h_n^{\beta_n} X_n^{\alpha_n - \beta_n} \right), \end{aligned}$$

qui se développe (cf. § III.1 et § III.2) en :

$$\sum_{\substack{(\beta_1, \dots, \beta_n) \in \mathbb{N}^n \\ \beta_i \leq \alpha_i \text{ pour tout } i}} \binom{\alpha_1}{\beta_1} \binom{\alpha_2}{\beta_2} \dots \binom{\alpha_n}{\beta_n} h_1^{\beta_1} h_2^{\beta_2} \dots h_n^{\beta_n} X_1^{\alpha_1 - \beta_1} X_2^{\alpha_2 - \beta_2} \dots X_n^{\alpha_n - \beta_n},$$

ce qui s'écrit, compte tenu du fait que la caractéristique de K est nulle :

$$(10) \quad A(X_1 + h_1, \dots, X_n + h_n) = \sum_{(\forall i) 0 \leq \beta_i \leq \alpha_i} \frac{1}{\beta_1! \dots \beta_n!} h_1^{\beta_1} \dots h_n^{\beta_n} (D_1^{\beta_1} \dots D_n^{\beta_n} \cdot A).$$

Si l'on permet, au second membre de (10), à (β_i) de varier dans \mathbb{N}^n , on ne fait que rajouter des termes nuls. Regroupons alors dans un même paquet d'indice p tous les termes du second membre de (10) tels que $\beta_1 + \beta_2 + \dots + \beta_n = p$ ($p \in \mathbb{N}$). On obtient :

$$A(X_1 + h_1, \dots, X_n + h_n) = \sum_{p \in \mathbb{N}} \frac{1}{p!} \left[\left(\sum_{\beta_1 + \beta_2 + \dots + \beta_n = p} \frac{p!}{\beta_1! \dots \beta_n!} h_1^{\beta_1} \dots h_n^{\beta_n} D_1^{\beta_1} \dots D_n^{\beta_n} \right) \cdot A \right],$$

ce qui démontre bien (9) (formule du multinôme). ■

Exercice 1 : Le corps K est supposé de caractéristique nulle. Soit F un polynôme à n lettres sur K . Montrer que si, pour $k \in \mathbb{N}$, on a :

$$\sum_{i=1}^n X_i \frac{\partial F}{\partial X_i} = kF, \quad \text{alors } F \text{ est homogène de degré } k.$$

Exercice 2 : La caractéristique de K étant nulle, montrer que dans l'algèbre des endomorphismes du K -ev $K[X_1, \dots, X_n]$ des polynômes à n lettres (X_i) , les dérivations D_1, D_2, \dots, D_n sont algébriquement libres sur K (i.e. les monômes $(D_1^{\alpha_1} \dots D_n^{\alpha_n})_{\alpha \in \mathbb{N}^n}$ sont linéairement indépendants sur K).

Exercice 3 : Soit F et G dans $K[X_1, \dots, X_n]$ des polynômes en les n lettres (X_i) . Démontrer la formule de Leibniz : si $p = (p_1, p_2, \dots, p_n) \in \mathbb{N}^n$,

$$\frac{\partial^{p_1} \dots \partial^{p_n} (FG)}{\partial X_1^{p_1} \dots \partial X_n^{p_n}} = \sum_{\substack{\lambda \in \mathbb{N}^n, \mu \in \mathbb{N}^n \\ \lambda + \mu = p}} \frac{p_1! \dots p_n!}{\lambda_1! \dots \lambda_n! \mu_1! \dots \mu_n!} \frac{\partial^{\lambda_1} \dots \partial^{\lambda_n} F}{\partial X_1^{\lambda_1} \dots \partial X_n^{\lambda_n}} \cdot \frac{\partial^{\mu_1} \dots \partial^{\mu_n} G}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}}.$$

Exercice 4 : Soit $K[X_1, \dots, X_n]$ une algèbre de polynômes, et F, G_1, G_2, \dots, G_n des éléments de cette algèbre. On pose $P = F(G_1, G_2, \dots, G_n)$. Démontrer :

$$\forall i \in \llbracket 1, n \rrbracket \quad \frac{\partial P}{\partial X_i} = \sum_{k=1}^n \frac{\partial G_k}{\partial X_i} \cdot \frac{\partial F}{\partial X_k}(G_1, G_2, \dots, G_n).$$

(Indication : il suffit de le prouver lorsque F est un monôme.)

Exercice 5 (cet exercice suppose connues les propriétés élémentaires des déterminants).

Le corps K est de caractéristique nulle. Soit F_1, F_2, \dots, F_n des éléments d'une algèbre $K[X_1, \dots, X_n]$ de polynômes en les n lettres X_i . On suppose que le déterminant

$$\text{Jac}(F_1, \dots, F_n) = \det \left[\frac{\partial F_i}{\partial X_j} \right]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \quad \text{est non nul.}$$

Démontrer que F_1, F_2, \dots, F_n sont algébriquement libres sur K (utiliser l'exercice 4).

Remarque : On peut prouver que la réciproque est vraie, mais il faut la notion de *degré de transcendance* d'une extension. Voir [27] par exemple.

§ X.3 FONCTIONS SYMÉTRIQUES

Jusqu'à la fin de ce chapitre, le corps de base sera \mathbb{C} .

$n \in \mathbb{N}^*$ étant supposé fixé, nous considérons donc une \mathbb{C} -algèbre de polynômes $\mathbb{C}[X_1, X_2, \dots, X_n]$ en les n lettres X_i .

Rappelons qu'en tant que \mathbb{C} -ev, cette algèbre se décompose en somme directe : $\mathbb{C}[X_1, \dots, X_n] = \bigoplus_{e \in \mathbb{N}} \mathcal{H}_e$, où \mathcal{H}_e est le \mathbb{C} -ev des polynômes homogènes de degré e .

Si $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$, on posera $\|\alpha\| = \alpha_1 + \alpha_2 + \dots + \alpha_n$ et $\delta(\alpha) = \max_{1 \leq i \leq n} (\alpha_i)$. Il sera commode d'ordonner \mathbb{N}^n lexicographiquement ;

cet ordre, noté \leq , est défini par : $\forall \alpha \in \mathbb{N}^n, \forall \beta \in \mathbb{N}^n$, si $\alpha = \beta$ on a $\alpha \leq \beta$, et si $\alpha \neq \beta$, en appelant i le plus petit indice $j \in \llbracket 1, n \rrbracket$ tel que $\alpha_j \neq \beta_j$, on a $\alpha \leq \beta$ ssi $\alpha_i < \beta_i$. L'ordre lexicographique est *total*. Nous utiliserons aussi sur \mathbb{N}^n une autre relation d'ordre, notée \leq , qui est l'ordre *naturel produit*, défini par $\alpha \leq \beta$ ssi $(\forall i) \alpha_i \leq \beta_i$, non total si $n \geq 2$.

Action du groupe \mathfrak{S}_n sur \mathbb{N}^n

On obtient une action à gauche, dite *naturelle*, du groupe \mathfrak{S}_n sur \mathbb{N}^n en associant, à $\sigma \in \mathfrak{S}_n$ et $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$, l'élément $\sigma * \alpha = (\alpha_{\sigma^{-1}(1)}, \alpha_{\sigma^{-1}(2)}, \dots, \alpha_{\sigma^{-1}(n)})$ de \mathbb{N}^n .

Si $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, posons $E_i(\alpha) = \{j \in \llbracket 1, n \rrbracket \mid \alpha_j = i\}$ et $\omega_i(\alpha) = \text{card}(E_i(\alpha))$, ($i \in \llbracket 0, \delta(\alpha) \rrbracket$). La famille $(E_i(\alpha))$ est un *partage* de $\llbracket 1, n \rrbracket$.

Le stabilisateur Γ_α de α est le groupe des $\sigma \in \mathfrak{S}_n$ tels que $\sigma(E_i(\alpha)) = E_i(\alpha)$ pour tout i . Si donc nous posons, pour abréger, $\omega_i = \omega_i(\alpha)$ et $d = \delta(\alpha)$, le groupe Γ_α s'identifie à $\mathfrak{S}_{\omega_0} \times \mathfrak{S}_{\omega_1} \times \dots \times \mathfrak{S}_{\omega_d}$; son cardinal est $\omega_0! \omega_1! \dots \omega_d!$, et on a :

$$\omega_0 + \omega_1 + \dots + \omega_d = n.$$

L'orbite \mathcal{O} de α possède donc $\frac{n!}{\omega_0! \omega_1! \dots \omega_d!}$ éléments. Si $\beta = (\beta_i)$ est un autre élément de cette orbite, on a $\delta(\beta) = \delta(\alpha)$, et $\omega_i(\beta) = \omega_i(\alpha)$ pour tout $i \in \llbracket 0, d \rrbracket$.

Chaque *classe à gauche* U de \mathfrak{S}_n suivant le groupe Γ_α contient une, et une seule, permutation s_U dont la restriction à chaque $E_i(\alpha)$ est *strictement croissante* : en effet une telle classe U est l'ensemble des $\sigma \in \mathfrak{S}_n$ tels que $\sigma * \alpha = \beta$, où $\beta \in \mathcal{O}$ est donné ; donc $U = \{\sigma \in \mathfrak{S}_n \mid \sigma(E_i(\alpha)) = E_i(\beta)\}$ pour $i \in \llbracket 0, d \rrbracket$ et l'on voit bien qu'il y a une et une seule $\sigma \in U$ dont la restriction à chaque $E_i(\alpha)$ soit une bijection croissante sur

Il est facile, à partir de là, de classer les orbites de \mathbb{N}^n pour l'action naturelle de \mathfrak{S}_n : soit \mathfrak{p}_n l'ensemble des $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{N}^n$ tels que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

LEMME 1

|| Chaque \mathfrak{S}_n -orbite \mathcal{O} de \mathbb{N}^n rencontre \mathfrak{p}_n suivant un singleton.

Démonstration :

Soit $\beta = (\beta_i) \in \mathcal{O}$. En réordonnant les β_i par ordre décroissant, on voit que β rencontre \mathfrak{p}_n .

Soit maintenant $\alpha \in \mathfrak{p}_n$ et $\beta \in \mathfrak{p}_n$. Les ensembles $E_i(\alpha)$ (resp. $E_i(\beta)$) sont consécutifs, i.e. si $i < j$, on a $x > y$ pour tous $x \in E_i(\alpha)$ et $y \in E_j(\alpha)$ (resp. $E_i(\beta)$). Supposons α et β dans la même \mathfrak{S}_n -orbite : l'unique $\sigma \in \mathfrak{S}_n$ telle que $\sigma * \alpha = \beta$ et que $\sigma|_{E_i(\alpha)}$ soit une bijection croissante sur $E_i(\beta)$ pour tout i est $\text{Id}_{[1, n]}$ par ce qui précède, donc $\alpha = \beta$.

Par suite $\text{card}(\mathcal{O} \cap \mathfrak{p}_n) = 1$. ■

Remarque 1 : Soit $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathcal{O}$. L'unique élément de $\mathcal{O} \cap \mathfrak{p}_n$ ainsi mis en évidence est la suite $(\lambda_1, \lambda_2, \dots, \lambda_n)$ formée en écrivant d'abord les ω_d termes α_i égaux à $d = \delta(\alpha)$, puis les ω_{d-1} termes égaux à $d-1$, et ainsi de suite jusqu'aux ω_0 termes égaux à 0. C'est donc *le plus grand terme de \mathcal{O} pour l'ordre lexicographique* (ce que nous noterons en abrégé P.G.O.L. (\mathcal{O})).

A $\lambda = (\lambda_i) \in \mathfrak{p}_n$, associons la suite $(\omega_0, \omega_1, \dots, \omega_d)$, où $d = \delta(\lambda)$ et $\omega_i = \omega_i(\lambda)$. Pour d fixé $\neq 0$, on obtient ainsi une bijection de l'ensemble des $\lambda \in \mathfrak{p}_n$ de $\delta(\lambda) = d$ sur la partie de \mathbb{N}^{d+1} formée des $(\omega_0, \omega_1, \dots, \omega_d)$ tels que $\sum_{i=0}^d \omega_i = n$ avec $\omega_d \geq 1$.

Polynômes symétriques, polynômes alternés

Soit $F = \sum_{\alpha \in \mathbb{N}^n} A_\alpha \mathcal{M}_\alpha(X) \in \mathbb{C}[X_1, X_2, \dots, X_n]$ et $\sigma \in \mathfrak{S}_n$.

Posons $\sigma * F = F(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)}) = \sum_{\alpha \in \mathbb{N}^n} A_\alpha \mathcal{M}_{\sigma * \alpha}(X)$.

On définit ainsi une action à gauche de \mathfrak{S}_n sur $\mathbb{C}[X_1, X_2, \dots, X_n]$ (les vérifications sont immédiates). Cette action laisse stables les \mathcal{H}_e ($e \in \mathbb{N}$) : si $\sigma \in \mathfrak{S}_n$ et $F \in \mathcal{H}_e$, alors $\sigma * F \in \mathcal{H}_e$. De plus, pour $\sigma \in \mathfrak{S}_n$ fixé, l'application $\mathbb{C}[X_1, \dots, X_n] \rightarrow \mathbb{C}[X_1, \dots, X_n]$, $F \mapsto \sigma * F$ est \mathbb{C} -linéaire (on dit que \mathfrak{S}_n opère linéairement sur $\mathbb{C}[X_1, \dots, X_n]$). Mais aussi, pour tous polynômes F et G , on a : $\sigma * (FG) = (\sigma * F)(\sigma * G)$. Finalement, pour $\sigma \in \mathfrak{S}_n$ fixé, $f_\sigma : F \mapsto \sigma * F$ est un automorphisme de la \mathbb{C} -algèbre $\mathbb{C}[X_1, \dots, X_n]$, dont la réciproque est $f_{\sigma^{-1}}$.

DÉFINITION X.3.1

Soit Γ un sous-groupe de \mathfrak{S}_n . L'ensemble \mathcal{A}_Γ des polynômes $F \in \mathbb{C}[X_1, \dots, X_n]$ tels que $\sigma * F = F$ pour tout $\sigma \in \Gamma$ s'appelle ensemble des polynômes Γ -invariants. Si $\Gamma = \mathfrak{S}_n$, \mathcal{A}_Γ est l'ensemble des polynômes symétriques. Si $\Gamma = \mathfrak{U}_n$, \mathcal{A}_Γ est l'ensemble des polynômes alternés.

Pour Γ fixé, \mathcal{A}_Γ est une sous- \mathbb{C} -algèbre de $\mathbb{C}[X_1, \dots, X_n]$. Pour $\Gamma_1 \subset \Gamma_2 \subset \mathfrak{S}_n$, on a $\mathcal{A}_{\Gamma_2} \subset \mathcal{A}_{\Gamma_1}$. En particulier, on a toujours $\mathcal{A}_{\mathfrak{S}_n} \subset \mathcal{A}_\Gamma$.

Exemple 1 : Pour tout $d \in \mathbb{N}$, le polynôme $S_d = X_1^d + X_2^d + \dots + X_n^d$ est symétrique. Ce polynôme s'appelle *somme de Newton* de degré d .

Exemple 2 : Pour $k \in \llbracket 1, n \rrbracket$, soit $\sigma_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} X_{i_1} X_{i_2} \dots X_{i_k}$. Le

polynôme σ_k est symétrique, car la fonction correspondante l'est.

Ce fait a été expliqué au § VII.5.

Les polynômes σ_k sont appelés *polynômes symétriques élémentaires* des X_i .

Exemple 3 : Soit $\Delta = \prod_{1 \leq i < j \leq n} (X_j - X_i)$. Nous avons vu, au § V.4, qu'on a

$$\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) \Delta(x_1, \dots, x_n),$$

pour $\sigma \in \mathfrak{S}_n$ et $(x_1, x_2, \dots, x_n) \in \mathbb{Q}^n$, $\varepsilon(\sigma)$ désignant la signature de la permutation σ .

Puisque \mathbb{Q} est infini, cela montre que $\sigma * \Delta = \varepsilon(\sigma) \Delta$. En particulier Δ est alterné. Δ s'appelle le *polynôme de Vandermonde* ⁽¹⁾. Remarquons que $D = \Delta^2$ est un polynôme symétrique. On l'appelle le **discriminant général**.

LEMME 2

|| Les polynômes $\sigma_1, \sigma_2, \dots, \sigma_n$ sont algébriquement libres sur \mathbb{C} .

Démonstration :

Considérons l'application $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$, $(x_1, \dots, x_n) = \underline{x} \mapsto (\sigma_1(\underline{x}), \sigma_2(\underline{x}), \dots, \sigma_n(\underline{x}))$. On a, en désignant par T une indéterminée sur \mathbb{C} :

$$(T - x_1)(T - x_2) \dots (T - x_n) = T^n - \sigma_1(\underline{x}) T^{n-1} + \dots + (-1)^n \sigma_n(\underline{x})$$

($\underline{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$). Puisque \mathbb{C} est algébriquement clos, cela montre que f est surjective.

⁽¹⁾ Alexandre Théophile Vandermonde, mathématicien et physicien franç

Considérons des nombres complexes $(c_\alpha)_{\alpha \in \mathbb{N}^n}$, formant une famille à support fini, tels que $\sum_{\alpha \in \mathbb{N}^n} c_\alpha \sigma_1^{\alpha_1} \dots \sigma_n^{\alpha_n} = 0$. La surjectivité de f montre que la fonction polynomiale $\mathbb{C}^n \rightarrow \mathbb{C}$, $(t_1, t_2, \dots, t_n) \mapsto \sum_{\alpha \in \mathbb{N}^n} c_\alpha t_1^{\alpha_1} t_2^{\alpha_2} \dots t_n^{\alpha_n}$, est nulle. Donc tous les c_α sont nuls (corollaire 1 du théorème X.1.2), ce qui démontre le lemme. ■

Fonctions monomiales

Considérons une \mathfrak{S}_n -orbite \mathcal{O} de \mathbb{N}^n pour l'action naturelle. Le polynôme $G = \sum_{\alpha \in \mathcal{O}} \mathcal{M}_\alpha(\underline{X})$ est **symétrique**, car si $\sigma \in \mathfrak{S}_n$, l'application $\mathcal{O} \rightarrow \mathcal{O}$, $\alpha \mapsto \sigma * \alpha$ est bijective, d'où $\sigma * G = \sum_{\alpha \in \mathcal{O}} \mathcal{M}_{\sigma * \alpha}(\underline{X}) = G$. Les coefficients de G sont tous égaux à 1.

On a vu que \mathcal{O} est caractérisée par son P.G.O.L. $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathfrak{p}_n$, ce qui conduit à noter le polynôme G tout simplement G_λ , et ce polynôme G_λ sera appelé **fonction monomiale** (associée à l'élément $\lambda \in \mathfrak{p}_n$).

Le nombre $J(\lambda) = \lambda_1$ sera appelé l'**ordre** de G_λ (cet ordre est nul ssi $\lambda = 0_{\mathbb{N}^n}$, c'est-à-dire si $G_\lambda = 1$).

Si $d = \delta(\lambda) \geq 1$, on note $\omega_i(\lambda) = \text{card} \{j \in \llbracket 1, n \rrbracket \mid \lambda_j = i\}$, d'où $\omega_0(\lambda) + \dots + \omega_d(\lambda) = n$. Nous aurons également besoin un peu plus loin du nombre $\varpi(\lambda) = (n+1)d + \omega_d(\lambda)$.

Si l'on écrit en abrégé ω_i à la place de $\omega_i(\lambda)$, on constate que le nombre de monômes de la fonction monomiale G_λ est $\text{card}(\mathcal{O}) = \frac{n!}{\omega_0! \omega_1! \dots \omega_n!}$.

Les fonctions monomiales G_λ sont les plus naturelles de toutes les fonctions polynomiales symétriques. Ce sont elles qui se présentent dans la plupart des calculs concrets concernant ces fonctions symétriques. Il importe donc de savoir conduire ces calculs sans inquiétude, le but de ce qui suit étant de montrer comment on y parvient.

Exemple 4 : Si $\lambda_1 = \lambda_2 = \dots = \lambda_k = 1$ et $\lambda_i = 0$ pour $i \in \llbracket k+1, n \rrbracket$, on retrouve $G_\lambda = \sigma_k$, k -ième polynôme symétrique élémentaire.

Si $\lambda_1 = d$ et $\lambda_i = 0$ pour $i \geq 2$, on obtient $G_\lambda = S_d$, d -ième polynôme de Newton.

Exemple 5 : Soit $n \geq 3$. Prenons $\lambda_1 = \lambda_2 = 2$, $\lambda_3 = 1$ et $\lambda_i = 0$ pour $i > 3$. On a :

$$G_\lambda = \sum_{i < j, (i, j, k) \text{ distincts}} X_i^2 X_j^2 X_k.$$

Le nombre de monômes de G_λ est $\frac{n!}{(n-3)! 1! 2!} = \frac{n(n-1)(n-2)}{2}$.

même pour $n \geq 4$ si l'on prend $\lambda_1 = \lambda_2 = 2$, $\lambda_3 = \lambda_4 = 1$, $\lambda_i = 0$ pour $i > 4$, on a :

$$G_\lambda = \sum_{i < j, k < l, (i, j, k, l) \text{ distincts}} X_i^2 X_j^2 X_k X_l.$$

Le nombre de monômes de G_λ est alors $\frac{n(n-1)(n-2)(n-3)}{4}$. Le plus grand monôme dans l'ordre lexicographique est $X_1^2 X_2^2 X_3 X_4$.

Si l'on veut représenter symboliquement une fonction monomiale quelconque G_λ , en posant pour abrégé $d = \delta(\lambda)$ et $\omega_i = \omega_i(\lambda)$ et en se souvenant que $\omega_d + \omega_{d-1} + \dots + \omega_0 = n$, on pourra écrire :

$$G_\lambda = \left\{ \begin{array}{ccccc} d & d-1 & \dots & 1 & 0 \\ \omega_d & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{array} \right\}$$

notation permettant de reconstituer

$$(1) \quad G_\lambda = \sum_{\sigma \in \mathcal{C}_\lambda} X_{\sigma(1)}^{\lambda_1} X_{\sigma(2)}^{\lambda_2} \dots X_{\sigma(n)}^{\lambda_n}$$

où \mathcal{C}_λ est l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ dont la restriction à chaque $(E_i(\lambda))_{i \in \llbracket 0, d \rrbracket}$ est *strictement croissante* ⁽¹⁾.

THÉOREME X.3.1

La famille $(G_\lambda)_{\lambda \in \mathfrak{p}_n}$ des fonctions monomiales est une base du \mathbb{C} -ev des polynômes symétriques. De plus, si F est un polynôme symétrique à coefficients dans un sous-anneau A de \mathbb{C} , ses coordonnées dans la base (G_λ) appartiennent à A .

Démonstration :

Soit Ω l'ensemble des \mathfrak{S}_n -orbites de \mathbb{N}^n pour l'action naturelle. On sait que les monômes $(\mathcal{M}_\alpha(\underline{X}))_{\alpha \in \mathbb{N}^n}$ forment une base du \mathbb{C} -ev $\mathbb{C}[X_1, \dots, X_n]$, et Ω est une partition de \mathbb{N}^n . Il est donc clair que les polynômes (G_λ) sont \mathbb{C} -linéairement indépendants, car si \mathcal{O} est l'orbite de $\lambda \in \mathfrak{p}_n$, on a $G_\lambda = \sum_{\alpha \in \mathcal{O}} \mathcal{M}_\alpha(\underline{X})$.

Considérons un polynôme symétrique

$$(2) \quad F = \sum_{\alpha \in \mathbb{N}^n} c_\alpha \mathcal{M}_\alpha(\underline{X}).$$

⁽¹⁾ Rappelons que $E_i(\lambda) = \{j \in \llbracket 1, n \rrbracket \mid \lambda_j = i\}$.

On a vu que si $\sigma \in \mathfrak{S}_n$, $\sigma * F = \sum_{\alpha \in \mathbb{N}^n} c_\alpha \mathcal{M}_{\sigma * \alpha}(\underline{X})$, d'où $c_\alpha = c_{\sigma^{-1} * \alpha}$ pour

tout $\sigma \in \mathfrak{S}_n$. En particulier, si α parcourt une même \mathfrak{S}_n -orbite \mathcal{O} , tous les c_α sont égaux à un même nombre $A_\mathcal{O}$ ne dépendant que de \mathcal{O} , d'où, en regroupant dans (2) les termes par paquets correspondant à une même orbite :

$$(3) \quad F = \sum_{\mathcal{O} \in \Omega} A_\mathcal{O} G_{\lambda_\mathcal{O}}, \quad \text{où } \lambda_\mathcal{O} \text{ est le P.G.O.L. de } \mathcal{O}.$$

Finalement la famille (G_λ) est libre et génératrice. C'est donc une base du \mathbb{C} -ev $\mathbb{C}[X_1, \dots, X_n]$, et cette preuve montre bien que si tous les c_α sont dans un même sous-anneau A de \mathbb{C} , les $A_\mathcal{O}$ y sont de même. ■

Le théorème fondamental sur les fonctions symétriques

Considérons $\lambda \in \mathfrak{p}_n$, qui définit la fonction monomiale

$$G_\lambda = \left\{ \begin{array}{cccccc} d & d-1 & \dots & 1 & 0 \\ \omega_d & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{array} \right\} \quad (d \geq 1)$$

soit $\omega \in \llbracket 1, \omega_d \rrbracket$. Etudions le produit $P = \sigma_\omega G_\lambda$, où σ_ω désigne la ω -ième fonction symétrique élémentaire ; P est un polynôme homogène symétrique de degré $N + \omega$, où $N = \deg(G_\lambda) = \sum_{i=0}^d i \omega_i$. Classons les

monômes $(\mathcal{M}_\alpha(\underline{X}))_{\alpha \in \mathbb{N}^n}$ par l'ordre lexicographique de \mathbb{N}^n . Si F est un polynôme homogène non nul, soit $\Lambda(F)$ le coefficient du plus grand monôme de F à coefficient $\neq 0$. Alors, pour tous F_1 et F_2 , on a : $\Lambda(F_1 F_2) = \Lambda(F_1) \Lambda(F_2)$ (F_1 et F_2 homogènes) (cf. la remarque 2 qui suit le théorème X.1.3). Or ici, $\Lambda(\sigma_\omega) = 1 = \Lambda(G_\lambda)$ car les plus grands monômes de σ_ω et de G_λ sont respectivement $X_1 X_2 \dots X_\omega$ et $X_1^{\lambda_1} X_2^{\lambda_2} \dots X_n^{\lambda_n}$. La décomposition du polynôme produit $P = \sigma_\omega G_\lambda$ dans la base monomiale $(G_\nu)_{\nu \in \mathfrak{p}_n}$ se réduit à deux parties bien distinctes : d'une part, avec le coefficient $1 = \Lambda(\sigma_\omega G_\lambda) = \Lambda(P)$, la fonction monomiale

$$G_\mu = \left\{ \begin{array}{cccccc} d+1 & d & d-1 & \dots & 1 & 0 \\ \omega & \omega_d - \omega & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{array} \right\},$$

car $\mu_i = \lambda_i + 1$ pour $i \leq \omega$ et $\mu_i = \lambda_i$ si $i > \omega$, ce qui entraîne $\varpi(\lambda) < \varpi(\mu)$; d'autre part des fonctions monomiales à coefficients $\neq 0$ du type G_ν avec $\nu \leq \mu$ mais $\nu \neq \mu$, c'est-à-dire des

$$G_\nu = \left\{ \begin{array}{cccccc} f & f-1 & \dots & 1 & 0 \\ p_f & p_{f-1} & \dots & p_1 & p_0 \end{array} \right\},$$

avec $1 \leq f \leq d+1$ et, si $f = d+1$, $p_{d+1} < \omega$, d'où $\varpi(\nu) \leq \omega - 1 + (n+1)(d+1)$ et si $f \leq d$, $\varpi(\nu) \leq n + (n+1)d$, d

les cas $\varpi(\nu) < \varpi(\mu) = \omega + (n+1)(d+1)$. (La décomposition exacte de P dans les (G_ν) est proposée en exercice). Il est clair que les coefficients de ces G_ν dans l'expression de P sont dans \mathbb{N} . Nous avons donc prouvé :

LEMME 3

Avec les notations ci-dessus, on a :

$$\sigma_\omega G_\lambda = G_\mu + \sum_{\nu \in \mathfrak{p}_n, \nu \leq \mu} A_\nu G_\nu, \quad \text{où } A_\nu \in \mathbb{N},$$

$$\text{où } G_\mu = \begin{Bmatrix} d+1 & d & d-1 & \dots & 1 & 0 \\ \omega & \omega_d - \omega & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{Bmatrix}$$

et on a $\varpi(\lambda) < \varpi(\mu)$, et $\varpi(\nu) < \varpi(\mu)$ si $A_\nu \neq 0$.

Nous pouvons maintenant en déduire :

THÉOREME X.3.2

(« Théorème fondamental sur les fonctions symétriques »). La \mathbb{C} -algèbre $\mathcal{A}_{\mathfrak{S}_n}$ des polynômes symétriques dans $\mathbb{C}[X_1, \dots, X_n]$ est $\mathbb{C}[\sigma_1, \sigma_2, \dots, \sigma_n]$. Tout $F \in \mathcal{A}_{\mathfrak{S}_n}$ s'écrit de manière unique sous la forme : $F(X_1, \dots, X_n) = S(\sigma_1, \dots, \sigma_n)$ avec $S \in \mathbb{C}[X_1, \dots, X_n]$, et si les coefficients de F sont dans un sous-anneau A de \mathbb{C} , il en est de même de ceux de S .

Démonstration :

On sait que : $\mathbb{C}[\sigma_1, \dots, \sigma_n] \subset \mathcal{A}_{\mathfrak{S}_n}$.

Si $F \in \mathbb{C}[\sigma_1, \dots, \sigma_n]$, l'unicité de $S \in \mathbb{C}[X_1, \dots, X_n]$ tel que $F[X_1, X_2, \dots, X_n] = S(\sigma_1, \sigma_2, \dots, \sigma_n)$ résulte du lemme 2.

Il ne reste plus à démontrer que l'existence de la décomposition. Compte tenu du théorème X.3.1, il suffit de prouver que toute fonction monomiale $F = G_\mu$ ($\mu \in \mathfrak{p}_n$) appartient à $\mathbb{C}[\sigma_1, \dots, \sigma_n]$. C'est évident si $\mu = 0_{\mathbb{N}^n}$ car alors $G_\mu = 1$ et $S = 1$. La récurrence peut démarrer avec $\varpi(\mu) = 1$. Supposons donc l'assertion prouvée pour tous les $F = G_\nu$ tels que $\varpi(\nu) \leq D$, où $D \geq 1$, et considérons $F = G_\mu$ avec $\mu \in \mathfrak{p}_n$ tel que $\varpi(\mu) = D+1$. On a $F = G_\mu = \begin{Bmatrix} e & e-1 & \dots & 1 & 0 \\ \rho_e & \rho_{e-1} & \dots & \rho_1 & \rho_0 \end{Bmatrix}$ avec $e \geq 1$, $\rho_e \geq 1$. Si $e = 1$, on a $F = \sigma_{\rho_e} \in \mathbb{C}[\sigma_1, \dots, \sigma_n]$. Si $e \geq 2$, posons $\omega = \rho_e$, $d = e-1$, $\omega_d = \omega + \rho_{e-1}$, $\omega_i = \rho_i$ pour $i \leq d-1$. Soit G_λ la fonction monomiale $G_\lambda = \begin{Bmatrix} d & d-1 & \dots & 1 & 0 \\ \omega_d & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{Bmatrix}$. Le lemme 3 donne

$$(4) \quad F = \sigma_\omega G_\lambda - \sum_{\nu \in \mathfrak{p}_n, \nu \leq \mu} A_\nu G_\nu, \quad \text{avec } A_\nu \in \mathbb{N}$$

et $\varpi(\nu) < \varpi(\mu)$ chaque fois que $A_\nu \neq 0$; de plus $\varpi(G_\lambda) < \varpi(\mu)$. L'hypothèse de récurrence s'applique donc à G_λ et à tous les G_ν , d'où $F \in \mathbb{C}[\sigma_1, \sigma_2, \dots, \sigma_n]$.

De plus la relation (3) du théorème X.3.1 permet de prouver par récurrence que toute fonction monomiale F appartient en fait à $\mathbb{Z}[\sigma_1, \dots, \sigma_n]$ (sous-anneau de $\mathbb{C}[\sigma_1, \dots, \sigma_n]$ formé des polynômes en les σ_k à coefficients dans \mathbb{Z}). La dernière assertion du théorème X.3.2 s'en déduit aussitôt. ■

Remarque 2 : Soit $F \in \mathbb{C}[X_1, X_2, \dots, X_n]$ symétrique et homogène de degré N . Du fait que chaque σ_k est homogène de degré k , on tire que le polynôme $S \in \mathbb{C}[X_1, \dots, X_n]$ tel que $F(X_1, \dots, X_n) = S(\sigma_1, \dots, \sigma_n)$ est combinaison \mathbb{C} -linéaire de monômes du type $\sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \dots \sigma_n^{\alpha_n}$ avec $\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = N$, ce qui, au moins pour de petites valeurs de N , permet d'alléger la recherche de S par une méthode d'identification.

Remarque 3 : Avec les notations et hypothèses de la *remarque 2*, en supposant $F \neq 0$, le degré partiel d de F en l'une quelconque des variables X_i est évidemment le même, puisque F est symétrique. La récurrence mise en œuvre pour prouver le théorème X.3.2 montre aisément que d est le maximum des ordres des fonctions monomiales intervenant dans la décomposition de F .

Nous avons déjà indiqué que le calcul des A_ν du lemme 3 est possible en toute généralité (cf. exercice 7). Par suite, la preuve du théorème X.3.2 fournit un *algorithme efficace* permettant le calcul *effectif*, de proche en proche, de toute fonction monomiale, et par suite de tout polynôme symétrique, comme polynôme en les (σ_k) .

Exemple 6 : Soit $n \geq 5$. Calculer $G = \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 1 & n-3 \end{Bmatrix}$ en fonction des (σ_k) .

Solution : Par définition $G = \sum_{i < j, (i, j, k) \text{ distincts}} X_i^2 X_j^2 X_k$. La méthode du théorème X.3.2 appliquée à $G = G_\mu$ conduit à considérer

$$G_\lambda = \begin{Bmatrix} 1 & 0 \\ 3 & n-3 \end{Bmatrix},$$

puis à calculer $\sigma_2 G_\lambda$, c'est-à-dire ici $\sigma_2 \sigma_3$, ce qui donne

$$G + 3H + 10K, \text{ avec } H = \begin{Bmatrix} 2 & 1 & 0 \\ 1 & 3 & n-4 \end{Bmatrix} \text{ et } K = \begin{Bmatrix} 1 & 0 \\ 5 & n-5 \end{Bmatrix} = \sigma_5.$$

De même pour avoir H on forme $\sigma_1 \sigma_4 = H + 5\sigma_5$, d'où $H = \sigma_1 \sigma_4 - 5\sigma_5$. Enfin $G = \sigma_2 \sigma_3 - 3\sigma_1 \sigma_4 + 5\sigma_5$. On contrôle le résultat en calculant $G(1, 1, \dots, 1) = \frac{n(n-1)(n-2)}{2}$. Or $\sigma_1 = n, \quad \sigma_2 = \frac{n(n-1)}{2}, \quad \sigma_3 = \frac{n(n-1)(n-2)}{6}, \quad \sigma_4 = \frac{n(n-1)(n-2)(n-3)}{24}, \quad \sigma_5 = \frac{n(n-1)(n-2)(n-3)(n-4)}{120}$.

$\sigma_3 = \binom{n}{3}$, $\sigma_4 = \binom{n}{4}$ et $\sigma_5 = \binom{n}{5}$, et en mettant $\binom{n}{3}$ en facteur on vérifie bien que $3 = \frac{n(n-1)}{2} - \frac{3(n-3)n}{4} + \frac{5(n-3)(n-4)}{20}$. *A priori*, compte tenu des remarques 2 et 3, il aurait pu figurer dans la décomposition de G des termes en $\sigma_1^2 \sigma_3$ et en $\sigma_1 \sigma_2^2$, mais leurs coefficients sont nuls.

Exemple 7 : Soit $n \geq 6$. Calculer $G = \begin{Bmatrix} 2 & 1 & 0 \\ 3 & 0 & n-3 \end{Bmatrix}$ en fonction des (σ_k) . Ici $G = G_\mu = \sum_{i < j < k} X_i^2 X_j^2 X_k^2$. La même méthode conduit à former $\sigma_3^2 = G + 2 U_1 + 6 U_2 + 20 \sigma_6$, avec

$$U_1 = \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 2 & n-4 \end{Bmatrix} \quad \text{et} \quad U_2 = \begin{Bmatrix} 2 & 1 & 0 \\ 1 & 4 & n-5 \end{Bmatrix},$$

puis $\sigma_2 \sigma_4 = U_1 + 4 U_2 + 15 \sigma_6$ et enfin $\sigma_1 \sigma_5 = U_2 + 6 \sigma_6$, d'où l'on tire d'abord $U_2 = \sigma_1 \sigma_5 - 6 \sigma_6$, puis $U_1 = \sigma_2 \sigma_4 - 4 \sigma_1 \sigma_5 + 9 \sigma_6$ et enfin

$$G = \sigma_3^2 - 2 \sigma_2 \sigma_4 + 2 \sigma_1 \sigma_5 - 2 \sigma_6.$$

L'algorithme exposé ci-dessus a l'avantage d'être absolument sûr, mais il est assez lourd. Nous verrons au § suivant que l'exploitation des sommes de Newton permet parfois de l'éviter.

Exercice 1 : Calculer en fonction des (σ_k) les fonctions monomiales suivantes :

$$\begin{aligned} a) F &= \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 2 & 0 & 0 & n-2 \end{Bmatrix} \quad (n \geq 6) & b) F &= \begin{Bmatrix} 4 & 3 & 2 & 1 & 0 \\ 1 & 0 & 0 & 2 & n-3 \end{Bmatrix} \quad (n \geq 6) \\ c) F &= \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 1 & 0 & 3 & n-4 \end{Bmatrix} \quad (n \geq 6) & d) F &= \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 1 & 2 & 0 & n-3 \end{Bmatrix} \quad (n \geq 7). \end{aligned}$$

Exercice 2 : On suppose $n \geq 4$. Soit F le polynôme $\sum_{i < j, k < l} (X_i^2 + X_j^2)(X_k^2 + X_l^2)$. Développer d'abord F dans la base des fonctions monomiales, puis calculer F en fonction des (σ_m) .

Exercice 3 : Pour $n = 3$ calculer en fonction de $\sigma_1, \sigma_2, \sigma_3$ les polynômes suivants :

$$a) F = \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 2 & 1 & 0 & 0 \end{Bmatrix} \quad b) F = \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{Bmatrix} \quad c) F = \begin{Bmatrix} 4 & 2 & 1 & 0 \\ 1 & 2 & 0 & 0 \end{Bmatrix}.$$

Pour $n = 4$, calculer en fonction de $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ les polynômes :

$$\begin{aligned} d) F &= \begin{Bmatrix} 4 & 3 & 2 & 1 & 0 \\ 2 & 2 & 0 & 0 & 0 \end{Bmatrix} & e) F &= \begin{Bmatrix} 2 & 1 & 0 \\ 4 & 0 & 0 \end{Bmatrix} \\ f) F &= (X_1 X_2 + X_3 X_4)(X_1 X_3 + X_2 X_4)(X_1 X_4 + X_2 X_3). \end{aligned}$$

Exercice 4 : Soit $n = 3$ et $\Delta = (X_1 - X_2)(X_2 - X_3)(X_3 - X_4)$.

Développer Δ . Soit U la somme des monômes de coefficient $+1$ (resp. V celle avec les coefficients -1) dans Δ . Calculer $U + V$ et UV en fonction de $\sigma_1, \sigma_2, \sigma_3$ et en déduire Δ^2 .

Réponse : $\Delta^2 = \sigma_1^2 \sigma_2^2 + 18 \sigma_1 \sigma_2 \sigma_3 - 4 \sigma_2^3 - 4 \sigma_1^3 \sigma_3 - 27 \sigma_3^2$.

Exercice 5 : a) Que deviennent les fonctions symétriques élémentaires quand on y remplace X_n par 0 ?

b) Soit $n \in \mathbb{N}$ ($n \geq 2$) et deux entiers p et q dans \mathbb{N}^* tels que $p + q = n$. On pose $Y_i = X_i$ ($1 \leq i \leq p$) et $Z_j = X_{p+j}$ ($1 \leq j \leq q$). Soit $(\sigma_k)_{1 \leq k \leq n}$, $(s_k)_{1 \leq k \leq p}$, et $(t_k)_{1 \leq k \leq q}$ les fonctions symétriques élémentaires respectives des X_i , des Y_i et des Z_j . On convient que $s_0 = t_0 = 1$. Montrer, pour $k \in \llbracket 1, n \rrbracket$: $\sigma_k = \sum_{i+j=k} s_i t_j$.

Exercice 6 : Soit $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$. Montrer que le P.G.O.L. de $F = \sigma_1^{\alpha_1} \sigma_2^{\alpha_2} \dots \sigma_n^{\alpha_n}$ est $X_1^{\lambda_1} X_2^{\lambda_2} \dots X_n^{\lambda_n}$ avec $\lambda_i = \alpha_i + \alpha_{i+1} + \dots + \alpha_n$ ($1 \leq i \leq n$). On posera : $\hat{\alpha} = (\lambda_1, \lambda_2, \dots, \lambda_n)$. En déduire que :

$$F = G_{\hat{\alpha}} + \sum_{\nu \in \mathfrak{p}_n, \nu \not\leq \hat{\alpha}} B_{\nu} G_{\nu},$$

où $B_{\nu} \in \mathbb{N}$ pour tout ν , et déduire de là une autre démonstration du théorème X.3.2 (cette démonstration est plutôt théorique car le calcul des B_{ν} se révèle vite impraticable). Calculer les B_{ν} pour $F = \sigma_1^2 \sigma_2^2$, pour $F = \sigma_3^2$.

Exercice 7 : (calcul exact des A_{ν} du lemme 3).

Soit une fonction monomiale $G_{\lambda} = \left\{ \begin{matrix} d & d-1 & \dots & 1 & 0 \\ \omega_d & \omega_{d-1} & \dots & \omega_1 & \omega_0 \end{matrix} \right\}$, avec $\omega_0 + \omega_1 + \dots + \omega_d = n \geq 2$ et $d \geq 1$. On donne $\omega \in \llbracket 1, d \rrbracket$. Démontrer :

$$\sigma_{\omega} G_{\lambda} = G_{\mu} +$$

$$+ \sum_{\left\{ \begin{matrix} k = (k_d, \dots, k_0) \triangleleft (\omega_d, \omega_{d-1}, \dots, \omega_0) \\ 1 \leq k_d \leq \omega, \sum k_i = \omega \end{matrix} \right\}} U_k \left\{ \begin{matrix} d+1 & d & \dots & 1 & 0 \\ k_d & \omega_d - k_d + k_{d-1} & \dots & \omega_1 - k_1 + k_0 & \omega_0 - k_0 \end{matrix} \right\}$$

$$+ \sum_{\left\{ \begin{matrix} l = (l_{d-1}, \dots, l_0) \triangleleft (\omega_{d-1}, \dots, \omega_0) \\ \sum l_i = \omega \end{matrix} \right\}} V_l \left\{ \begin{matrix} d & d-1 & \dots & 1 & 0 \\ \omega_d + l_{d-1} & \omega_{d-1} - l_{d-1} + l_{d-2} & \dots & \omega_1 - l_1 + l_0 & \omega_0 - l_0 \end{matrix} \right\},$$

avec pour tous k et l : $U_k = \prod_{j=1}^d \binom{\omega_j - k_j + k_{j-1}}{k_{j-1}}$ et $V_l = \prod_{j=1}^d \binom{\omega_j - l_j + l_{j-1}}{l_{j-1}}$ (en posant ici $l_d = 0$).

Indication : On utilisera l'expression des fonctions monomiales donnée après l'exemple 5 (relation (1)).

Application : Faire le calcul explicite lorsque $G_{\lambda} = \left\{ \begin{matrix} 3 & 2 & 1 & 0 \\ 3 & 3 & 1 & n-7 \end{matrix} \right\}$ ($n \geq 7$) et $\omega = 2$.

Exercice 8 : Soit a_1, a_2, \dots, a_n éléments de \mathbb{Z} ($n \geq 1$). On suppose que toutes les racines dans \mathbb{C} du polynôme $F = X^n + a_1 X^{n-1} + \dots + a_n$ sont de module ≤ 1 . Démontrer qu'il existe $N \in \mathbb{N}^*$ tel que toute racine de F soit une racine N -ième de 1. **Indication :** Soit $(\zeta_1, \zeta_2, \dots, \zeta_n)$ une liste des racines de F ; étudier les polynômes $F_k = \prod_{i=1}^n (X - \zeta_i^k)$ pour $k \in \mathbb{N}$ en appliquant le théorème fondamental des fonctions symétriques.

Exercice 9 : Soit $z \in \mathbb{C}^*$ tel que $|z| \neq 1$ et soit $n \in \mathbb{N}^*$. On cherche les valeurs $g_1(z), g_2(z), \dots, g_n(z)$ des polynômes symétriques élémentaires $\sigma_1, \sigma_2, \dots, \sigma_n$ sur la liste $(1, z, z^2, \dots, z^{n-1})$. Pour cela soit T une indéterminée sur \mathbb{C} et

$$F(z, T) = (1 - T)(1 - zT) \dots (1 - z^{n-1}T) \in \mathbb{C}[T].$$

Calculer $F(z, zT)$. En déduire une relation de récurrence simple entre les $g_i(z)$, et enfin donner les valeurs des $g_i(z)$. **Réponse :** $g_i(z) = z^{\frac{i(i-1)}{2}} \frac{(1 - z^n) \dots (1 - z^{n-i+1})}{(1 - z) \dots (1 - z^i)}$

Application : Soit $\theta \in \mathbb{R} \setminus \mathbb{Q}$ et $\zeta = e^{i\theta}$. Calculer les fonctions symétriques élémentaires de $\zeta^{1-m}, \zeta^{3-m}, \dots, \zeta^{m-3}, \zeta^{m-1}$ pour $m \in \mathbb{N}^*, m \geq 2$.

Réponse : $\sigma_k = \frac{\sin m\theta \sin (m-1)\theta \dots \sin (m-k+1)\theta}{\sin \theta \sin 2\theta \dots \sin k\theta}$.

Exercice 10 : (Algèbre des polynômes alternés).

On suppose acquis les résultats de l'exercice 11 du § X.1. On note \mathcal{A} la \mathbb{C} -algèbre des *polynômes alternés* dans $\mathbb{C}[X_1, \dots, X_n]$ et \mathcal{S} celle des polynômes symétriques. Δ désigne le polynôme de Vandermonde $\prod_{1 \leq i < j \leq n} (X_j - X_i)$. Prouver :

- $\Delta \in \mathcal{A} \setminus \mathcal{S}$.
- \mathcal{A} contient la \mathbb{C} -algèbre \mathcal{A}' formée des polynômes $S + \Delta T$, $S \in \mathcal{S}$, $T \in \mathcal{S}$.
- Si $P \in \mathcal{A} \setminus \mathcal{S}$, la \mathbb{G}_n -orbite de P (pour l'action $(\sigma, Q) \mapsto \sigma * Q$) contient exactement deux éléments U et V , et l'on a : $U + V \in \mathcal{S}$ et $U - V = \Delta W$ avec $W \in \mathcal{S}$ (utiliser l'exercice cité).
- En déduire que $\mathcal{A}' = \mathcal{A}$ et que $\mathcal{S} \times \mathcal{S} \rightarrow \mathcal{A}$, $(S, T) \mapsto S + \Delta T$ est une bijection.

Exercice 11 : Soit trois entiers n, k, r avec $n \geq 2, r \in \llbracket 1, n \rrbracket$. On note $(\zeta_1, \dots, \zeta_n)$ une liste des racines n -ièmes de 1 dans \mathbb{C} . Calculer le nombre $\sum_{j_1 < j_2 < \dots < j_r} \zeta_{j_1}^k \dots \zeta_{j_r}^k$.

§ X.4 FORMULES DE NEWTON

Nous reprenons les notations du § X.3, avec un entier $n \in \mathbb{N}^*$. Les sommes de Newton $S_d = \sum_{i=1}^n X_i^d$ se présentent souvent dans les calculs de fonctions polynomiales symétriques, et nous allons voir qu'on passe aisément des fonctions symétriques élémentaires σ_k aux S_d , et *vice versa*.

THÉORÈME X.4.1 (Formules de Newton)

$$\begin{aligned} \text{(I)} \quad & \text{si } k \leq n: \quad S_k - \sigma_1 S_{k-1} + \cdots + (-1)^{k-1} \sigma_{k-1} S_1 + \\ & \quad \quad \quad + (-1)^k \times k \sigma_k = 0 \\ \text{(II)} \quad & \text{si } k > n: \quad S_k - \sigma_1 S_{k-1} + \cdots + (-1)^n S_{k-n} \sigma_n = 0. \end{aligned}$$

En convenant que $S_0 = n$ il y a concordance de (I) et (II) pour $k = n$.

Démonstration :

Pour obtenir (II) il suffit de développer, pour chaque $i \in \llbracket 1, n \rrbracket$, le produit

$$\prod_{i=1}^n (X_i - X_j) = X_i^n - \sigma_1 X_i^{n-1} + \cdots + (-1) \sigma_n = 0,$$

de multiplier par X_i^{k-n} ($k \geq n$) et d'additionner les n relations obtenues en faisant varier i dans $\llbracket 1, n \rrbracket$. Pour démontrer (I), fixons i

notons s_1, s_2, \dots, s_{n-1} les fonctions symétriques élémentaires des lettres $(X_j)_{j \neq i}$; complétons avec $s_0 = 1, s_n = 0$. Pour $k \in \llbracket 1, n \rrbracket$, on a : $\sigma_k = X_i s_{k-1} + s_k$, et en particulier $\frac{\partial \sigma_k}{\partial X_i} = s_{k-1}$. D'où

$$X_i \frac{\partial \sigma_k}{\partial X_i} = \sigma_k - s_k = \sigma_k - \frac{\partial \sigma_{k+1}}{\partial X_i}.$$

Alors une récurrence immédiate permet de déduire de là que :

$$(1) \quad \frac{\partial \sigma_k}{\partial X_i} = (-1)^{k-1} X_i^{k-1} + (-1)^{k-2} \sigma_1 X_i^{k-2} + \dots + \sigma_{k-1}.$$

En multipliant les relations (1) par X_i , et en sommant pour $i \in \llbracket 1, n \rrbracket$ il vient :

$$(2) \quad \sum_{i=1}^n X_i \frac{\partial \sigma_k}{\partial X_i} = (-1)^{k-1} S_{k-1} + (-1)^{k-2} \sigma_1 S_{k-2} + \dots + \sigma_{k-1} S_1.$$

En utilisant la *formule d'Euler* vue au § X.2, qui s'applique car σ_k est homogène de degré k , on peut remplacer $\sum_{i=1}^n X_i \frac{\partial \sigma_k}{\partial X_i}$ par $k\sigma_k$ au premier membre de (2), et l'on trouve finalement (I). ■

Exemple 1 : Pour $n \geq 3$, calculons S_3 . Les relations (I) donnent : $S_1 = \sigma_1$, $S_2 - \sigma_1 S_1 + 2\sigma_2 = 0$, d'où $S_2 = \sigma_1^2 - 2\sigma_2$, et $S_3 - \sigma_1 S_2 + \sigma_2 S_1 - 3\sigma_3 = 0$, d'où

$$S_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3.$$

Pour $n \geq 4$, on obtiendrait de même, en reportant de proche en proche S_1, S_2 et S_3 dans la formule $S_4 - \sigma_1 S_3 + \sigma_2 S_2 - \sigma_3 S_1 + 4\sigma_4 = 0$:

$$S_4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2 - 4\sigma_4.$$

Exemple 2 : Inversement, ces formules fournissent les σ_k en fonction de S_1, S_2, \dots, S_n . Par exemple

$$\sigma_1 = S_1, \quad \sigma_2 = \frac{1}{2} (S_1^2 - S_2), \quad \sigma_3 = \frac{1}{6} (2S_3 + S_1^3 - 3S_1 S_2), \text{ etc.}$$

Plus généralement :

COROLLAIRE

La \mathbb{C} -algèbre \mathcal{A}_{∞} des polynômes symétriques est $\mathbb{C}[S_1, S_2, \dots, S_n]$ et S_1, \dots, S_n sont algébriquement libres sur \mathbb{C} . Donc, si F est un polynôme symétrique, il existe un et un seul polynôme $T \in \mathbb{C}[X_1, \dots, X_n]$ tel que $F(X_1, \dots, X_n) = T(S_1, \dots,$

Démonstration :

L'inclusion $\mathbb{C}[S_1, \dots, S_n] \subset \mathcal{A}_{\mathfrak{S}_n}$ est évidente.

En résolvant les relations (I) de proche en proche par rapport aux σ_k , il est clair que $\sigma_k \in \mathbb{C}[S_1, \dots, S_n]$ pour tout k , $1 \leq k \leq n$, d'où $\mathbb{C}[\sigma_1, \dots, \sigma_n] = \mathcal{A}_{\mathfrak{S}_n} \subset \mathbb{C}[S_1, \dots, S_n]$, et finalement $\mathbb{C}[S_1, \dots, S_n] = \mathcal{A}_{\mathfrak{S}_n}$.

L'application $\mathbb{C}^n \rightarrow \mathbb{C}^n$, $(x_1, \dots, x_n) \mapsto (S_1(x_1, \dots, x_n), \dots, S_n(x_1, \dots, x_n))$ est surjective, car si on remplace dans (I) S_1, \dots, S_n par y_1, \dots, y_n ($y_i \in \mathbb{C}$), on en déduit de manière unique s_1, s_2, \dots, s_n dans \mathbb{C} tels que

$$y_k - s_1 y_{k-1} + \dots + (-1)^k k s_k = 0, \quad (1 \leq k \leq n).$$

Puis on a $(x_1, \dots, x_n) \in \mathbb{C}^n$ tel que $s_k = \sigma_k(x_1, \dots, x_n)$ pour $1 \leq k \leq n$ (cf. lemme 2, § X.3), d'où enfin $y_k = S_k(x_1, \dots, x_n)$ pour $1 \leq k \leq n$. Cette surjectivité entraîne l'indépendance algébrique sur \mathbb{C} de S_1, \dots, S_n (même preuve que lemme 2, § X.3). ■

Formules de Waring ⁽¹⁾

Soit T une indéterminée sur \mathbb{C} . Si $(x_1, \dots, x_n) \in \mathbb{C}^n$, les séries formelles $\text{Log}[(1 - x_1 T) \dots (1 - x_n T)]$ et $\sum_{i=1}^n \text{Log}(1 - x_i T)$ sont égales (cf. § VIII.5). Or

$$(1 - x_1 T) \dots (1 - x_n T) = 1 - \tilde{\sigma}_1 T + \tilde{\sigma}_2 T^2 + \dots + (-1)^n \tilde{\sigma}_n T^n,$$

et $\sum_{i=1}^n \text{Log}(1 - x_i T) = - \sum_{i=1}^n \sum_{p=1}^{\infty} \frac{x_i^p}{p} T^p = - \sum_{p=1}^{\infty} \frac{\tilde{S}_p}{p} T^p$, où $\tilde{\sigma}_k = \sigma_k(x_1, \dots, x_n)$ et $\tilde{S}_p = S_p(x_1, \dots, x_n)$ avec $k \in \llbracket 1, n \rrbracket$ et $p \in \mathbb{N}^*$. On a donc, par superposition de la série formelle $\text{Log}(1 - T)$ et $\tilde{\sigma}_1 T - \tilde{\sigma}_2 T^2 + \dots + (-1)^n \tilde{\sigma}_n T^n$:

$$\text{Log}[(1 - x_1 T) \dots (1 - x_n T)] = - \sum_{k \geq 1} \frac{1}{k} (\tilde{\sigma}_1 T - \tilde{\sigma}_2 T^2 + \dots + (-1)^{n-1} \tilde{\sigma}_n T^n)$$

= (en développant chaque terme avec la formule du multinôme)

$$\begin{aligned} &= - \sum_{k \geq 1} \frac{1}{k} \left(\sum_{\alpha_1 + \dots + \alpha_n = k} (-1)^k \frac{k!}{\alpha_1! \dots \alpha_n!} \times \right. \\ &\quad \left. \times (-1)^{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n} \tilde{\sigma}_1^{\alpha_1} \tilde{\sigma}_2^{\alpha_2} \dots \tilde{\sigma}_n^{\alpha_n} T^{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n} \right) \\ &= - \sum_{p \in \mathbb{N}^*} (-1)^p T^p \left(\sum_{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = p} (-1)^{\sum \alpha_i} \frac{(\alpha_1 + \dots + \alpha_n - 1)!}{\alpha_1! \dots \alpha_n!} \tilde{\sigma}_1^{\alpha_1} \dots \tilde{\sigma}_n^{\alpha_n} \right) \end{aligned}$$

⁽¹⁾ Edward Waring (1734-1798), mathématicien anglais, auteur des *Miscellanea analytica* (1762) et des *Meditationes algebraicae* (Cambridge, 1770).

On obtient donc, par identification des séries $\sum_{i=1}^n \text{Log}(1 - x_i T)$ et $\text{Log} \prod_{i=1}^n (1 - x_i T)$:

$$(\forall p \geq 1) \quad \frac{1}{p} \tilde{S}_p = (-1)^p \sum_{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = p} (-1)^{\sum \alpha_i} \frac{(\alpha_1 + \dots + \alpha_n - 1)!}{\alpha_1! \dots \alpha_n!} \tilde{\sigma}_1^{\alpha_1} \dots \tilde{\sigma}_n^{\alpha_n},$$

ce qui fournit les \tilde{S}_p en fonction des $\tilde{\sigma}_i$. Ces relations étant vraies pour tout $(x_1, x_2, \dots, x_n) \in \mathbb{C}^n$, ce sont des identités polynomiales entre les S_k et les σ_k :

(3) $(\forall p \geq 1)$

$$\boxed{\frac{1}{p} S_p = (-1)^p \sum_{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = p} (-1)^{\sum \alpha_i} \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_n - 1)!}{\alpha_1! \alpha_2! \dots \alpha_n!} \sigma_1^{\alpha_1} \dots \sigma_n^{\alpha_n}}.$$

Ces relations (3) s'appellent **formules de Waring** et résolvent le problème du calcul des S_p en fonction des σ_k plus rapidement qu'avec les formules de Newton où il fallait opérer de proche en proche.

Exemple 3 : Soit à calculer S_5 en utilisant la formule (3), avec $n \geq 5$. Il suffit de former un tableau des $(\alpha_i) \in \mathbb{N}^n$ tels que $\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = 5$, en remarquant que $\alpha_k = 0$ si $k \geq 6$, en réservant une colonne pour l'exposant $r = \sum \alpha_i$ et une pour le coefficient correspondant $\frac{(\alpha_1 + \alpha_2 + \dots + \alpha_5 - 1)!}{\alpha_1! \dots \alpha_5!}$, ce qui donne

α_1	α_2	α_3	α_4	α_5	r	$\frac{(r-1)!}{\Pi(\alpha_i)!}$
0	0	0	0	1	1	1
1	0	0	1	0	2	1
0	1	1	0	0	2	1
2	0	1	0	0	3	1
1	2	0	0	0	3	1
3	1	0	0	0	4	1
5	0	0	0	0	5	1/5

d'où $\frac{1}{5} S_5 = \frac{1}{5} \sigma_1^5 - \sigma_1^3 \sigma_2 + \sigma_1 \sigma_2^2 + \sigma_1^2 \sigma_3 - \sigma_2 \sigma_3 - \sigma_1 \sigma_4 + \sigma_5.$

De la même manière, on obtient :

$$1 - \tilde{\sigma}_1 T + \tilde{\sigma}_2 T^2 + \dots + (-1)^n \tilde{\sigma}_n T^n = \exp \left(\text{Log} \prod_{i=1}^n (1 - x_i T) \right) \quad (\text{cf. § VIII.5})$$

$$= \exp \left(- \sum_{p \geq 1} \frac{1}{p} \tilde{S}_p T^p \right) = \exp(-\tilde{S}_1 T) \times \exp \left(\frac{-1}{2} \tilde{S}_2 T^2 \right) \dots \times \exp \left(\frac{-1}{n} \tilde{S}_p T^p \right)$$

et l'identification, à partir de $\exp(U) = \sum_{k \geq 0} \frac{1}{k!} U^k$, conduit à

$$(4) \quad (-1)^p \sigma_p = \sum_{\alpha_1 + 2\alpha_2 + \dots + n\alpha_n = p} \frac{(-1)^{\sum_{i=1}^p \alpha_i}}{\alpha_1! \dots \alpha_p! 2^{\alpha_2} \dots p^{\alpha_p}} S_1^{\alpha_1} S_2^{\alpha_2} \dots S_p^{\alpha_p}$$

qui résout le problème du calcul des σ_k en fonction des S_k ($1 \leq k \leq n$).

Exemple 4 : Pour $n \geq 5$ soit à calculer $F = \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 1 & 0 & 2 & n-3 \end{Bmatrix}$ en fonction des S_k .

Solution : On a $F = \sum_{j < k, (i, j, k) \text{ distincts}} X_i^3 X_j X_k$. On est conduit à calculer $S_3 \sigma_2 = F + H$ avec $H = \sum_{i \neq j} X_i^4 X_j$, d'où $S_4 \sigma_1 = S_5 + H$. Cela donne $H = S_4 \sigma_1 - S_5$, d'où $F = S_3 \sigma_2 - S_4 \sigma_1 + S_5$.

En utilisant les valeurs $\sigma_1 = S_1$ et $2 \sigma_2 = S_1^2 - S_2$, on en déduit F en fonction des S_k : $F = \frac{1}{2} S_1^2 S_3 - \frac{1}{2} S_2 S_3 - S_1 S_4 + S_5$.

Si maintenant on utilise les résultats des exemples 1 et 3 on peut aussi exprimer F en fonction des σ_k :

$$F = \sigma_1^2 \sigma_3 - \sigma_1 \sigma_4 - 2 \sigma_2 \sigma_3 + 5 \sigma_5.$$

Exercice 1 : Soit $n \in \mathbb{N}^*$, $(x_1, \dots, x_n) \in \mathbb{C}^n$ et T une indéterminée sur \mathbb{C} . On pose $P = \prod_{j=1}^n (1 - x_j T) \in \mathbb{C}[T]$. On note $\tilde{\sigma}_k$ (resp. \tilde{S}_k) les valeurs au point (x_1, \dots, x_n) des fonctions σ_k (resp. S_k). Développer en série formelle dans $\mathbb{C}[[T]]$ la fraction $-\frac{P'(T)}{P(T)} = \sum_{j=1}^n \frac{x_j}{1 - x_j T}$. En déduire : $-P'(T) = P(T) \left(\sum_{k \geq 0} \tilde{S}_{k+1} T^k \right)$ et obtenir ainsi une nouvelle démonstration des formules de Newton.

Exercice 2 : Calculer S_6 en fonction des $(\sigma_k)_{1 \leq k \leq 6}$ pour $n \geq 6$. Calculer également $\sigma_1, \dots, \sigma_6$ en fonction de S_1, \dots, S_6 en utilisant les formules de Waring.

Exercice 3 : Soit des entiers m, p, q, r avec $m > p > q > r \geq 1$. Calculer en fonction des S_k , pour $n \geq 4$, la fonction $F = \sum_{(i, j, k, l) \text{ distincts}} X_i^m X_j^p X_k^q X_l^r$.

Exercice 4 : Soit a, b, c dans \mathbb{Z}^* des entiers premiers entre eux deux à deux, non congrus à 0 mod (17), et tels que $a^{17} + b^{17} + c^{17} = 0$.

a) Montrer d'abord l'existence de 3 entiers x, y, z dans \mathbb{Z}^* , non congrus à 0 mod (17), tels que $x + y + z = 0$ et $x^{17} + y^{17} + z^{17} \equiv 0 \pmod{289}$.

b) Montrer qu'on peut même faire en sorte dans les égalités précédentes que x, y et z soient premiers entre eux deux à deux.

c) On note $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$ les valeurs des fonctions $\sigma_1, \sigma_2, \sigma_3$ au point (x, y, z) , et (\tilde{S}_k) les sommes de Newton de x, y, z . Calculer \tilde{S}_{17} à l'aide des $\tilde{\sigma}_i$. En déduire que $\tilde{\sigma}_2 \tilde{\sigma}_3 [\tilde{\sigma}_2^3 + 5 \tilde{\sigma}_3^2] [\tilde{\sigma}_2^3 + 7 \tilde{\sigma}_3^2] \equiv 0 \pmod{17}$ et en tirer une contradiction avec !

départ. On obtient ainsi une démonstration du « premier cas » du grand théorème de Fermat avec $p = 17$ (le « second cas », plus difficile, est celui où l'un des 3 nombres a , b ou c serait multiple de 17).

Exercice 5 : En utilisant les formules de Newton, résoudre dans \mathbb{C} les systèmes suivants :

a) $x^2 + y^2 + z^2 = 2$, $x^3 + y^3 + z^3 = 2$, $x^4 + y^4 + z^4 = 2$

b) $x^2 + y^2 + z^2 = 0$, $x^4 + y^4 + z^4 = 0$, $x^5 + y^5 + z^5 = 0$

c) $x + y + z = 1$, $x^2 + y^2 + z^2 = 9$, $\frac{1}{x} + \frac{1}{y} + \frac{1}{z} = 1$.

Exercice 6 : Soit m et n dans \mathbb{N}^* et $(x_1, \dots, x_m) \in \mathbb{C}^m$, $(y_1, \dots, y_n) \in \mathbb{C}^n$. On suppose $(\forall k \in \mathbb{N}^*) \sum_{i=1}^m x_i^k = \sum_{j=1}^n y_j^k$. Comparer les listes (x_i) et (y_j) .

Exercice 7 : On suppose $n \geq 4$. Démontrer, pour $k \in \mathbb{N}^*$:

$$\sum_{1 \leq i_1 < i_2 < i_3 \leq n} X_{i_1}^k X_{i_2}^k X_{i_3}^k = \frac{1}{6} (S_k^3 - 3 S_k S_{2k} + 2 S_{3k})$$

et $\sum_{1 \leq i_1 < i_2 < i_3 < i_4 \leq n} X_{i_1}^k X_{i_2}^k X_{i_3}^k X_{i_4}^k = \frac{1}{24} [S_k^4 - 6 S_k^2 S_{2k} + 8 S_k S_{3k} + 3 S_{2k}^2 - 6 S_{4k}]$.

Exercice 8 : On donne $k \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$. Pour $r \in \llbracket 1, n \rrbracket$, soit $L_r = \sum_{i_1 < i_2 < \dots < i_r} X_{i_1}^k X_{i_2}^k \dots X_{i_r}^k$. Calculer à l'aide de la formule de Waring (4) L_r en fonction de $S_k, S_{2k}, \dots, S_{rk}$, et retrouver ainsi les résultats de l'exercice 7.

Exercice 9 : Soit $n \in \mathbb{N}^*$. Pour chaque polynôme $P \in \mathbb{C}[X_1, \dots, X_n]$, on note $\frac{\partial P}{\partial \sigma_k}$ (resp. $\frac{\partial P}{\partial S_k}$) le polynôme $\frac{\partial P}{\partial X_k}(\sigma_1, \sigma_2, \dots, \sigma_n)$ (resp. $\frac{\partial P}{\partial X_k}(S_1, S_2, \dots, S_n)$).

a) Soit r et k dans \mathbb{N} , $r \in \llbracket 1, n \rrbracket$, $r + k \leq n$. On note P le polynôme tel que $\sigma_{r+k} = P(S_1, S_2, \dots, S_n)$, donné par (4). Prouver : $\frac{\partial P}{\partial S_n} = \frac{(-1)^{n-1}}{n}$ et, si $k \geq 1$:

$$\frac{\partial P}{\partial S_r} = \frac{(-1)^{r-1}}{r} \sigma_k.$$

b) Soit $P \in \mathbb{C}[X_1, \dots, X_n]$ et soit Q le polynôme tel que $P(\sigma_1, \dots, \sigma_n) = Q(S_1, \dots, S_n)$. Utiliser a) et l'exercice 4 du § X.2 pour démontrer que :

$$(\forall r \in \llbracket 1, n \rrbracket) \quad \frac{\partial Q}{\partial S_r} = \frac{(-1)^{r-1}}{r} \left(\frac{\partial P}{\partial \sigma_r} + \sigma_1 \frac{\partial P}{\partial \sigma_{r+1}} + \dots + \sigma_{n-r} \frac{\partial P}{\partial \sigma_n} \right).$$

c) *Application.* On suppose $n \geq 6$. Soit la fonction monomiale $G = \begin{Bmatrix} 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & n-3 \end{Bmatrix} = \sum_{i < j < k} X_i^3 X_j^2 X_k$. Etablir successivement : $G = S_1 S_2 S_3 - S_1 S_5 - S_3^2 - S_2 S_4 + 2 S_6$, puis

$$G = \lambda_0 \sigma_1^6 + \lambda_1 \sigma_5 \sigma_1 + \lambda_2 \sigma_4 \sigma_2 + \lambda_3 \sigma_4 \sigma_1^2 + \lambda_4 \sigma_3^2 + \lambda_5 \sigma_1 \sigma_2 \sigma_3 + \lambda_6 \sigma_2^3,$$

avec des $\lambda_i \in \mathbb{Z}$. Utiliser b) pour trouver des équations liant les σ_i , et en déduire G en fonction des σ_i . (Réponse : $\lambda_0 = -12$, $\lambda_1 = 7$, $\lambda_2 = 4$, $\lambda_3 = -3$, $\lambda_4 = -3$, $\lambda_5 = 1$, $\lambda_6 = 0$).

§ X.5 ÉQUATIONS ALGÈBRIQUES. ÉQUATIONS DE DEGRÉ 3

Résolution d'une équation algébrique

Nous appellerons *équation algébrique* une équation du type $f(z) = 0$ où f est un polynôme donné de $\mathbb{C}[X]$ et où l'inconnue z est un nombre complexe, élément de \mathbb{C} . Tout $z \in \mathbb{C}$ tel que $f(z) = 0$ est appelé une *racine* de l'équation. Nous avons vu au § VI.6 comment résoudre une équation du second degré. Mais nous savons plus généralement que, pour $n \in \mathbb{N}^*$, en vertu du théorème de d'Alembert, le polynôme

$$P = X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 \quad (a_i \in \mathbb{C})$$

est dissocié sur \mathbb{C} , et s'écrit donc $P = \prod_{k=1}^n (X - z_k)$, où (z_1, z_2, \dots, z_n) est une liste des racines de P .

La question qui se pose à qui veut résoudre l'équation $P(z) = 0$ est dès lors d'exprimer les z_k connaissant a_1, a_2, \dots, a_n . Cela peut s'entendre de diverses façons. Par exemple, on peut chercher des algorithmes permettant d'obtenir avec toute la précision désirée des valeurs approchées des z_k : ce problème a été étudié depuis fort longtemps et a reçu des solutions variées, souvent ingénieuses. Si nous ne l'abordons pas ici, c'est pour le réserver au Cours d'Analyse (cf. tome 4). On peut aussi chercher à exprimer les z_k par une formule théorique en fonction des a_k , cette formule devant indiquer le moyen de parvenir aux z_k à l'aide d'une suite *finie* d'opérations « autorisées » portant sur les a_k (cf. la résolution de l'équation du second degré). Il s'agit alors de bien s'accorder sur un petit nombre d'opérations algébriques suffisamment simples pour être communément autorisées : il y aura bien sûr *les quatre opérations rationnelles* (addition, soustraction, multiplication et division) et on accepte généralement *l'extraction d'une racine k -ième d'un nombre complexe*, avec $k \in \mathbb{N}^*$. Les équations algébriques dont il est possible d'exprimer les racines par une *succession finie* d'opérations de ce type, à partir d'une part de *nombre rationnels*, d'autre part *des a_k donnés*, sont dites **résolubles par radicaux**.

Par exemple, au sens qui vient d'être précisé ci-dessus, il est évident que n'importe quelle équation binôme du type $z^n = a$ où $n \in \mathbb{N}^*$ et $a \in \mathbb{C}$ est résoluble par radicaux (cf. § VI.6). Mais il a fallu les travaux d'Abel et de Galois pour démontrer que l'équation « générale » de degré $n \geq 5$ n'est pas résoluble par radicaux dans le sens ci-dessus.

On peut même s'imposer des conditions encore plus draconiennes en n'acceptant sous chaque radical que des nombres réels positifs et l'on parlera alors d'équations **résolubles par radicaux réels**, comme le sont toutes les équations de degré 1 ou 2 (cf. § VI.6), si l'on peut expliciter les racines par une succession finie d'additions, soustractions, multiplications, divisions, extractions d'une racine k -ième d'un nombre réel ≥ 0 , toutes ces opérations portant sur des rationnels, sur les a_k , sur i , et sur des nombres obtenus dans une étape précédente. Pour prendre comme exemple simple celui de la division du cercle en n arcs égaux

l'équation $X^n - 1 = 0$, on peut prouver que les équations $X^3 - 1 = 0$, $X^5 - 1 = 0$, $X^{17} - 1$ (Gauss) sont résolubles par radicaux réels (et même dans ces trois cas en n'utilisant que des racines carrées, ce qui entraîne la possibilité de construire les racines « à la règle et au compas »), alors que les équations $X^7 - 1 = 0$, $X^{11} - 1 = 0$ ne le sont pas (cf. exercice 23, § VII.4), mais nous ne prétendons pas dans les pages qui suivent résoudre le problème de reconnaître, une équation algébrique étant donnée, si elle est ou non résoluble par radicaux. Cela exigerait des connaissances en théorie des groupes nettement supérieures au niveau de cet ouvrage (théorie fine de Galois), et même pour les équations du troisième degré le problème de la résolution par radicaux réels recèle encore des questions difficiles. Les définitions données ci-dessus doivent donc être comme un simple guide pour ce qui suit, sans autre prétention mathématique.

Fonctions symétriques des racines

Soit (x_1, x_2, \dots, x_n) une liste des racines du polynôme

$$P = X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{C}[X] \quad (n \geq 1).$$

Si $S \in \mathbb{C}[X_1, \dots, X_n]$ est un *polynôme symétrique*, nous savons que $S = Q(\sigma_1, \dots, \sigma_n)$, où $Q \in \mathbb{C}[X_1, \dots, X_n]$. La valeur de S sur la liste (x_1, \dots, x_n) est donc (compte tenu de l'étude du § VII.5) : $S(x_1, \dots, x_n) = Q(-a_1, a_2, \dots, (-1)^n a_n)$. Cette valeur est indépendante de la numérotation (x_1, \dots, x_n) choisie pour les racines, comme il fallait évidemment s'y attendre puisque S est symétrique.

Un des polynômes symétriques S particulièrement intéressant est $D(X_1, \dots, X_n) = \prod_{1 \leq i < j \leq n} (X_i - X_j)^2$, le *discriminant général d'ordre n* (voir § X.3) ; sa valeur sur (x_1, \dots, x_n) s'appelle le **discriminant du polynôme P** . Il est clair que **ce discriminant est non nul ssi P possède n racines simples**.

Le calcul du discriminant général en fonction de $\sigma_1, \dots, \sigma_n$ est très lourd, mais sa valeur pour des polynômes particuliers est plus facile à obtenir. Voici une remarque très simple dont on pourra s'aider : puisque $P(X) = \prod_{k=1}^n (X - x_k)$ on a : $P'(x_k) = \prod_{l \neq k} (x_k - x_l)$ pour tout $k \in \llbracket 1, n \rrbracket$,

d'où

$$\prod_{k=1}^n P'(x_k) = \varepsilon_n \prod_{k < l} (x_k - x_l)^2 = \varepsilon_n D(x_1, \dots, x_n), \quad \text{avec} \quad \varepsilon_n = (-1)^{\frac{n(n-1)}{2}}$$

Exemple 1 : Soit à calculer le discriminant D_n de $P = X^n - 1$. D'après ce qui précède, on a :

$$D_n = \varepsilon_n \prod_{\zeta \in \mu_n} P'(\zeta) = \varepsilon_n \prod_{\zeta \in \mu_n} n \zeta^{n-1} = \varepsilon_n n^n \prod_{\zeta \in \mu_n} \zeta = (-1)^{\frac{(n-1)(n+2)}{2}} n^n.$$

Exemple 2 : Calculons le discriminant D de $P = X^n + pX + q$, $n \geq 2$, $q \neq 0$. On a $D = \varepsilon_n \prod_{k=1}^n (nx_k^{n-1} + p)$, où (x_k) est une liste des racines de P .

Les x_k sont $\neq 0$ et on a : $nx_k^{n-1} + p = \frac{-qn}{x_k} + (1-n)p$. L'équation dont

une liste des racines est $\left(\frac{1}{x_k}\right)_{1 \leq k \leq n}$ est $Q = 0$ avec $Q = qX^n + pX^{n-1} + 1$.

Posant $Z = p(1-n) - qnX$, on obtient l'équation $R = 0$ dont une liste des racines est $\left(p(1-n) - qn \frac{1}{x_k}\right)_k = (nx_k^{n-1} + p)_k$, à savoir

$$R = (p(1-n) - Z)^n + pn(p(1-n) - Z)^{n-1} + n^n q^{n-1} = 0,$$

d'où en prenant le produit des racines de R :

$$D = [n^n q^{n-1} + p^n(1-n)^{n-1}] (-1)^{\frac{n(n-1)}{2}}.$$

Equation du troisième degré

L'équation générale de degré 3 : $x^3 + ax^2 + bx + c = 0$ se ramène, par le changement d'inconnue $x \mapsto x - \frac{a}{3}$ à la forme réduite :

$$(1) \quad x^3 + px + q = 0 \quad \text{avec} \quad (p, q) \in \mathbb{C}^2.$$

Montrons d'abord que (1) est *résoluble par radicaux* : pour cela cherchons s'il existe $(\alpha, \beta) \in \mathbb{C}^2$ tel que l'on ait, dans $\mathbb{C}[X]$, l'identité :

$$X^3 + pX + q = [X - (\alpha + \beta)][X - (\alpha j + \beta j^2)][X - (\alpha j^2 + \beta j)].$$

La condition nécessaire et suffisante sur le couple (α, β) est fournie par le système :

$$(2) \quad \begin{cases} p = -3\alpha\beta \\ q = -(\alpha^3 + \beta^3). \end{cases}$$

,

Ce système (2) entraîne le suivant :

$$(3) \quad \begin{cases} \alpha^3 \beta^3 = -\frac{p^3}{27} \\ \alpha^3 + \beta^3 = -q \end{cases}$$

et les solutions de (3) sont évidemment les couples (α, β) tels que $\alpha^3 = U$ et $\beta^3 = V$, ou $\alpha^3 = V$ et $\beta^3 = U$, en désignant par (U, V) une liste des racines de l'équation de degré 2 ci-dessous

$$(4) \quad X^2 + qX - \frac{p^3}{27} = 0$$

que l'on sait résoudre par radicaux (et même par radicaux réels). Pour résoudre (1), il nous suffit de trouver *un* couple (α, β) solution de (2). Soit donc $(\alpha_0, \beta_0) \in \mathbb{C}^2$ tel que $\alpha_0^3 = U$ et $\beta_0^3 = V$. On a donc $\alpha_0 \beta_0 = \frac{-1}{3} j^k p$, avec $k \in \llbracket 0, 2 \rrbracket$, et $\alpha_0^3 + \beta_0^3 = -q$. Il suffit de poser $\alpha = \alpha_0$ et $\beta = \beta_0 j^{2k}$ pour être sûr que (α, β) vérifie le système (2), et par suite une liste des racines de (1) sera : $\alpha + \beta, \alpha j + \beta j^2, \alpha j^2 + \beta j$. Cette liste s'exprime par radicaux à partir des coefficients p, q et de rationnels. En effet si nous notons $\Delta = q^2 + \frac{4}{27} p^3$ le discriminant de (4), et δ l'une quelconque des racines carrées de Δ dans \mathbb{C} , on peut prendre pour couple (U, V) : $U = \frac{-q + \delta}{2}, V = \frac{-q - \delta}{2}$. Si maintenant on choisit parmi les racines cubiques de U et de V un couple vérifiant $\sqrt[3]{U} \times \sqrt[3]{V} = \frac{-p}{3}$, la liste des racines de (1) peut s'écrire :

(5) $\alpha + \beta, \alpha j + \beta j^2, \alpha j^2 + \beta j$ avec :

$$\alpha = \sqrt[3]{\frac{-q + \delta}{2}}, \quad \beta = \sqrt[3]{\frac{-q - \delta}{2}} \quad \text{et} \quad \delta^2 = \frac{1}{27} (4p^3 + 27q^2).$$

Les relations (5) constituent les **formules de Cardan** ⁽¹⁾ qui présentent surtout un intérêt théorique car elles ne servent pratiquement jamais pour la résolution numérique de l'équation (1).

Plaçons-nous dans le cas général où $U \neq V$ et $p \neq 0$ (c.-à-d. $\Delta \neq 0$ et $UV \neq 0$). Le système (3) admet 18 solutions (α_0, β_0) , mais la contrainte $\alpha\beta = \frac{-p}{3}$ montre que le système (2) n'admet plus que 6 solutions, ce qui donne 6 manières d'écrire les racines de (1) sous la forme (5). Soit (x_1, x_2, x_3) la liste obtenue avec la solution (α, β) de (2) ; les 6 solutions de (2) sont alors $\{(\alpha \zeta, \beta \zeta^2)\}_{\zeta \in \mu_3}$ et $\{(\beta \zeta, \alpha \zeta^2)\}_{\zeta \in \mu_3}$. Comme il fallait s'y attendre, les listes obtenues à partir de ces 6 solutions sont exactement les 6 listes $(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)})_{\sigma \in \mathfrak{S}_3}$. Le discriminant de $P = X^3 + pX + q$ est $D = -(4p^3 + 27q^2) = -27\Delta$ (il a été calculé dans l'exemple 2 et on peut le vérifier à partir des formules (5)). Ainsi, *pour que (1) admette 3 racines simples, il faut et il suffit que : $\Delta \neq 0$.*

La résolution par radicaux de l'équation (1) a réussi, mais on aimerait comprendre pourquoi. Pour cela introduisons dans $\mathbb{C}[X_1, X_2, X_3]$ les polynômes

$$A = X_1 + jX_2 + j^2 X_3, \quad B = X_1 + j^2 X_2 + jX_3, \\ V = (X_1 - X_2)(X_1 - X_3)(X_2 - X_3).$$

⁽¹⁾ Hieronimo Cardano (en français Jérôme Cardan) fut un célèbre médecin, astrologue, poète et mathématicien italien (1501-1576). Il inventa également la suspension

Avec les notations habituelles, on a : $A^3 = S_3 + 6\sigma_3 + 3jE + 3j^2F$, où $E = X_1^2X_2 + X_2^2X_3 + X_3^2X_1$ et $F = X_1X_2^2 + X_2X_3^2 + X_3X_1^2$, et

$$B^3 = S_3 + 6\sigma_3 + 3j^2E + 3jF$$

(on a remplacé j par j^2 dans A^3), d'où :

$$A^3 + B^3 = 2S_3 + 12\sigma_3 - 3(E + F), \quad A^3 - B^3 = 3j\sqrt{3}(E - F) = 3j\sqrt{3}V.$$

On peut calculer directement $E + F = \sum_{i \neq j} X_i^2 X_j = \sigma_1 \sigma_2 - 3\sigma_3$, ainsi que $S_3 = \sigma_1^3 - 3\sigma_1 \sigma_2 + 3\sigma_3$ (cf. exemple 1 du § X.4), d'où :

$$A^3 + B^3 = 2\sigma_1^3 - 9\sigma_1 \sigma_2 + 27\sigma_3,$$

et par suite :

$$A^3 = \sigma_1^3 - \frac{9}{2}\sigma_1 \sigma_2 + \frac{27}{2}\sigma_3 + 3j\frac{\sqrt{3}}{2}V \quad \text{et} \quad B^3 = \sigma_1^3 - \frac{9}{2}\sigma_1 \sigma_2 + \frac{27}{2}\sigma_3 - 3j\frac{\sqrt{3}}{2}V.$$

Sous l'action naturelle du groupe \mathfrak{S}_3 dans $\mathbb{C}[X_1, X_2, X_3]$, l'orbite de V est de cardinal 2 : c'est $\{V, -V\}$. Donc celle de A^3 est $\{A^3, B^3\}$. Mais il est clair que $X_1 = \frac{1}{3}(\sigma_1 + A + B)$, $X_2 = \frac{1}{3}(\sigma_1 + j^2A + jB)$, $X_3 = \frac{1}{3}(\sigma_1 + jA + j^2B)$.

Enfin, V^2 est le *discriminant général* D (pour $n = 3$), dont la valeur en fonction des σ_k est $V^2 = -(4\sigma_2^3 + 27\sigma_3^2) + \sigma_1 G$, où G est un polynôme en $\sigma_1, \sigma_2, \sigma_3$ qu'il est inutile d'expliciter ici.

Dans tout ce qui précède, prenons les *valeurs* des divers polynômes sur une liste (x_1, x_2, x_3) des racines de (1) : on retrouve exactement les *formules de Cardan* (à cela près que $A^3 = 27\alpha^3$ et $B^3 = 27\beta^3$). La circonstance favorable a été, en dernier ressort, le fait que X_1, X_2, X_3 ont pu être exprimés comme combinaisons \mathbb{C} -linéaires de A, B et σ_1 , que A et B sont exprimables par radicaux en fonction de A^3 et B^3 , eux-mêmes exprimables par radicaux à partir de $\sigma_1, \sigma_2, \sigma_3$ puisque $V^2 \in \mathbb{C}[\sigma_1, \sigma_2, \sigma_3]$.

L'orbite de A sous \mathcal{U}_3 est $\{A, jA, j^2A\}$, et sous \mathfrak{S}_3 , c'est

$$\{A, jA, j^2A, B, jB, j^2B\}.$$

Ainsi A^3 et B^3 ne prennent chacun que les *deux* valeurs A^3 et B^3 sous l'action de \mathfrak{S}_3 .

Equation de degré 3 et radicaux réels

Conservons les notations précédentes et supposons toujours $\Delta \neq 0$ (le cas où l'équation (1) admet une racine double étant immédiat à résoudre). Les formules de Cardan (5) montrent qu'une *condition suffisante* pour pouvoir exprimer les racines de (1) par radicaux réels est de savoir exprimer les racines cubiques des nombres $\frac{-q \pm \delta}{2}$. Si ces nombres sont tous deux réels, et cela arrive ssi $q \in \mathbb{R}, p \in \mathbb{R}$ et $\Delta > 0$, alors dans les formules (5) on peut choisir pour α et β les racines cubiques *réelles* de $\frac{-q \pm \delta}{2}$ et comme

$j = \frac{-1 + i\sqrt{3}}{2}$ et $j^2 = \frac{-1 - i\sqrt{3}}{2}$, il est clair que dans ce cas on a réussi à exprimer les 3 racines par radicaux réels. Remarquons de plus que $\alpha j + \beta j^2 \notin \mathbb{R}$ et $\alpha j^2 + \beta j \notin \mathbb{R}$, car $\alpha \neq \beta$, et l'équation (1) admet donc deux racines complexes conjuguées. En résumé, si $(p, q) \in \mathbb{R}^2$ et $\Delta > 0$, (1) admet une et une seule racine réelle, et toutes les racines sont exprimables par radicaux réels à partir de p et q (et de $\mathbb{Q}[i]$).

Trisection de l'angle

En dehors du cas simple que nous venons de considérer ci-dessus, nous n'étudierons pas la résolution par radicaux réels d'une équation de degré 3, car c'est une question fort délicate. Nous évoquerons cependant en quelques lignes le célèbre problème de la trisection de l'angle qui se ramène à la résolution de l'équation $z^3 = a$ ($a \in \mathbb{C} \setminus \mathbb{R}$). On peut supposer $|a| = 1$ (car $z^3 = a$ ssi $z = \sqrt[3]{|a|} t$ avec $t^3 = \frac{a}{|a|}$), c'est-à-dire $a = e^{i\varphi}$, avec $\varphi \in]0, \pi[$ (le cas où $\varphi \in]-\pi, 0[$ se ramenant au précédent par le changement de z en $-z$). Les racines de l'équation $z^3 = e^{i\varphi}$ sont alors $z_k = e^{i\frac{\varphi}{3} + 2ik\frac{\pi}{3}}$, $0 \leq k \leq 2$. Leurs parties réelles sont les

$$\gamma_k = 2 \cos \left(\frac{\varphi}{3} + \frac{2k\pi}{3} \right),$$

et $(\gamma_0, \gamma_1, \gamma_2)$ est une liste des racines du polynôme $f(T) = T^3 - 3T - 2\cos\varphi$ dont le discriminant est $\Delta = -108\sin^2\varphi < 0$. On peut prouver qu'en général, quand $\Delta < 0$, les racines de (1) ne sont pas exprimables par radicaux réels. C'est le cas ici des γ_k qui ne sont pas exprimables par radicaux réels à partir de $\cos\varphi$ et de $\mathbb{Q}[i]$. (Il y a cependant beaucoup d'exceptions, par exemple $\varphi = \frac{k\pi}{2^n}$, avec $0 < k < 2^n$, pour ne citer que les plus évidentes.)

Résolution trigonométrique de (1)

Les considérations précédentes nous conduisent à réexaminer la résolution de l'équation (1) dans le cas où $(p, q) \in \mathbb{R}^2$ et $\Delta < 0$ (ce qui implique $p < 0$). Les formules de Cardan (5) montrent que dans ce cas l'équation (1) admet 3 racines réelles : en effet on peut choisir dans (5) α et β complexes conjugués, puisque leurs cubes le sont, et après un tel choix, on a $x_1 = \alpha + \beta$, $x_2 = \alpha j + \beta j^2 = \alpha j + \overline{\alpha j}$, $x_3 = \alpha j^2 + \beta j = \alpha j^2 + \overline{\alpha j^2}$ qui sont bien dans \mathbb{R} . Une façon simple d'obtenir ces 3 racines est de transformer le polynôme $P = X^3 + pX + q$ par le changement de variable

$$X = Y \sqrt{\frac{-4p}{3}},$$

ce qui donne $P(X) = Q(Y) = \frac{-p}{3} \sqrt{\frac{-4p}{3}} \left[4Y^3 - 3Y - \frac{3q\sqrt{3}}{2p\sqrt{-p}} \right]$.

Comme l'hypothèse $\Delta < 0$ entraîne $\left| \frac{3q\sqrt{3}}{2p\sqrt{-p}} \right| < 1$ rien n'empêche de

poser $\frac{3q\sqrt{3}}{2p\sqrt{-p}} = \cos \varphi$, avec $\varphi \in]0, \pi[$, et l'équation $P = 0$ équivaut

alors à $\cos 3\theta = \cos \varphi$, en posant $X = \sqrt{\frac{-4p}{3}} \cos \theta$. Les trois racines

réelles sont alors les trois nombres $x_k = \sqrt{\frac{-4p}{3}} \cos \left(\frac{\varphi}{3} + \frac{2k\pi}{3} \right)$, $k \in \llbracket 0, 2 \rrbracket$. En résumé, si $(p, q) \in \mathbb{R}^2$, une condition nécessaire et suffisante pour que P soit dissocié sur \mathbb{R} est : $\Delta \leq 0$, résultat qu'on pourrait bien entendu obtenir par l'Analyse, en étudiant les variations de $x \mapsto P(x)$.

Exercice 1 : Calculer le discriminant des polynômes suivants de $\mathbb{C}[X]$:

a) $(X+a)^n + bX + c$ ($n \geq 3$, a, b, c dans \mathbb{C}). En déduire le discriminant général en degré 3.

b) $(X+a)^n + bX^n + cX^{n-1}$ ($n \geq 3$, a, b, c dans \mathbb{C}).

c) $\frac{X^n}{n!} + \frac{X^{n-1}}{(n-1)!} + \dots + 1$.

Exercice 2 : Calculer le discriminant des polynômes de Tchebychev de 1^{re} et de 2^e espèce d'ordre donné $n \geq 2$: $T_n(X)$ tel que $T_n(\cos \theta) = \cos n\theta$ et $U_n(X)$ tel que

$$U_n(\cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}.$$

Exercice 3 : Déterminer pour quelles valeurs de λ les équations suivantes possèdent au moins une racine double :

a) $x^3 - 3x + \lambda = 0$; b) $x^3 - 8x^2 + (13 - \lambda)x - 6 - 2\lambda = 0$;

c) $x^4 - 4x^3 + (2 - \lambda)x^2 + 2x - 2 = 0$.

Exercice 4 : Soit (α, β, γ) une liste des racines du polynôme $P \in \mathbb{C}[X]$ tel que : $P = X^3 + aX^2 + bX + c$. Calculer les fonctions f suivantes de (α, β, γ) (quand elles sont définies) :

a) $f = \sum (\beta^3 - \gamma^3)^2$; b) $f = \sum \frac{\beta^2 + \gamma^2}{\beta + \gamma} = \frac{\beta^2 + \gamma^2}{\beta + \gamma} + \frac{\gamma^2 + \alpha^2}{\gamma + \alpha} + \frac{\alpha^2 + \beta^2}{\alpha + \beta}$;

c) $f = \sum \frac{2\beta\gamma - \alpha^2}{\beta + \gamma - \alpha}$; d) $f = \sum \left(\frac{\beta - \gamma}{\beta + \gamma} \right)^2$; e) $f = \sum (\beta + \gamma - \alpha)^3$.

Exercice 5 : Calculer $f = \sum_{i < j} \frac{\alpha_i^2 + \alpha_j^2}{\alpha_i \alpha_j}$ et $g = \sum_{i \neq j} \frac{\alpha_i}{\alpha_j^2}$ lorsque $(\alpha_1, \dots, \alpha_n)$ est une liste des racines de $P = X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{C}[X]$ ($a_k \in \mathbb{C}$, $n \geq 2$, $a_n \neq 0$).

Exercice 6 : Résoudre une équation algébrique de degré n sachant que ses racines forment une progression arithmétique.

Exercice 7 : Pour le polynôme $P = X^4 + aX^3 + bX^2 + cX + d$, dont $(x_k)_{1 \leq k \leq 4}$ est une liste des racines, calculer les fonctions f suivantes des (x_k) :

a) $f = \sum_{\{i, j, k, l\} = \llbracket 1, 4 \rrbracket} (x_i - x_j)^2 (x_k - x_l)^2$;

$$b) f = \sum_{\{i, j, k, l\} = \llbracket 1, 4 \rrbracket} (x_j + x_k - x_i - x_l);$$

$$c) f = \sum_{i < j, (i, j, k) \text{ distincts}} \frac{x_i x_j}{x_k^2}.$$

Exercice 8 : Soit $P = X^4 - 3X^2 + 5X - 1 \in \mathbb{C}[X]$, $P = \prod_{k=1}^4 (X - x_k)$.

On pose $f = \sum_{1 \leq i < j \leq 4} \frac{x_i^4 x_j^4}{x_i^4 x_j^4 - x_i^4 - x_j^4 + 1}$. Montrer que $f = \frac{303}{136}$.

Exercice 9 : Résoudre dans \mathbb{C} les équations suivantes, en cherchant une racine rationnelle :

- a) $10x^3 - 17x^2 + x + 28 = 0$; b) $3x^3 - 4x^2 + x + 88 = 0$;
 c) $27x^3 + 42x^2 - 28x - 8 = 0$; d) $x^3 - 5x^2 - 16x + 80 = 0$;
 e) $x^3 - 9x^2 + 14x + 24 = 0$; f) $x^4 + 2x^3 - 21x^2 - 22x + 40 = 0$;
 g) $81x^3 - 18x^2 - 36x + 8 = 0$; h) $3x^4 - 25x^3 + 50x^2 - 50x + 12 = 0$;
 i) $x^4 - 5x^3 + 11x^2 - 13x + 6 = 0$.

Exercice 10 : Soit l'équation du 3^e degré : $x^3 + ax^2 + bx + c = 0$.

a) CNS pour qu'elle admette deux racines α et β telles que $\beta = 2\alpha$.

b) Cette condition étant satisfaite, ramener l'équation à une du second degré.

Exercice 11 : Soit l'équation $x^3 + ax^2 + bx + c = 0$ et (α, β, γ) une liste des racines. Condition nécessaire et suffisante pour que (α, β, γ) forment (après renumérotage si besoin est) :

- a) une progression géométrique ;
 b) une progression harmonique (i.e. $\frac{2}{\beta} = \frac{1}{\alpha} + \frac{1}{\gamma}$) ;
 c) vérifient $\alpha\beta + 1 = 0$. Dans chacun de ces cas, en déduire une résolution de l'équation évitant les formules de Cardan.

Exercice 12 : Former un polynôme de degré 3 à coefficients dans \mathbb{Z} dont deux des racines soient :

- a) 1 et $3 + 2i$; b) 1 et $5 - i$.

Exercice 13 : Former l'équation de degré 4 ayant les 4 racines

$$-9 + \varepsilon_1 \sqrt{137} + 3 \sqrt{34 - 2\varepsilon_2 \sqrt{137}} \quad \text{où} \quad (\varepsilon_1, \varepsilon_2) \in \{-1, +1\}^2.$$

Exercice 14 : Soit $(\alpha, \beta) \in \mathbb{C}^2$. Former l'équation de degré 9 ayant pour racines les nombres $\alpha j^k + \beta j^l$, $(k, l) \in \llbracket 0, 2 \rrbracket^2$.

Exercice 15 : Soit (α, β, γ) une liste des racines de l'équation $P(x) = 0$, où $P = X^3 + aX^2 + bX + c$. Former l'équation du second degré ayant pour racines (u, v) :

$$u = \frac{\beta\gamma + j\gamma\alpha + j^2\alpha\beta}{\alpha + j\beta + j^2\gamma}, \quad v = \frac{\beta\gamma + j^2\gamma\alpha + j\alpha\beta}{\alpha + j^2\beta + j\gamma}.$$

Exercice 16 : On donne deux équations de degré 3 : $P_1(x) = 0$, $P_2(x) = 0$, où : $P_k(x) = X^3 + p_k X + q_k$ ($k \in \{1, 2\}$), et l'on note $(\alpha_k, \beta_k, \gamma_k)$ une liste des racines de P_k . Former l'équation de degré 6 admettant pour racines les 6 valeurs prises par $\varphi = \alpha_1 \alpha_2 + \beta_1 \beta_2 + \gamma_1 \gamma_2$ lorsqu'on change arbitrairement les numérotations.

Exercice 17 : On donne $(\alpha, \beta, \gamma) \in \mathbb{C}^3$. Résoudre le système à l'unique inconnue $x \in \mathbb{C}$: $(\beta - \gamma)^3 (x - \alpha)^3 = (\gamma - \alpha)^3 (x - \beta)^3 = (\alpha - \beta)^3 (x - \gamma)^3$.

Exercice 18 : Soit $(x_k)_{1 \leq k \leq 3}$ une liste des racines de $P = X^3 + aX^2 + bX + c$. Former l'équation de degré 6 ayant pour racines $(x_{\sigma(1)} + jx_{\sigma(2)} + j^2 x_{\sigma(3)})_{\sigma \in \mathfrak{S}_3}$.

Exercice 19 : Soit $\gamma \in \mathbb{Q}$ tel que le nombre $\Delta = \frac{4}{27}(\gamma - 1)^3 + \gamma^2$ (relatif à l'équation $(x - 1)(x^2 + x + \gamma) = 0$) soit > 0 . En utilisant les formules de Cardan, démontrer que $\sqrt[3]{3} = \sqrt[3]{2\sqrt{7} + 3\sqrt{3}} - \sqrt[3]{2\sqrt{7} - 3\sqrt{3}}$ (indication : faire $\gamma = 2$). Vérifier directement cette égalité et en trouver d'autres analogues.

Exercice 20 : Soit $x \in \mathbb{C}^*$ et $n \in \mathbb{Z}$. On pose $y = x + \frac{1}{x}$ et $S_n(x) = x^n + \frac{1}{x^n}$. Sachant que deux valeurs consécutives prises par la suite S_n sont des entiers rationnels, montrer que la suite $(S_n)_{n \in \mathbb{Z}}$ ne prend que des valeurs entières.

§ X.6 ÉQUATIONS DE DEGRÉ 4, ÉQUATIONS PARTICULIÈRES

Equation de degré 4 et radicaux

Nous allons montrer que toute équation algébrique de degré 4 est résoluble par radicaux (mais nous n'aborderons pas la difficile étude des radicaux réels). Pour cela raisonnons dans la \mathbb{C} -algèbre de polynômes $\mathcal{P} = \mathbb{C}[X_1, X_2, X_3, X_4]$ où nous utilisons les notations des §§ antérieurs.

Il existe des polynômes $G \in \mathcal{P}$ possédant les propriétés suivantes :

- (\mathcal{R}_1) la \mathfrak{S}_4 -orbite de G pour l'action naturelle de \mathfrak{S}_4 est de cardinal 3 ;
- (\mathcal{R}_2) si $\{G_1, G_2, G_3\}$ est cette orbite, les X_k sont exprimables par radicaux en fonction des G_i .

Il est facile d'exhiber de tels polynômes, par exemple $X_1 X_2 + X_3 X_4$, $(X_1 + X_2 - X_3 - X_4)^2$, $(X_1 - X_2)^2 + (X_3 - X_4)^2$, ...

Si G est un tel polynôme et $\{G_k\}_{1 \leq k \leq 3}$ sa \mathfrak{S}_4 -orbite, le polynôme en T :

$$R_G(T) = (T - G_1)(T - G_2)(T - G_3)$$

s'écrit

$$R_G(T) = T^3 + AT^2 + BT + C,$$

où A, B, C sont des polynômes symétriques des X_k . Donc, en vertu du théorème X.3.2, A, B et C sont des polynômes en les $(\sigma_k)_{1 \leq k \leq 4}$ à coefficients dans \mathbb{C} . D'autre part, l'étude du § X.5 montre que les G_k sont exprimables par radicaux en fonction des (σ_k) . En revenant aux (X_k) , on en déduit qu'ils sont eux-mêmes exprimables par radicaux en fonction des σ_k , ce qui explique que les racines de l'équation générale du quatrième degré peuvent s'exprimer par radicaux en fonction des coefficients de l'équation.

On peut se demander pourquoi il existe des fonctions G du type ci-dessus pour des polynômes à $n = 4$ lettres, et pas pour $n \geq 5$. Cela résulte en fait de la structure très particulière du groupe \mathfrak{S}_4 (qui a été étudié au § V.7), et notamment (avec les notations du § V.7) de l'existence de la suite de sous-groupes, $\{e\} \subset K \subset \mathcal{U}_4 \subset \mathfrak{S}_4$ où chaque sous-groupe est distingué dans le suivant, les groupes

cycliques, de cardinal 3 pour \mathcal{U}_4/K , de cardinal 2 pour $\mathfrak{S}_4/\mathcal{U}_4$, le groupe de Klein K étant lui-même isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.

Au contraire, dès que $n \geq 5$, on peut démontrer que le groupe alterné \mathcal{U}_n ne possède pas de sous-groupe distingué non trivial (on dit que \mathcal{U}_n est « simple »), et c'est au fond ce qui empêche que l'équation générale de degré ≥ 5 puisse être résolue par radicaux (cf. § V.7, exercice 15).

Méthode d'Euler

Parmi les polynômes G vérifiant (\mathcal{R}_1) et (\mathcal{R}_2) choisissons

$$G = (X_1 + X_2 - X_3 - X_4)^2,$$

d'où $G_1 = G$, $G_2 = (X_1 + X_3 - X_2 - X_4)^2$ et $G_3 = (X_1 + X_4 - X_2 - X_3)^2$.

Posons $U_1 = X_1 X_2 + X_3 X_4$, $U_2 = X_1 X_3 + X_2 X_4$ et $U_3 = X_1 X_4 + X_2 X_3$, ce qui permet d'écrire $G_k = \sigma_1^2 - 4\sigma_2 + 4U_k$ ($k \in \llbracket 1, 3 \rrbracket$). Calculons les fonctions symétriques élémentaires de U_k et des G_k à l'aide des (σ_k) , en commençant par celles des U_k :

$$(1) \quad U_1 + U_2 + U_3 = \sigma_2$$

$$(2) \quad U_1 U_2 + U_2 U_3 + U_3 U_1 = \sum_{\substack{(i,j,k) \text{ distincts} \\ j < k}} X_i^2 X_j X_k = \sigma_1 \sigma_3 - 4\sigma_4$$

et $U_1 U_2 U_3 = \varphi + \sigma_4 S_2$, avec $\varphi = \sum_{i < j < k} X_i^2 X_j^2 X_k^2$ et $S_2 = \sigma_1^2 - 2\sigma_2$. Or $\sigma_3^2 = \varphi + 2\sigma_2 \sigma_4$, d'où

$$(3) \quad U_1 U_2 U_3 = \sigma_1^2 \sigma_4 + \sigma_3^2 - 4\sigma_2 \sigma_4.$$

A l'aide de (1), (2) et (3), un calcul élémentaire donne :

$$(4) \quad G_1 + G_2 + G_3 = 3\sigma_1^2 - 8\sigma_2$$

$$(5) \quad G_1 G_2 + G_2 G_3 + G_3 G_1 = 3\sigma_1^4 - 16\sigma_1^2 \sigma_2 + 16\sigma_2^2 + 16\sigma_1 \sigma_3 - 64\sigma_4$$

$$(6) \quad G_1 G_2 G_3 = \sigma_1^6 - 8\sigma_1^4 \sigma_2 + 16\sigma_1^3 \sigma_3 + 16\sigma_1^2 \sigma_2^2 - 64\sigma_1 \sigma_2 \sigma_3 + 64\sigma_3^2.$$

Remarquons (ce sera utile pour la suite) que :

$$G_2 - G_3 = 4(X_4 - X_3)(X_2 - X_1), \quad \text{etc.,} \quad \text{d'où}$$

$$(7) \quad \prod_{1 \leq i < j \leq 3} (G_j - G_i) = 64 \prod_{1 \leq i < j \leq 4} (X_j - X_i).$$

Remarquons également que, si l'on pose $\Gamma_1 = X_1 + X_2 + X_3 + X_4$

$\Gamma_2 = X_1 + X_3 - X_2 - X_4$, $\Gamma_3 = X_1 + X_4 - X_2 - X_3$, alors

$$(8) \quad \Gamma_1 \Gamma_2 \Gamma_3 = 8 \sigma_3 + \sigma_1 E,$$

où E est un polynôme symétrique en les (X_k) qu'il est inutile de calculer.

Il est temps de considérer maintenant l'équation de degré 4 : $P_1(x) = 0$, où $P_1(X) = X^4 + a_1 X^3 + b_1 X^2 + c_1 X + d_1$, que l'on ramène tout de suite à la forme réduite

$$(9) \quad P(x) = 0, \quad \text{où} \quad P(X) = X^4 + bX^2 + cX + d, \quad (b, c, d) \in \mathbb{C}^3$$

par le changement de variable $X \mapsto X - \frac{a_1}{4}$.

Si $d = 0$, l'équation (9) se ramène au degré 3. Si $c = 0$, l'équation (9) est *bicarrée* et se ramène au degré 2 en posant $Y = X^2$. Nous examinerons la résolution de l'équation (9) dans le cas général où $cd \neq 0$.

Analyse : Soit (x_1, x_2, x_3, x_4) une liste des racines de (9). Notons (u_k) , (g_k) , (γ_k) les valeurs prises par les polynômes (U_k) , (G_k) , (Γ_k) sur cette liste, et notons $(\tilde{\sigma}_k)$ les fonctions symétriques élémentaires des (x_i) .

On a : $\tilde{\sigma}_1 = 0$, $\tilde{\sigma}_2 = b$, $\tilde{\sigma}_3 = -c$, $\tilde{\sigma}_4 = d$, et les relations (1) à (6) fournissent :

$$u_1 + u_2 + u_3 = b, \quad u_1 u_2 + u_2 u_3 + u_3 u_1 = -4d, \quad u_1 u_2 u_3 = c^2 - 4bd,$$

et

$$g_1 + g_2 + g_3 = -8b, \quad g_1 g_2 + g_2 g_3 + g_3 g_1 = 16(b^2 - 4d), \\ g_1 g_2 g_3 = 64c^2.$$

Par suite, on a dans $\mathbb{C}[X]$:

(10)

$$(X - g_1)(X - g_2)(X - g_3) = X^3 + 8bX^2 + 16(b^2 - 4d)X - 64c^2.$$

Il est clair d'autre part, en tenant compte de $\tilde{\sigma}_1 = 0$, que

$$(11) \quad x_1 = \frac{\gamma_1 + \gamma_2 + \gamma_3}{4}, \quad x_2 = \frac{\gamma_1 - \gamma_2 - \gamma_3}{4}, \\ x_3 = \frac{\gamma_2 - \gamma_3 - \gamma_1}{4}, \quad x_4 = \frac{\gamma_3 - \gamma_1 - \gamma_2}{4}.$$

Rappelons aussi que par définition

$$(12) \quad (\gamma_k)^2 = g_k \quad (1 \leq k \leq 3).$$

Enfin, de (8), on déduit :

$$(13) \quad \gamma_1 \gamma_2 \gamma_3 = -8c.$$

En résumé : soit $R(X)$ le polynôme donné par (10) (nous dirons que c'est la **résolvante** de l'équation (9)). Alors (g_1, g_2, g_3) est une liste des racines de $R(X)$; les γ_k sont des racines carrées appropriées, respectivement, des g_k , telles que $\gamma_1 \gamma_2 \gamma_3 = -8c$; les x_k sont donnés par (11) en fonction des γ_k .

Si D_P et D_R sont les discriminants respectifs de P et R , on déduit de (7) que $D_R = 2^{12} D_P$ (ce qui fournit une méthode de calcul de D_P) et par suite : pour que (9) ait des racines multiples, il faut et il suffit que sa résolvante en ait.

Synthèse : La résolvante $R(X)$ donnée par (10) est un polynôme du troisième degré dont nous avons appris au § X.5 à chercher les racines. Soit (g_1, g_2, g_3) une liste de ces racines. On a $g_1 g_2 g_3 = 64c^2$. On peut donc choisir, et c'est ce que nous ferons, des racines carrées (ρ_k) des (g_k) telles que $\rho_1 \rho_2 \rho_3 = -8c$. Posons :

$$y_1 = \frac{\rho_1 + \rho_2 + \rho_3}{4}, \quad y_2 = \frac{\rho_1 - \rho_2 - \rho_3}{4},$$

$$y_3 = \frac{\rho_2 - \rho_3 - \rho_1}{4}, \quad y_4 = \frac{\rho_3 - \rho_1 - \rho_2}{4}.$$

Pour chaque $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \{-1, +1\}^3$, introduisons $\varepsilon * y_i =$ la valeur obtenue en remplaçant (ρ_1, ρ_2, ρ_3) par $(\varepsilon_1 \rho_1, \varepsilon_2 \rho_2, \varepsilon_3 \rho_3)$ dans y_i .

De même, pour chaque $\sigma \in \mathfrak{S}_3$, introduisons : $y_i^{\langle \sigma \rangle} =$ la valeur obtenue en remplaçant dans y_i la suite (ρ_1, ρ_2, ρ_3) par $(\rho_{\sigma(1)}, \rho_{\sigma(2)}, \rho_{\sigma(3)})$ (ce qui revient à changer la numérotation des (g_k) à l'aide de σ).

On constate d'abord que : si $\varepsilon_1 \varepsilon_2 \varepsilon_3 = +1$, la suite $(\varepsilon * y_k)_{1 \leq k \leq 4}$ est égale à $(y_{\tau(1)}, y_{\tau(2)}, y_{\tau(3)}, y_{\tau(4)})$, où $\tau \in K$ (i.e. τ est une « double transposition ») ; et si $\varepsilon_1 \varepsilon_2 \varepsilon_3 = -1$, la suite $(\varepsilon * y_k)_{1 \leq k \leq 4}$ est $(-y_{\tau(k)})_{1 \leq k \leq 4}$ avec $\tau \in K$.

Ensuite, on voit que, pour chaque $\sigma \in \mathfrak{S}_3$, on a :

$$(y_k^{\langle \sigma \rangle})_{1 \leq k \leq 4} = (y_{\tau(k)})_{1 \leq k \leq 4}, \quad \text{où } \tau \in \mathfrak{S}_4.$$

Finalement, soit $(z_k)_{1 \leq k \leq 4}$ une liste obtenue à partir de (y_k) en changeant les ρ_k en $(\varepsilon_k \rho_k)$ ($\varepsilon \in \{-1, +1\}^3$), puis en modifiant arbitrairement l'ordre des ρ_k . Alors le polynôme $Q = \prod_{k=1}^4 (X - z_k)$ est l'un ou l'autre des

deux polynômes : $L = \prod_{k=1}^4 (X - y_k)$ ou bien $M = \prod_{k=1}^4 (X + y_k)$. De manière précise, c'est L (resp. M) ssi $\varepsilon_1 \varepsilon_2 \varepsilon_3 = +1$ (resp. $\varepsilon_1 \varepsilon_2 \varepsilon_3 = -1$). Or l'analyse précédente nous a prouvé que si $(x_k)_{1 \leq k \leq 4}$ est une lis

de (9), $P = \prod_{k=1}^4 (X - x_k)$ coïncide nécessairement avec l'un des polynômes Q . En tenant compte de (13), on en déduit que : $P = L$, c'est-à-dire $(y_k)_{1 \leq k \leq 4}$ est une liste des racines de P .

Autres méthodes pour l'équation du 4^e degré

L'équation de degré 4 a donné lieu à une abondante littérature, et nombreuses sont les méthodes de résolution. Bornons-nous à en signaler deux, l'une contemporaine d'Euler, la seconde bien antérieure, toutes deux remarquables.

1° On peut utiliser le polynôme $G = X_1 X_2 + X_3 X_4$ qui, lui aussi, ne prend que 3 valeurs quand on permute les racines de (9). Les relations (1), (2) et (3) fournissent alors la résolvante

$$(14) \quad R_1 = X^3 - bX^2 - 4dX - (c^2 - 4bd)$$

Si (g_1, g_2, g_3) est une liste des racines de R_1 , on utilise pour obtenir les racines (x_k) de P le système $x_1 x_2 + x_3 x_4 = g_1$, $x_1 x_3 + x_2 x_4 = g_2$, $x_1 x_4 + x_2 x_3 = g_3$, $x_1 + x_2 + x_3 + x_4 = 0$, $\sum_{i < j} x_i x_j = b$, mais la remontée de (g_1, g_2, g_3) aux (x_k) est sensiblement plus délicate que dans la méthode d'Euler.

2° *Méthode de Ferrari* ⁽¹⁾. On introduit un paramètre $y \in \mathbb{C}$ pour écrire l'équation (9) sous la forme $x^4 + x^2 y + \frac{1}{4} y^2 = (y - b) x^2 - cx + \frac{1}{4} y^2 - d$, soit

$$(15) \quad \left(x^2 + \frac{y}{2}\right)^2 = (y - b) x^2 - cx + \frac{1}{4} y^2 - d$$

et on cherche à déterminer $y \in \mathbb{C}$ pour que le second membre de (15) soit un carré parfait $(mx + n)^2$. Pour cela, il faut et il suffit que le discriminant du trinôme en x soit nul, c'est-à-dire : $c^2 - (y^2 - 4d)(y - b) = 0$, ou encore

$$y^3 - by^2 - 4dy + 4bd - c^2 = 0$$

c'est-à-dire que y soit l'une des racines de la même résolvante (14) que dans la méthode précédente. Pour une telle racine y , on calcule m et n tels que $(y - b) x^2 - cx + \frac{1}{4} y^2 - d = (mx + n)^2$, et (15) se décompose alors en les

⁽¹⁾ Ludovico Ferrari (1522-1565), mathématicien italien, élève et secrétaire

deux équations de degré 2 :

$$(16) \quad x^2 - \varepsilon mx + \frac{1}{2}y - \varepsilon n = 0 \quad (\varepsilon \in \{-1, +1\}).$$

Si (x_1, x_2) et (x_3, x_4) sont les couples respectifs des racines des équations (16), on a : $x_1 x_2 = \frac{y}{2} + n$, $x_3 x_4 = \frac{y}{2} - n$, d'où $y = x_1 x_2 + x_3 x_4$. Il n'est donc pas étonnant que les racines de (14) soient exactement les nombres $y \in \mathbb{C}$ qui rendent le second membre de (15) carré parfait : autrement dit, la méthode de Ferrari est une variante élégante du 1°.

Si nous notons $y_1 = x_1 x_2 + x_3 x_4$, $y_2 = x_1 x_3 + x_2 x_4$, $y_3 = x_1 x_4 + x_2 x_3$, on constate que $y_1 - y_2 = (x_4 - x_1)(x_3 - x_2)$, etc., d'où $\prod_{1 \leq i < j \leq 3} (y_j - y_i)^2 =$

$\prod_{1 \leq i < j \leq 4} (x_j - x_i)^2$: les discriminants de (9) et de sa résolvante (14) sont donc égaux.

Equations réciproques

Nous avons déjà signalé qu'une équation algébrique de degré $n \geq 5$ ne peut, *en général*, pas être résolue par radicaux. Cela résulte des travaux d'Abel et de Galois et on pourra trouver des explications détaillées dans tout traité, élémentaire ou non, sur la Théorie de Galois (cf. [6], [13] ou [28]).

Il existe cependant des cas où l'on peut, par transformations simples, *abaisser le degré* d'une équation, c'est-à-dire la ramener à une équation algébrique de degré moindre. Nous avons déjà signalé au passage le cas de l'équation bicarrée $x^4 + bx^2 + d = 0$ qui est du second degré en $y = x^2$. Voici un autre exemple simple :

DÉFINITION X.6.1

$\left\{ \begin{array}{l} \text{Un polynôme } P = a_0 X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{C}[X] \text{ (} n \geq 1 \text{) est} \\ \text{dit } \textbf{réciproque de première espèce} \text{ ssi } \left(\forall k \leq \frac{n}{2} \right) a_k = a_{n-k} \\ \text{et } \textbf{réciproque de deuxième espèce} \text{ ssi } \left(\forall k \leq \frac{n}{2} \right) a_k + a_{n-k} = 0. \end{array} \right.$

Si P est réciproque de deuxième espèce, le nombre 1 est racine de P et après division de $P(X)$ par $(X - 1)$ on obtient un polynôme $Q(X)$ qui est réciproque de première espèce.

Si P est réciproque de première espèce et de degré n impair, alors -1 est racine de P , et après division par $X + 1$ on obtient un polynôme réciproque de première espèce et de degré pair, d'où :

PROPOSITION X.6.1

$\left\| \begin{array}{l} \text{Si un polynôme réciproque n'admet ni } 1 \text{ ni } -1 \text{ pour racines, il est de} \\ \text{première espèce et de degré pair.} \end{array} \right.$

Après avoir débarrassé un polynôme réciproque de ses racines éventuelles 1 et -1 , on est donc ramené à étudier, pour $n \geq 1$, un polynôme du type

$$P(X) = a_0 X^{2n} + a_1 X^{2n-1} + \dots + a_{n-1} X^{n+1} + a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

n'admettant ni 1 ni -1 pour racine, avec $a_0 \neq 0$.

On constate que $P(X) = X^{2n} P\left(\frac{1}{X}\right)$, d'où l'on déduit que si x est une racine de P avec l'ordre de multiplicité $\alpha \geq 1$, alors $\frac{1}{x}$ est aussi racine de P , de même multiplicité α . De plus 0 n'est pas racine puisque $a_0 \neq 0$. Posons alors : $S_n = x^n + \frac{1}{x^n}$ pour $n \geq 0$ ($S_0 = 2$, $S_1 = U$). L'équation $P(x) = 0$ équivaut à

$$(17) \quad a_0 S_n + a_1 S_{n-1} + \dots + a_{n-1} U + a_n = 0$$

à la nouvelle inconnue U (il suffit de diviser P par X^n). Or les S_n sont fournis par la relation de récurrence

$$(18) \quad S_0 = 2, S_1 = U, \text{ et } (\forall n \geq 2) \quad S_n = US_{n-1} - S_{n-2}$$

qui donne l'expression de S_n comme polynôme en U .

Si l'on désire l'expression générale de S_n , il suffit d'introduire la série formelle $S = \sum_{n \geq 0} S_n X^n \in \mathbb{C}[[X]]$ qui vérifie $S(1 - XU + X^2) = 2 - UX$,

d'où $S = \frac{2 - UX}{1 - (UX - X^2)} = \left(\sum_{k \geq 0} (UX - X^2)^k \right) (2 - UX)$, d'où par identification :

$$(19) \quad S_n = 2 A_n(U) - U A_{n-1}(U)$$

avec, pour tout p , $A_p(U) = \sum_{0 \leq 2k \leq p} (-1)^k \binom{p-k}{k} U^{p-2k}$.

Comme on pouvait s'y attendre, on voit que S_p est, pour tout $p \in \mathbb{N}$, un polynôme de degré p en U . En reportant dans (17) on obtient donc une équation de degré n en U appelée *résolvante* de P . A chaque racine u de cette résolvante correspondent deux racines inverses l'une de l'autre x et $\frac{1}{x}$ de P , données par l'équation $x + \frac{1}{x} = u$, soit $x^2 - ux + 1 = 0$. D'ailleurs si

le premier membre de (17) est $\Phi(U) = a_0 \prod_{k=1}^r (U - u_k)^{\alpha_k}$, on a :

$$\begin{aligned} P(X) &= X^n \Phi(U) = a_0 X^n \prod_{k=1}^r \left(X + \frac{1}{X} - u_k \right)^{\alpha_k} = \\ &= a_0 \prod_{k=1}^r (X^2 - u_k X + 1)^{\alpha_k} . \end{aligned}$$

ce qui prouve qu'une racine u de multiplicité α dans Φ conduit à deux racines x et $\frac{1}{x}$ de même multiplicité α dans P .

Exemple 1 : (racines 5-ièmes de l'unité)

Partons de $P(X) = X^5 - 1 = (X - 1)(X^4 + X^3 + X^2 + X + 1)$. Le polynôme $Q(X) = X^4 + X^3 + X^2 + X + 1$ est réciproque de première espèce et n'admet ni 1 ni -1 pour racine. L'équation (17) s'écrit ici : $X^2 + \frac{1}{X^2} + X + \frac{1}{X} + 1 = 0$, c'est-à-dire $S_2 + U + 1 = 0$. Or $S_2 = U^2 - 2$ et la résolvante est donc

$$(20) \quad U^2 + U - 1 = 0$$

dont les racines sont : $u_1 = \frac{-1 + \sqrt{5}}{2}$ et $u_2 = \frac{-1 - \sqrt{5}}{2}$.

Les racines de Q (qui ne sont autres que $\exp\left(\pm 2ik \frac{\pi}{5}\right)$, $k \in \llbracket 1, 2 \rrbracket$) sont fournies par les équations $x_1 + \frac{1}{x_1} = u_1$ et $x_2 + \frac{1}{x_2} = u_2$, d'où l'on déduit facilement :

$$x_1 = \exp\left(\frac{2i\pi}{5}\right) = \frac{1}{4}[(\sqrt{5} - 1) + i\sqrt{10 + 2\sqrt{5}}], \quad x'_1 = \bar{x}_1$$

$$x_2 = \exp\left(\frac{4i\pi}{5}\right) = \frac{-1}{4}[\sqrt{5} + 1 - i\sqrt{10 - 2\sqrt{5}}], \quad x'_2 = \bar{x}_2,$$

ce qui donne les 5 racines de $X^5 - 1$ par radicaux réels.

Remarquons au passage qu'en s'arrêtant à la résolvante (20), on obtient déjà par radicaux réels,

$$2 \cos \frac{2\pi}{5} = u_1 = \frac{-1 + \sqrt{5}}{2} > 0 \quad \text{et} \quad 2 \cos \frac{4\pi}{5} = \frac{-1 - \sqrt{5}}{2} < 0.$$

Transformation des équations

Nous avons déjà vu comment des transformations simples ($Y = X^2$ pour une équation bicarrée, $U = X + \frac{1}{X}$ pour une équation réciproque) permettaient parfois d'abaisser le degré d'une équation, ou aussi de la ramener à une forme plus intéressante (par exemple le changement $Z = X + \frac{a_1}{4}$ suffit pour annuler le coefficient de X^3 dans l'équation générale

degré, ce qui nous a été ensuite utile dans sa résolution car alors $\tilde{\sigma}_1 = 0$).

Cela nous conduit à définir, de manière générale, pour un polynôme $P(X) = X^n + a_1 X^{n-1} + \dots + a_n$ dont une liste des racines est (x_1, x_2, \dots, x_n) la transformée de l'équation $P(x) = 0$ ($P \in \mathbb{C}[X]$) par la fonction rationnelle $y = F(x)$: c'est l'équation

$$(Y - F(x_1))(Y - F(x_2)) \dots (Y - F(x_n)) = 0$$

qui admet le même nombre de racines que l'équation de départ, à chaque racine x_i correspondant la racine $F(x_i)$ avec le même ordre de multiplicité. Ce qui est remarquable, c'est que les coefficients des puissances de Y sont des fonctions *symétriques* des x_i et peuvent donc se calculer à partir des coefficients (a_k) de l'équation $P(x) = 0$.

Théoriquement on peut toujours se ramener au cas où F est une fonction polynomiale. En effet si $F(x) = \frac{G(x)}{H(x)}$, avec $H(x_i) \neq 0$ (sinon la transformation ne serait pas définie), alors $\text{pgcd}(P, H) = 1$. D'après le théorème de Bezout, il existe U et V dans $\mathbb{C}[X]$ tels que $UP + VH = 1$, donc $F = \frac{G(UP + VH)}{H} = \frac{G}{H} UP + GV = FUP + Q$, avec $Q \in \mathbb{C}[X]$. Mais pour tout i , $F(x_i) = 0 + Q(x_i)$, $1 \leq i \leq n$, ce qui prouve l'assertion.

Exemple 2 : Transformons l'équation $X^3 + pX + q = 0$ (où $p \neq 0$) par

$$Y = 3X^2 + aX + b.$$

L'équation transformée est

$$(Y - 3x_1^2 - ax_1 - b)(Y - 3x_2^2 - ax_2 - b)(Y - 3x_3^2 - ax_3 - b) = 0.$$

Le coefficient de Y^2 est

$$-3S_2 - 3a\sigma_1 - 3b = -3\sigma_1^2 + 6\sigma_2 - 3a\sigma_1 - 3b = 6p - 3b.$$

Il est facile de l'annuler en posant : $b = 2p$.

Essayons maintenant d'annuler le coefficient de Y :

$$9 \sum_{i < j} x_i^2 x_j^2 + a^2 \sigma_2 + 3b^2 + 3a \sum_{(i,j) \text{ distincts}} x_i^2 x_j + 6bS_2 + 2ab\sigma_1 = 0,$$

ce qui donne la condition $9p^2 + pa^2 + 12p^2 + 9aq - 24p^2 = 0$, soit $pa^2 + 9aq - 3p^2 = 0$. Il est donc possible, par une transformation $Y = 3X^2 + aX + 2p$, où a est bien choisi, de transformer l'équation $X^3 + pX + q = 0$ en une équation de la forme $Y^3 - B = 0$ dont la résolution par radicaux est immédiate. On a ainsi obtenu une nouvelle méthode de résolution de l'équation générale du 3^e degré. Le lecteur pourra par une méthode analogue transformer l'équation générale du 4^e degré en équation bicarrée.

Exercice 1 : Résoudre, par la méthode de Ferrari, l'équation $x^4 + \frac{2}{27}x - \frac{1}{108} = 0$.

Exercice 2 : Soit $P(X) = X^4 + aX^3 + bX^2 + cX + d \in \mathbb{C}[X]$. Former les polynômes

$$\prod_{i < j} (X - x_i x_j) \quad \text{et} \quad \prod_{i < j} \left(X - \frac{x_i + x_j}{2} \right), \quad \text{où} \quad P(X) = \prod_{i=1}^4 (X - x_i).$$

Exercice 3 : Résoudre l'équation $X^4 - 2X^3 + 5X^2 - 6X + 3 = 0$, en cherchant un changement de variable $X = aY + b$ ($a \in \mathbb{C}^*$, $b \in \mathbb{C}$) qui la transforme en une équation réciproque. Préciser tous les couples (a, b) qui conviennent.

Exercice 4 : Soit $(\lambda, \mu, \nu) \in \mathbb{C}^3$. Résoudre avec soin l'équation

$$[x^4 - 6\lambda x^2 + 3(4\mu\nu - \lambda^2)]^2 - 64(\lambda^3 + \mu^3 + \nu^3 - 3\lambda\mu\nu)x^2 = 0.$$

Exercice 5 : Soit $R(X)$ la résolvante, obtenue par la méthode d'Euler, de l'équation $X^4 + bX^2 + cX + d = 0$ ($b, c, d \in \mathbb{C}^3$). Montrer, P désignant le premier membre de l'équation :

- si P admet exactement une racine double, il en est de même de R , et réciproquement.
- Si P admet une racine triple et une simple, alors $R(X) = X^3$.
- Si P admet deux racines doubles, alors $R(X)$ admet 0 pour racine double.

Exercice 6 : On reprend l'exercice 5, mais en supposant que $(b, c, d) \in \mathbb{R}^3$. Prouver :

- Si toutes les racines de P sont réelles, celles de R sont réelles positives.
- Si P n'a pas de racine réelle, R admet deux racines réelles négatives et une positive.
- Si P admet deux racines réelles et deux complexes conjuguées, R admet une racine réelle positive et deux racines complexes conjuguées.

Exercice 7 : Résoudre dans \mathbb{C} l'équation $x^4 + \frac{7}{3}x^2 + 30x - \frac{100}{3} = 0$ en commençant par chercher une racine rationnelle (terminer avec les formules de Cardan).

Exercice 8 : Résoudre dans \mathbb{C} l'équation $4(x^2 - x + 1)^3 - 27x^2(x - 1)^2 = 0$.

Exercice 9 : Résoudre dans \mathbb{C} : $4x^4 - 85x^3 + 357x^2 - 340x + 64 = 0$ en la ramenant à une équation réciproque.

Exercice 10 : Soit $(a_1, a_2, \dots, a_n) \in \mathbb{C}^n$ ($n \geq 1$), avec $a_n \neq 0$. On écrit

$$X^n + a_1 X^{n-1} + \dots + a_n = \prod_{k=1}^n (X - x_k).$$

Montrer comment on peut calculer en fonction des a_k les coefficients du polynôme

$$\prod_{k=1}^n \left(Z - \left(x_k + \frac{1}{x_k} \right) \right). \quad \text{Achever les calculs pour } n \leq 5.$$

Exercice 11 : On considère dans \mathbb{C} l'équation $f(x) = x^4 + 3x^3 - x^2 - 3x + 11 = 0$ (\mathcal{E}).

a) soit $g(y) = y^4 f\left(y + \frac{2}{y}\right)$. Former $g(y)$ et écrire la division euclidienne de g par $y^5 - 1$.

b) En déduire les racines de (\mathcal{E}).

Exercice 12 : a) Soit $(\alpha, \beta) \in \mathbb{C}^2$ et $(p, q) \in \mathbb{C}^2$ tels que $\alpha^5 + \beta^5 = -2q$ et $\alpha\beta = p$, et $f(X) = X^5 - 5pX^3 + 5p^2X + 2q$. Montrer que les racines de l'équation $f(x) = 0$ sont : $(\zeta\alpha + \bar{\zeta}\beta)_{\zeta \in \mu_5}$.

b) En supposant $(p, q) \in \mathbb{R}_+ \times \mathbb{R}$, montrer :

- si $p^5 < q^2$, 4 des racines de f sont non réelles, et une est réelle,
- si $p^5 > q^2$, toutes les racines de f sont réelles,
- si $p^5 = q^2$, f est divisible par le carré d'un polynôme.

Exercice 13 : Déterminer $(a, b, c) \in \mathbb{C}^3$ de manière que chacun des nombres a, b, c soit solution de l'équation $x^3 + ax^2 + bx + c = 0$.

Exercice 14 : Calculer $\lambda \in \mathbb{C}$ pour que deux des racines de $x^4 - 2x^2 + \lambda x + 3 = 0$ aient pour produit 1. Résoudre alors l'équation.

Exercice 15 : Trouver toutes les équations de degré 3 invariantes par la transformation $y = x - \frac{1}{x}$.

Exercice 16 : Résoudre l'équation $X^8 + 1 = 0$ de deux façons et en déduire les fonctions circulaires de $\frac{k\pi}{8}$. Généraliser pour exprimer par radicaux $\cos \frac{\pi}{2^n}$.

Exercice 17 : Soit (x_1, x_2, x_3) une liste des racines de $x^3 + px^2 + qx + r = 0$. Former l'équation dont les racines sont $x_1^2 + x_2^2$, $x_2^2 + x_3^2$ et $x_3^2 + x_1^2$. *Indication :* il suffit de transformer l'équation donnée par $y = p^2 - 2q - x^2$.

Exercice 18 : Pour quelles valeurs de $\lambda \in \mathbb{C}$ l'équation $x^4 - 4\lambda x^3 + 3 = 0$ a-t-elle une racine multiple ? Pour chacune des valeurs trouvées, résoudre l'équation.

Exercice 19 : (une démonstration du théorème de d'Alembert).

a) Soit K un corps commutatif et $P \in K[X]$ un polynôme non constant. Montrer qu'il existe une extension L du corps K , algébrique et de degré fini (cf. § IX.7) telle que P se décompose dans $L[X]$ en facteurs du premier degré. *Indication :* Si F est un facteur irréductible de P , considérer $K[X]/FK[X]$ et désigner par x la classe de X . Ensuite raisonner par récurrence.

b) Soit $P \in \mathbb{C}[X]$. En considérant $F = P\bar{P}$ (\bar{P} est le polynôme dont les coefficients sont conjugués de ceux de P), montrer qu'il suffit de démontrer le théorème de d'Alembert sous la forme : tout polynôme à coefficients réels de degré $d \geq 1$ admet au moins un zéro dans \mathbb{C} .

c) Démontrer par l'Analyse que tout polynôme de $\mathbb{R}[X]$ de degré $d \geq 1$, d étant impair, admet au moins une racine dans \mathbb{R} (théorème des valeurs intermédiaires).

d) Soit alors $d = 2^n q$ où q est impair le degré de $F \in \mathbb{R}[X]$. Supposons $n \geq 1$. D'après a) il existe une extension L de \mathbb{C} telle que $F = \prod_{i=1}^d (X - x_i)$, avec les x_i dans L . Considérer

alors les éléments $y_{ij} = x_i + x_j - cx_i x_j$ de L ($i \leq j$), où $c \in \mathbb{R}$. Montrer que le polynôme $G(X) = \prod_{i \leq j} (X - y_{ij})$ est à coefficients dans \mathbb{R} . Quel est son degré ? Par hypothèse de

récurrence G admet donc une racine $z_c \in \mathbb{C}$.

e) On a donc $y_{i(c), j(c)} = x_{i(c)} + x_{j(c)} - cx_{i(c)} x_{j(c)} = z_c$. Mais \mathbb{R} étant infini, montrer qu'il existe 2 réels distincts c et c' tels que $i(c) = i(c')$ et $j(c) = j(c')$. En déduire l'existence de x_r et de x_s dans \mathbb{C} tels que $x_r + x_s \in \mathbb{C}$ et $x_r x_s \in \mathbb{C}$. Conclure enfin que F a une racine dans \mathbb{C} .

Chapitre XI

MATRICES

§ XI.1 MATRICES DE TYPE (m, n)

Dans ce §, nous désignerons par A un *anneau* non nul.

DÉFINITION XI.1.1

Soit m et n deux entiers ≥ 1 . On appelle **(m, n) -matrice à coefficients dans A** toute application $\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket \rightarrow A$, $(i, j) \mapsto a_{ij}$. Une telle matrice se note $[a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$, ou

$$[a_{ij}]_{(i,j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket},$$

ou $[a_{ij}]$ ou encore (a_{ij}) .
L'ensemble de ces matrices pour m et n fixés sera noté $\mathfrak{M}_{m,n}(A)$.

Soit $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$. Pour i fixé, $i \in \llbracket 1, m \rrbracket$, la suite $j \mapsto a_{ij}$, $\llbracket 1, n \rrbracket \rightarrow A$, s'appelle la **ligne d'indice i** de M : c'est un élément de A^n , que nous noterons $\mathcal{L}_i(M)$. De même, pour j fixé, $j \in \llbracket 1, n \rrbracket$, la suite $i \mapsto a_{ij}$, $\llbracket 1, m \rrbracket \rightarrow A$ s'appelle la **colonne d'indice j** de M : c'est un élément de A^m que nous noterons $\mathcal{C}_j(M)$. Dans le cas fréquent où A est un *corps commutatif* K , on parle de **vecteurs lignes** et de **vecteurs colonnes**.

La matrice M est dite **carrée** ssi $m = n$. Pour m fixé, l'ensemble $\mathfrak{M}_{m,m}(A)$ sera plus simplement noté $\mathfrak{M}_m(A)$.

Il est commode de convenir, si $m = 0$, ou $n = 0$, qu'une (m, n) -matrice est l'application vide $\emptyset \rightarrow A$. Ainsi, lorsque $m = 0$ ou $n = 0$, $\mathfrak{M}_{m,n}(A)$ est constitué d'un seul élément, la *matrice vide*.

Lorsque $m = 1$, M est appelée une **matrice ligne**. Il y a une bijection naturelle évidente entre $\mathfrak{M}_{1,n}(A)$ et A^n , qui associe à $M \in \mathfrak{M}_{1,n}(A)$ son unique ligne.

De même, si $n = 1$, M est appelée une **matrice colonne**. En associant, à $M \in \mathfrak{M}_{m,1}(A)$ son unique colonne, on a une bijection naturelle entre $\mathfrak{M}_{m,1}(A)$ et A^m .

Toutes ces définitions proviennent du fait qu'on représente habituellement une matrice $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$ sous forme d'un *tabl*

laire où les indices des lignes (resp. des colonnes) croissent du haut vers le bas (resp. de gauche à droite).

$$M = \begin{array}{c} \text{col. } j \\ \begin{array}{c} \left[\begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ \hline a_{21} & & & & & & \\ \hline \vdots & & & & & & \\ \hline a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\ \hline \vdots & & & & & & \\ \hline a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn} \end{array} \right] \end{array} \\ \text{ligne } i \end{array} .$$

Les lignes (resp. colonnes) de ce tableau sont précisément celles de M . En traçant des lignes de séparation dans ce tableau, on peut aussi le concevoir comme un *tableau rectangulaire de boîtes carrées*, où chacune des mn boîtes contient un coefficient a_{ij} .

Les $(1, 1)$ -matrices à coefficients dans A sont identifiées aux éléments de A lui-même.

Groupe additif $\mathfrak{M}_{m,n}(A)$

Fixons les entiers m et n ($m \geq 1, n \geq 1$). On définit de manière évidente l'addition dans $\mathfrak{M}_{m,n}(A)$:

$$\text{si } M = [a_{ij}] \in \mathfrak{M}_{m,n}(A), \text{ et si } N = [b_{ij}] \in \mathfrak{M}_{m,n}(A),$$

$M + N$ est, par définition, la matrice $[c_{ij}]$, où $c_{ij} = a_{ij} + b_{ij}$ pour tous $(i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$. Cette addition s'identifie à celle du groupe abélien $A^{\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket}$, et par suite : **$\mathfrak{M}_{m,n}(A)$ muni de l'addition des matrices est un groupe abélien.** L'élément neutre $0_{m,n}$ est la matrice dont tous les coefficients sont nuls (on l'appelle la **matrice nulle** de $\mathfrak{M}_{m,n}(A)$). L'opposé de $M = [a_{ij}]$ est $-M = [-a_{ij}]$.

Cas où A est une K -algèbre

Soit K un corps commutatif, et supposons que A soit une K -algèbre. On peut alors définir le *produit par un scalaire* $\lambda \in K$ d'une matrice $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$: c'est la matrice, notée λM , égale à $[\lambda a_{ij}]$.

On vérifie immédiatement que, muni de ce produit par les scalaires, et de son addition, $\mathfrak{M}_{m,n}(A)$ est un **K -ev.** Cette structure de K -ev est dite **naturelle** sur $\mathfrak{M}_{m,n}(A)$. Ce n'est autre que la structure de K -ev produit sur $A^{\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket}$ que l'on identifie à $\mathfrak{M}_{m,n}(A)$.

Base canonique du K -ev $\mathfrak{M}_{m,n}(K)$

Le cas le plus simple est celui où $A = K$, K corps commutatif. Considérons alors, pour $(i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$, la matrice E_{ij} , élément de $\mathfrak{M}_{m,n}(K)$, définie par $E_{ij} = [a_{\alpha\beta}]_{(\alpha,\beta) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket}$, avec $a_{ij} = 1_K$ et $a_{\alpha\beta} = 0$ si $(\alpha, \beta) \neq (i, j)$. Le K -ev $\mathfrak{M}_{m,n}(K)$ s'identifie au K -ev produit $K \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$ et (E_{ij}) n'est autre que la *base canonique* de ce K -ev (cf. § VI.3). Par suite, (E_{ij}) est une base (dite canonique) du K -ev $\mathfrak{M}_{m,n}(K)$. Il en résulte en particulier :

PROPOSITION XI.1.1

|| Si K est un corps commutatif, le K -ev $\mathfrak{M}_{m,n}(K)$ est de dimension finie égale à mn .

Une matrice quelconque $M = [a_{ij}] \in \mathfrak{M}_{m,n}(K)$ n'est autre que la combinaison linéaire $\sum_{(i,j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket} a_{ij} E_{ij}$.

En utilisant les symboles de Kronecker $\delta_{k,l}$ ($((k, l) \in \mathbb{N}^2, \delta_{l,l} = 1_K$ si $l \in \mathbb{N}, \delta_{kl} = 0$ si $k \neq l$), on a pour tous i et j ($(i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$) :

$$E_{ij} = [\delta_{\alpha i} \delta_{\beta j}]_{(\alpha, \beta) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket}.$$

Remarque 1 : Il est parfaitement évident que si A est un anneau quelconque, et pas nécessairement un corps, on peut encore définir les matrices E_{ij} de manière analogue ($a_{ij} = 1_A$ et tous les autres coefficients nuls). On les appelle alors des matrices *élémentaires*, et il est encore vrai que toute matrice $M \in \mathfrak{M}_{m,n}(A)$ s'écrit de façon unique sous la forme $\sum a_{ij} E_{ij}$ si $M = [a_{ij}]$, les a_{ij} étant dans A .

Sous-matrices

Considérons deux parties non vides de $\llbracket 1, m \rrbracket$ et $\llbracket 1, n \rrbracket$ respectivement : $I \subset \llbracket 1, m \rrbracket$ et $J \subset \llbracket 1, n \rrbracket$. Si p et q sont leurs cardinaux, I et J s'écrivent de manière unique $I = \{\alpha_1, \dots, \alpha_p\}$, $J = \{\beta_1, \dots, \beta_q\}$, avec $1 \leq \alpha_1 < \alpha_2 < \dots < \alpha_p \leq m$ et $1 \leq \beta_1 < \beta_2 < \dots < \beta_q \leq n$. Soit alors $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$ (où A est l'anneau de base).

DÉFINITION XI.1.2

|| Avec ces notations, on appelle *sous-matrice de M à lignes indexées dans I et colonnes indexées dans J* , la matrice $S \in \mathfrak{M}_{p,q}(A)$ égale à $[a_{\alpha_i, \beta_j}]_{(i,j) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket}$. Nous noterons $\mathcal{M}_{I,J}(M)$ cette matrice.

Soit $\tilde{I} = \llbracket 1, m \rrbracket \setminus I$ et $\tilde{J} = \llbracket 1, n \rrbracket \setminus J$. Dans le tableau rectangulaire qui représente la matrice M , il suffit de rayer toutes les lignes (resp. colonnes) dont l'indice appartient à \tilde{I} (resp. \tilde{J}) et de regrouper les boîtes restantes sans modifier leur ordre relatif pour obtenir la sous-matrice $\mathcal{M}_{I,J}$.

Exemple 1 : $m = n = 4$, $I = \{2, 3\}$, $J = \{1, 4\}$

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \quad \mathcal{M}_{I,J}(M) = \begin{bmatrix} a_{21} & a_{24} \\ a_{31} & a_{34} \end{bmatrix}.$$

Permutation de lignes et de colonnes

Soit $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$. Donnons-nous deux permutations $\sigma \in \mathfrak{S}_m$ et $\tau \in \mathfrak{S}_n$, et définissons $(\sigma, \tau) * M$ comme étant la matrice $[b_{ij}] \in \mathfrak{M}_{m,n}(A)$ telle que $b_{ij} = [a_{\sigma^{-1}(i), \tau^{-1}(j)}]$ pour tous i, j .

On a, pour toute $M \in \mathfrak{M}_{m,n}(A)$ et tous σ_1, σ_2 dans \mathfrak{S}_m et τ_1, τ_2 dans \mathfrak{S}_n :

$$(1) \quad (\text{Id}, \text{Id}) * M = M ; \quad (\sigma_1 \sigma_2, \tau_1 \tau_2) * M = (\sigma_1, \tau_1) * [(\sigma_2, \tau_2) * M].$$

Ces relations (1) traduisent donc le fait que l'opération

$$((\sigma, \tau), M) \mapsto (\sigma, \tau) * M$$

est une *action à gauche* du groupe produit $\mathfrak{S}_m \times \mathfrak{S}_n$ sur l'ensemble $\mathfrak{M}_{m,n}(A)$. Pour (σ, τ) fixé, l'application $f_{\sigma, \tau} : M \mapsto (\sigma, \tau) * M$ est un *automorphisme* du groupe abélien $\mathfrak{M}_{m,n}(A)$; et si A est une K -algèbre (K : corps commutatif), $f_{\sigma, \tau}$ est un *automorphisme* du K -ev $\mathfrak{M}_{m,n}(A)$.

Produit de matrices

DÉFINITION XI.1.3

Soit m, n, p trois entiers ≥ 1 , et des matrices $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$, $N = [b_{ij}] \in \mathfrak{M}_{n,p}(A)$. On appelle **produit des matrices M et N dans cet ordre**, et on note MN ou $M \cdot N$, la matrice $P = [c_{ij}] \in \mathfrak{M}_{m,p}(A)$ telle que

$$(\forall (i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, p \rrbracket) \quad c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

On ne définit donc le produit MN que lorsque le nombre de colonnes de la matrice de gauche M est égal au nombre de lignes de la matrice de droite N .

Exemple 2 : Prenons deux matrices carrées d'ordre 2 à coefficients dans \mathbb{C} :

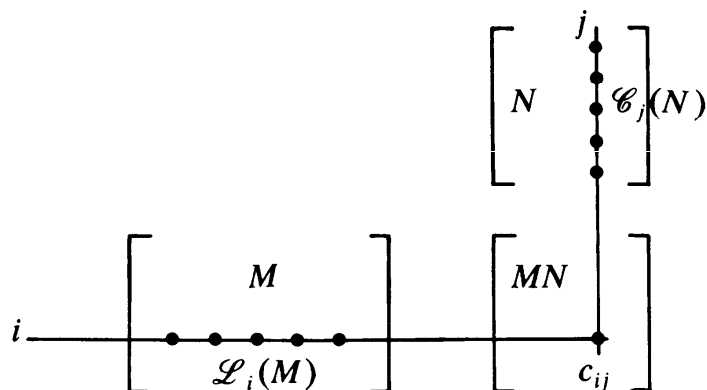
$$M = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} \quad \text{et} \quad N = \begin{bmatrix} 1 & 0 \\ \mu & 1 \end{bmatrix}.$$

Alors

$$MN = \begin{bmatrix} 1 + \lambda \mu & \lambda \\ \mu & 1 \end{bmatrix} \quad \text{et} \quad NM = \begin{bmatrix} 1 & \lambda \\ \mu & 1 + \lambda \mu \end{bmatrix}.$$

Visiblement le produit n'est pas commutatif, même pour des matrices carrées.

Remarque 2 : Pour calculer le produit MN , il est commode de disposer les tableaux rectangulaires représentant M et N de la façon indiquée ci-dessous :



Le terme c_{ij} de MN s'obtient à l'intersection de la ligne i de M et de la colonne j de N et se calcule comme « produit scalaire » $a_{i1} b_{1j} + \dots + a_{in} b_{nj}$ de la ligne $\mathcal{L}_i(M)$ par la colonne $\mathcal{C}_j(N)$, qui a un sens puisque toutes deux s'identifient à des éléments de A^n . On retient très facilement que le produit s'effectue « ligne par colonne », mais une précaution supplémentaire est à prendre si l'anneau A n'est pas commutatif : il faut faire attention d'écrire à gauche (resp. à droite) les termes de $\mathcal{L}_i(M)$ (resp. $\mathcal{C}_j(N)$).

THÉORÈME XI.1.1

Le produit de matrices possède les propriétés suivantes :

(I) Si m, n, p sont fixés dans \mathbb{N}^* , pour toutes M_1, M_2 dans $\mathfrak{M}_{m,n}(A)$ et N_1, N_2 dans $\mathfrak{M}_{n,p}(A)$, on a :

$$(M_1 + M_2) N_1 = M_1 N_1 + M_2 N_1 ;$$

$$M_1 (N_1 + N_2) = M_1 N_1 + M_1 N_2$$

(biadditivité du produit de matrices).

(II) Soit m, n, p, q dans \mathbb{N}^* et $M \in \mathfrak{M}_{m,n}(A)$, $N \in \mathfrak{M}_{n,p}(A)$, $P \in \mathfrak{M}_{p,q}(A)$ (d'où $MN \in \mathfrak{M}_{m,p}(A)$ et $NP \in \mathfrak{M}_{n,q}(A)$). Alors :

$$M(NP) = (MN)P \quad (\text{associativité du produit de matrices}).$$

(III) Si A est une K -algèbre, où K est un corps commutatif, pour $\lambda \in K$, $M \in \mathfrak{M}_{m,n}(A)$ et $N \in \mathfrak{M}_{n,p}(A)$, on a :

$$\lambda (MN) = (\lambda M) \cdot N = M \cdot (\lambda N).$$

Démonstration :

Les propriétés (I) et (III) se réduisent à des vérifications simples. Démontrons l'associativité. Pour cela posons :

$$M = [a_{ij}], N = [b_{ij}], P = [c_{ij}], MN = [x_{ij}], \\ NP = [y_{ij}], (MN)P = [u_{ij}], M(NP) = [v_{ij}].$$

On a : $x_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ ($1 \leq i \leq m, 1 \leq j \leq p$), d'où :

$$u_{ij} = \sum_{l=1}^p x_{il} c_{lj} = \sum_{l=1}^p \left(\sum_{k=1}^n a_{ik} b_{kl} \right) c_{lj} = \\ = \sum_{(k,l) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket} (a_{ik} b_{kl}) c_{lj} \quad (1 \leq i \leq m, 1 \leq j \leq q).$$

On calcule de même v_{ij} :

$$v_{ij} = \sum_{(k,l) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket} a_{ik} (b_{kl} c_{lj}) \quad (1 \leq i \leq m, 1 \leq j \leq q).$$

On utilise alors l'associativité du produit dans l'anneau A pour voir que $u_{ij} = v_{ij}$ pour tous i, j et l'assertion (II) est démontrée. ■

Supposons en particulier que A soit une K -algèbre (K : corps commutatif). Une conséquence du théorème XI.1.1 est que, si $M_0 \in \mathfrak{M}_{m,n}(A)$ est fixée (resp. $N_0 \in \mathfrak{M}_{n,p}(A)$ est fixée), l'application $N \mapsto M_0 N$,

$$\mathfrak{M}_{n,p}(A) \rightarrow \mathfrak{M}_{m,p}(A) \quad (\text{resp. } M \mapsto MN_0, \mathfrak{M}_{m,n}(A) \mapsto \mathfrak{M}_{m,p}(A))$$

est K -linéaire.

Transposition

DÉFINITION XI.1.4

Si $M \in \mathfrak{M}_{m,n}(A)$, $M = [a_{ij}]$, on appelle **transposée de M** , et on note tM , la matrice $N = [b_{ij}] \in \mathfrak{M}_{n,m}(A)$ telle que $b_{ij} = a_{ji}$ pour tous i, j .

On dit parfois que tM s'obtient à partir de M par échange des lignes et des colonnes.

Il est clair que, pour m et n fixés, l'application $\mathfrak{M}_{m,n}(A) \rightarrow \mathfrak{M}_{n,m}(A)$, $M \mapsto {}^tM$ est bijective, son application réciproque étant $N \mapsto {}^tN$.

C'est un *isomorphisme de groupes abéliens* (en particulier ${}^t(M_1 + M_2) = {}^tM_1 + {}^tM_2$) et, si A est une K -algèbre (K : corps commutatif), c'est un *isomorphisme de K -ev*.

Donnons-nous m, n, p dans \mathbb{N}^* , et soit $M = [a_{ij}] \in \mathfrak{M}_{m,n}(A)$ et $N = [b_{ij}] \in \mathfrak{M}_{n,p}(A)$. Soit $MN = [c_{ij}] \in \mathfrak{M}_{m,p}(A)$, d'où $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ ($1 \leq i \leq m, 1 \leq j \leq p$).

Puisque ${}^tM \in \mathfrak{M}_{n,m}(A)$ et ${}^tN \in \mathfrak{M}_{p,n}(A)$, on peut calculer le produit ${}^tN {}^tM \in \mathfrak{M}_{p,m}(A)$. Si c'_{ij} est son terme général ($1 \leq i \leq p, 1 \leq j \leq m$) il vaut $c'_{ij} = \sum_{k=1}^n b_{ki} a_{jk}$ qu'on compare avec le terme général c''_{ij} de ${}^t(MN)$ qui vaut $c''_{ij} = \sum_{k=1}^n a_{jk} b_{ki}$, d'où on déduit :

PROPOSITION XI.1.2

Si $M \in \mathfrak{M}_{m,n}(A)$, $N \in \mathfrak{M}_{n,p}(A)$, et si l'anneau A est commutatif, on a ${}^t(MN) = {}^tN {}^tM$.

Produit par blocs

Fixons des entiers m, n, r, s tous ≥ 1 et supposons données des suites d'entiers non nuls (m_1, m_2, \dots, m_r) , (n_1, n_2, \dots, n_s) telles que $m_1 + m_2 + \dots + m_r = m$ et $n_1 + n_2 + \dots + n_s = n$. Notons $s_0 = t_0 = 0$, $s_i = m_1 + m_2 + \dots + m_i$ et $t_j = n_1 + n_2 + \dots + n_j$ ($1 \leq i \leq r, 1 \leq j \leq s$) et

$$I_{i+1} = \llbracket s_i + 1, s_{i+1} \rrbracket, \quad J_{j+1} = \llbracket t_j + 1, t_{j+1} \rrbracket$$

($0 \leq i \leq r-1, 0 \leq j \leq s-1$). La suite (I_1, I_2, \dots, I_r) est un *partage de* $\llbracket 1, m \rrbracket$ en r intervalles consécutifs, ce qui signifie que pour $k < l$, si $\alpha \in I_k$ et $\beta \in I_l$, alors $\alpha < \beta$.

De même (J_1, J_2, \dots, J_s) est un partage de $\llbracket 1, n \rrbracket$ en s intervalles consécutifs.

Les ensembles $\Lambda_{i,j} = (I_i \times J_j)_{(i,j) \in \llbracket 1,r \rrbracket \times \llbracket 1,s \rrbracket}$ forment un partage de $\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$. Si nous visualisons $\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$ comme le tableau des indices des matrices de type (m, n) , ce partage correspond à une subdivision de ce tableau par des lignes horizontales correspondant aux entiers (s_i) et des lignes verticales correspondant aux entiers (t_j) .

Exemple 3 : $m = 4, n = 5, r = 2, s = 2, m_1 = 2, m_2 = 2, n_1 = 2, n_2 = 3$:

(1, 1)			(1, 5)
•	•	•	•
•	•	•	•
•	•	•	•
(4, 1)			(4, 5)

Pour (i, j) fixé dans $\llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket$, l'application

$$(\alpha, \beta) \mapsto (s_{i-1} + \alpha, t_{j-1} + \beta), \quad \llbracket 1, m_i \rrbracket \times \llbracket 1, n_j \rrbracket \rightarrow \Lambda_{i,j}$$

est bijective.

Soit alors une matrice $M \in \mathfrak{M}_{m,n}(A)$, $M = [a_{\lambda,\mu}]$. Nous lui associons, pour chaque $(i, j) \in \llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket$, la sous-matrice $\mathcal{M}_{I_i, J_j}(M) = [a_{s_{i-1} + \alpha, t_{j-1} + \beta}]$ $1 \leq \alpha \leq m_i$, $1 \leq \beta \leq m_j$, appartenant à $\mathfrak{M}_{m_i, n_j}(A)$, sous-matrice que nous noterons pour abréger $\text{Bl}_{i,j}(M)$, et que nous appellerons **bloc d'indice (i, j) de M** (défini par les suites (m_1, \dots, m_r) ; (n_1, \dots, n_s)).

La correspondance $M \rightarrow (\text{Bl}_{i,j}(M))_{(i,j) \in \llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket}$ établit une bijection de $\mathfrak{M}_{m,n}(A)$ sur l'ensemble des familles $(B_{ij})_{(i,j) \in \llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket}$ telles que $B_{i,j} \in \mathfrak{M}_{m_i, n_j}(A)$ pour tous i, j . Si (B_{ij}) est une telle famille, la matrice $M \in \mathfrak{M}_{m,n}(A)$ qui lui correspond sera appelée la **matrice de blocs $B_{i,j}$** (bien entendu cette expression n'a de sens que si les suites (m_i) et (n_j) ont été préalablement fixées). En abrégé cela revient à considérer la matrice M comme une matrice dont les éléments sont eux-mêmes des matrices. Si chacun des blocs a une seule ligne et une seule colonne, on retrouve la définition initiale de M , en identifiant $[a] \in \mathfrak{M}_{1,1}(A)$ avec $a \in A$. Il est évident que le passage d'une matrice à ses blocs est compatible avec l'addition matricielle, et avec le produit par un scalaire, mais il est bien plus remarquable (et cela se révèle très utile en calcul matriciel) que le passage d'une matrice à ses blocs soit aussi compatible avec le produit des matrices, en prenant bien sûr la précaution que tous les produits puissent s'effectuer :

THÉORÈME XI.1.2

Soit m, n, p ; r, s, t des entiers ≥ 1 ; $(m_i)_{1 \leq i \leq r}$, $(n_j)_{1 \leq j \leq s}$, $(p_k)_{1 \leq k \leq t}$ des suites d'entiers ≥ 1 telles que $\sum_{i=1}^r m_i = m$, $\sum_{j=1}^s n_j = n$, $\sum_{k=1}^t p_k = p$ et soit $M \in \mathfrak{M}_{m,n}(A)$, $N \in \mathfrak{M}_{n,p}(A)$.
Notons $(B_{ij})_{1 \leq i \leq r, 1 \leq j \leq s}$ les blocs de M définis par les suites (m_i) et (n_j) , $(C_{j,k})_{1 \leq j \leq s, 1 \leq k \leq t}$ les blocs de N définis par les suites (n_j) et (p_k) , (de sorte que si $(i, j, k) \in \llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket \times \llbracket 1, t \rrbracket$, le produit matriciel $B_{i,j} C_{j,k}$ est défini). Pour tous $(i, k) \in \llbracket 1, r \rrbracket \times \llbracket 1, t \rrbracket$, soit $P_{i,k}$ la matrice $\sum_{j=1}^s B_{i,j} C_{j,k} \in \mathfrak{M}_{m_i, p_k}(A)$.
Alors la matrice $P = MN$ est la matrice $P \in \mathfrak{M}_{m,p}(A)$ dont les $P_{i,k}$ sont les blocs associés aux suites (m_i) et (p_k) .

Autrement dit, pour calculer le produit MN , on décompose ces matrices en blocs $(B_{i,j})$ et $(C_{j,k})$ et on effectue le produit des blocs, puis on considère les $(P_{i,k})$ obtenus comme les blocs de la matrice cherchée MN .

Démonstration :

Nous poserons $M = [a_{\lambda\mu}]$, $N = [b_{\mu\nu}]$, $MN = [c_{\lambda\nu}]$,
 $P = [c'_{\lambda\nu}]$, $s_0 = t_0 = u_0 = 0$, $s_i = \sum_{l \leq i} m_l$, $t_j = \sum_{l \leq j} n_l$, $u_k = \sum_{l \leq k} p_l$
($i \leq r, j \leq s, k \leq t$).

Fixons $(i, k) \in \llbracket 1, r \rrbracket \times \llbracket 1, t \rrbracket$ et $(\alpha, \beta) \in \llbracket 1, m_i \rrbracket \times \llbracket 1, p_k \rrbracket$.

Le terme du produit MN d'indice $(\lambda, \nu) = (s_{i-1} + \alpha, u_{k-1} + \beta)$ est

$$c_{s_{i-1} + \alpha, u_{k-1} + \beta} = \sum_{\mu=1}^n a_{\lambda\mu} b_{\mu\nu} = c_{\lambda\nu}.$$

Dans la matrice P dont les blocs sont les (P_{uv}) , le terme $c'_{\lambda\nu}$ est le terme d'indice (α, β) de la matrice $P_{i,k} \in \mathfrak{M}_{m_i, p_k}(A)$. Or $P_{i,k} = \sum_{j=1}^s B_{i,j} C_{j,k}$. Si

$j \in \llbracket 1, s \rrbracket$, le terme d'indice (α, β) de $B_{i,j} C_{j,k}$ est $\sum_{\rho=t_{j-1}+1}^{t_j} a_{\lambda\rho} b_{\rho\nu}$, d'où

$$c'_{\lambda\nu} = \sum_{j=1}^s \left(\sum_{\rho=t_{j-1}+1}^{t_j} a_{\lambda\rho} b_{\rho\nu} \right) = \sum_{\rho=1}^n a_{\lambda\rho} b_{\rho\nu} = c_{\lambda\nu}. \quad \blacksquare$$

Remarque 3 : Avec les notations qui précèdent le théorème XI.1.2, supposons $m = n$, $r = s$ et $m_i = n_i$ pour $1 \leq i \leq r$. Alors $M \in \mathfrak{M}_n(A)$ est dite **diagonale par blocs** ssi $\text{Bl}_{i,j}(M) = 0$ pour $i \neq j$.

On dit que les $\text{Bl}_{i,i}(M)$ sont les **blocs diagonaux** de M et on écrit :

$$M = \text{Diag}(\text{Bl}_{1,1}(M), \text{Bl}_{2,2}(M), \dots, \text{Bl}_{r,r}(M)).$$

En particulier si les blocs ne contiennent chacun qu'un élément, on définit ainsi les *matrices carrées diagonales* $M = \text{Diag}(d_1, d_2, \dots, d_n)$ où $d_i \in A$ ($1 \leq i \leq n$).

Par exemple $M = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}$ est diagonale par blocs (avec $r = s = 2$).

On vérifie immédiatement que la somme et le produit de matrices diagonales par blocs le sont encore dès que les dimensions sont compatibles.

Dans les exercices A désigne un anneau commutatif et K un corps commutatif.

Exercice 1 : Soit n_1, n_2, \dots, n_{k+1} des entiers ≥ 1 ($k \geq 2$). On donne des matrices $M_r \in \mathfrak{M}_{n_r, n_{r+1}}(A)$ ($1 \leq r \leq k$), $M_r = [a_{ij}^{(r)}]$. Expliciter le terme général de la matrice $M_1 M_2 \dots M_k$.

Application : Lorsque $A = \mathbb{R}$, si tous les coefficients des M_r sont dans \mathbb{R}_+ (resp. \mathbb{R}_+^*), il en est de même de ceux de $M_1 M_2 \dots M_k$.

Exercice 2 : On décompose les matrices $M \in \mathfrak{M}_{m,n}(A)$ par blocs relativement aux suites (m_1, \dots, m_r) ; (n_1, \dots, n_s) . Faire opérer le groupe $\mathfrak{S}_r \times \mathfrak{S}_s$ sur $\mathfrak{M}_{m,n}(A)$ en permutant les blocs par lignes et par colonnes.

Exercice 3 : Soit n un entier ≥ 2 . Une matrice $M = [a_{ij}] \in \mathfrak{M}_n(A)$ est dite *en damier* ssi $a_{ij} = 0$ pour $j - i$ impair. Montrer que si $M \in \mathfrak{M}_n(A)$, et $N \in \mathfrak{M}_n(A)$ sont en damier, il en est de même de leur somme et de leur produit. Généraliser à des matrices en d

Exercice 4 : Soit K un corps commutatif fini, de cardinal q . Pour $(m, n) \in \llbracket 1, 4 \rrbracket^2$, dénombrer les matrices $M \in \mathfrak{M}_{m,n}(K)$ dont toutes les lignes sont distinctes ainsi que toutes les colonnes.

Exercice 5 : Soit E_{ij} une matrice élémentaire de $\mathfrak{M}_{m,n}(A)$ et $F_{k,l}$ une matrice élémentaire de $\mathfrak{M}_{n,p}(A)$. Étudier le produit $E_{i,j} \times F_{k,l}$ et retrouver ainsi la formule générale de multiplication des matrices en utilisant la « distributivité ».

Exercice 6 : Dans $\mathfrak{M}_4(\mathbb{C})$ considérons les matrices

$$A = \begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 1 \\ 0 & 0 & 1 & \beta \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 1 & \alpha & 0 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & 1 & \beta \end{bmatrix}.$$

Calculer A^2 , ABA et B^3 en opérant par blocs.

§ XI.2 MATRICES CARRÉES

Dans ce qui suit, A désignera un *anneau commutatif* non nul, et K un corps commutatif.

Si l'entier $n \geq 1$ est fixé, le produit des matrices définit sur $\mathfrak{M}_n(A)$ une loi interne associative, et l'on constate immédiatement que la matrice $I_n = [\delta_{ij}]$, où $\delta_{ii} = 1_A$ pour tout i et $\delta_{ij} = 0$ si $i \neq j$ est élément neutre pour ce produit. Compte tenu du théorème XI.1.1 et de la proposition XI.1.1, nous avons donc :

THÉORÈME XI.2.1

L'ensemble $\mathfrak{M}_n(A)$, muni de son addition et du produit des matrices, est un **anneau non nul**, dont l'élément unité est I_n . Si A est une K -algèbre, le K -ev $\mathfrak{M}_n(A)$, muni de cette structure d'anneau est une **K -algèbre unifère**. Le K -ev $\mathfrak{M}_n(K)$ est de **dimension finie**, égale à n^2 et une **base** de ce K -ev est formée des **matrices élémentaires** (E_{ij}) .

Le groupe des éléments inversibles de l'anneau $\mathfrak{M}_n(A)$ se note $GL(n, A)$ ou $GL_n(A)$ et s'appelle **groupe des matrices inversibles à coefficients dans A** .

Les matrices de la forme aI_n , où $a \in A$, sont dites **scalaires** : elles forment un sous-anneau de $\mathfrak{M}_n(A)$ isomorphe à A par l'application $a \mapsto aI_n$.

L'anneau $\mathfrak{M}_n(A)$ n'est en général pas commutatif si $n \geq 2$, comme l'a déjà prouvé l'exemple 2 du § XI.1. Cependant les matrices scalaires sont *permutables* avec toute matrice.

Si $a \in A$ est *inversible* dans A , on a : $aI_n \in GL(n, A)$.

Pour M et N dans $\mathfrak{M}_n(A)$, on a : ${}^t(MN) = {}^tN {}^tM$. En conséquence tM est inversible ssi M l'est, et si c'est le cas : $({}^tM)^{-1} = {}^t(M^{-1})$ et si M et N sont toutes deux inversibles $({}^t(MN))^{-1} = {}^t[(MN)^{-1}] = ({}^tN$

DÉFINITION XI.2.1

Soit $M = [a_{ij}] \in \mathfrak{M}_n(A)$. On dit que M est **trigonale supérieure** (resp. **inférieure**) ssi $a_{ij} = 0$ pour $i > j$ (resp. pour $i < j$). On dit que M est **unipotente supérieure** (resp. **inférieure**) ssi $a_{ij} = 0$ pour $i > j$ (resp. pour $i < j$) et $a_{ii} = 1_A$ pour tout $i \in \llbracket 1, n \rrbracket$. On dit que M est **diagonale** ssi $a_{ij} = 0$ pour $i \neq j$.

Nous noterons $\mathcal{T}_+(n, A)$, $\mathcal{T}_-(n, A)$, $\mathcal{U}_+(n, A)$, $\mathcal{U}_-(n, A)$ et $\text{Diag}(n, A)$ respectivement l'ensemble des matrices trigonales (on dit aussi triangulaires) supérieures, trigonales inférieures, unipotentes supérieures, unipotentes inférieures, diagonales. Une matrice diagonale M est définie de manière unique par la suite de ses termes $(a_{ii})_{1 \leq i \leq n}$ dits *diagonaux*. Si $(b_1, \dots, b_n) \in A^n$, nous noterons $\text{Diag}(b_1, \dots, b_n)$ la matrice diagonale dont $(b_i)_{1 \leq i \leq n}$ est la suite des termes diagonaux. Une matrice scalaire est diagonale, mais dès que $n \geq 2$ il y a bien plus de matrices diagonales que de matrices scalaires. Il est clair que $\mathcal{T}_+(n, A) \cap \mathcal{T}_-(n, A) = \text{Diag}(n, A)$ et $\mathcal{U}_+(n, A) \cap \mathcal{T}_-(n, A) = \mathcal{U}_-(n, A) \cap \mathcal{T}_+(n, A) = \{I_n\}$. Rien n'empêche d'étendre ces définitions à des matrices décomposées en blocs : par exemple $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ est trigonale supérieure par blocs, mais attention ! elle peut avoir des coefficients non nuls en dessous de sa diagonale principale.

DÉFINITION XI.2.2

Une matrice $M = [a_{ij}] \in \mathfrak{M}_n(A)$ est dite **symétrique** ssi $a_{ij} = a_{ji}$ pour tout $(i, j) \in \llbracket 1, n \rrbracket$, i.e. ssi ${}^tM = M$. Elle est dite **antisymétrique** ssi $a_{ij} = -a_{ji}$ pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, i.e. ssi ${}^tM = -M$ (ce qui entraîne $a_{ii} = 0$ si 2 est régulier dans A).

Les matrices symétriques (resp. antisymétriques) forment un *sous-groupe additif* de $\mathfrak{M}_n(A)$, et si A est une K -algèbre, un *sous- K -ev*, qu'on notera $\text{Sym}(n, A)$ (resp. $\text{Asym}(n, A)$). Notons $(E_{ij})_{(i,j) \in \llbracket 1, n \rrbracket^2}$ les matrices élémentaires. Une matrice est symétrique ssi elle s'écrit sous la forme $\left(\sum_{i=1}^n a_i E_{ii} \right) + \sum_{i < j} a_{ij} (E_{ij} + E_{ji})$, avec des a_i et des a_{ij} dans A , cette écriture étant alors unique.

De même, lorsque 2 est régulier dans A , une matrice M est antisymétrique ssi elle s'écrit $\sum_{i < j} a_{ij} (E_{ij} - E_{ji})$, cette écriture étant alors unique, et l'on voit que ses termes diagonaux sont nuls. Lorsque l'indicateur de torsion de A est égal à 2, toute matrice symétrique est simultanément antisymétrique.

THÉORÈME XI.2.2

Soit K un corps commutatif de caractéristique $\neq 2$, et $n \geq 2$. Alors les sous- K -ev $\text{Sym}(n, K)$ et $\text{Asym}(n, K)$ de $\mathfrak{M}_n(K)$ sont **supplémentaires**, leurs dimensions respectives étant $\frac{n(n+1)}{2}$ et $\frac{n(n-1)}{2}$.

Démonstration :

Tout d'abord, si $M \in \text{Sym}(n, K) \cap \text{Asym}(n, K)$, on a $M = {}^tM = -{}^tM$ d'où $M = -M$, d'où $M = 0$ puisque K est de caractéristique $\neq 2$. Donc $\text{Sym}(n, K) \cap \text{Asym}(n, K) = \{0\}$.

Ensuite, si $M \in \mathfrak{M}_n(K)$, la matrice $N = \frac{1}{2}(M + {}^tM)$ est symétrique, la matrice $P = \frac{1}{2}(M - {}^tM)$ est antisymétrique, et on a : $M = N + P$, d'où finalement $\mathfrak{M}_n(K) = \text{Sym}(n, K) \oplus \text{Asym}(n, K)$.

On a vu précédemment que $\text{Sym}(n, K)$ (resp. $\text{Asym}(n, K)$) est engendré comme K -ev par la famille $((E_{ii})_{1 \leq i \leq n}, (E_{ij} + E_{ji})_{i < j})$ qui possède $\frac{n(n+1)}{2}$ termes (resp. par la famille $(E_{ij} - E_{ji})_{i < j}$ qui possède $\frac{n(n-1)}{2}$ termes). Donc les dimensions d et e de ces K -ev vérifient : $d \leq \frac{n(n+1)}{2}$, $e \leq \frac{n(n-1)}{2}$, et comme $d + e = \dim \mathfrak{M}_n(K) = n^2$, on en déduit que $d = \frac{n(n+1)}{2}$, $e = \frac{n(n-1)}{2}$, ce qui entraîne que les familles génératrices mises en évidence sont des bases de ces sous- K -ev. ■

Matrices diagonales

Soit $M = \text{Diag}(a_1, a_2, \dots, a_n)$ et $N = \text{Diag}(b_1, b_2, \dots, b_n)$ deux matrices diagonales dans $\mathfrak{M}_n(A)$. Il est immédiat que

$$M + N = \text{Diag}(a_1 + b_1, \dots, a_n + b_n)$$

et $MN = \text{Diag}(a_1 b_1, \dots, a_n b_n)$, d'où :

PROPOSITION XI.2.1

|| L'ensemble $\text{Diag}(n, A)$ est un **sous-anneau** de $\mathfrak{M}_n(A)$, isomorphe à l'anneau produit A^n . Si A est une K -algèbre, $\text{Diag}(n, A)$ est une **sous- K -algèbre** de $\mathfrak{M}_n(A)$.

Une matrice diagonale $\text{Diag}(a_1, a_2, \dots, a_n)$ est **inversible** dans $\mathfrak{M}_n(A)$ ssi tous les a_i sont **inversibles dans A** . En particulier, si $A = K$, une matrice $\text{Diag}(a_1, \dots, a_n)$ est dans $\text{GL}(n, K)$ ssi $(\forall i) a_i \neq 0$.

Matrices trigonales**THÉORÈME XI.2.3**

|| (I) L'ensemble $\mathcal{T}_+(n, A)$ est un **sous-anneau** de $\mathfrak{M}_n(A)$ et, si A est une K -algèbre, c'est une **sous- K -algèbre** de $\mathfrak{M}_n(A)$.
 (II) Soit $M \in \mathcal{T}_+(n, A)$. On a : $M \in \text{GL}(n, A)$ ssi M est **inversible dans l'anneau $\mathcal{T}_+(n, A)$** , et cela se produit ssi les éléments diagonaux de M sont **inversibles dans A** .

(III) Prenons $A = K$. La sous- K -algèbre $\mathcal{T}_+(n, A)$ est un sous- K -ev de dimension $\frac{n(n+1)}{2}$ de $\mathfrak{M}_n(K)$. Une matrice $M \in \mathcal{T}_+(n, K)$ est **inversible dans $\mathcal{T}_+(n, K)$** ssi elle appartient à $GL(n, K)$, et cela se produit ssi ses éléments diagonaux sont dans K^* .

Démonstration :

(I) Soit $(E_{i,j})_{(i,j) \in \llbracket 1, n \rrbracket^2}$ les matrices élémentaires dans $\mathfrak{M}_n(A)$. L'ensemble $\mathcal{T}_+(n, A)$ est formé des matrices $\sum_{1 \leq i \leq j \leq n} a_{ij} E_{i,j}$ avec des a_{ij} dans A , cette écriture étant d'ailleurs unique.

La table de multiplication des (E_{ij}) est immédiate à écrire : pour tous i, j, k, l , on a $E_{ij} E_{kl} = 0$ si $j \neq k$ et $E_{ij} E_{jl} = E_{il}$. En particulier, si $i \leq j$ et $k \leq l$, on a toujours $E_{ij} E_{kl} \in \mathcal{T}_+(n, A)$ ce qui prouve que $\mathcal{T}_+(n, A)$ est stable pour le produit de matrices. De plus $I_n \in \mathcal{T}_+(n, A)$ et $\mathcal{T}_+(n, A)$ est évidemment un sous-groupe additif de $\mathfrak{M}_n(A)$, et une sous- K -algèbre si A est une K -algèbre.

(II) Remarquons d'abord que, si $M = \sum_{i \leq j} a_{ij} E_{ij}$ et $N = \sum_{i \leq j} b_{ij} E_{ij}$ sont trigonales supérieures, alors $MN = \sum_{i \leq j, k \leq l} a_{ij} b_{kl} E_{ij} E_{kl} = \sum_{i \leq j \leq l} a_{ij} b_{jl} E_{il}$, et qu'en particulier les éléments diagonaux de MN sont

$$(a_{11} b_{11}, a_{22} b_{22}, \dots, a_{nn} b_{nn}).$$

Si donc M est inversible dans $\mathcal{T}_+(n, A)$ et que N est son inverse dans $\mathcal{T}_+(n, A)$, alors $MN = I_n$, donc d'une part $M \in GL(n, A)$ et d'autre part $a_{ii} b_{ii} = 1_A$ pour tout i , donc les a_{ii} sont inversibles dans A .

Si maintenant M est inversible dans $\mathfrak{M}_n(A)$, son inverse s'écrit $N = \sum_{(k,l) \in \llbracket 1, n \rrbracket^2} b_{kl} E_{kl}$. Alors $I_n = MN = \sum_{i \leq j, 1 \leq l \leq n} a_{ij} b_{jl} E_{il}$, d'où

$$\sum_{i \leq j} a_{ij} b_{jl} = \delta_{il}$$

pour tous i et l . On en déduit d'abord $a_{nn} b_{nn} = 1$ et $a_{nn} b_{nl} = 0$ si $l < n$: a_{nn} est donc inversible dans A (et donc régulier dans A), et $b_{nl} = 0$ si $l < n$. On considère ensuite $i = n - 1$, puis $i = n - 2$, etc... et par récurrence on voit de même que $a_{ii} b_{ii} = 1$ pour tout i et que $b_{kl} = 0$ si $k > l$, d'où d'une part le fait que $N \in \mathcal{T}_+(n, A)$, d'autre part le fait que les a_{ii} sont inversibles dans A .

Il reste à prouver que la condition : les a_{ii} sont inversibles dans A , est suffisante pour assurer que $M = \sum_{i \leq j} a_{ij} E_{ij}$ est inversible dans $\mathcal{T}_+(n, A)$.

On cherche l'inverse éventuel de M sous la forme $N = \sum_{k \leq l \leq n} b_{kl} E_{kl}$. Les conditions nécessaires et suffisantes sur les b_{kl} sont : a_{ii}

$1 \leq i \leq n$, $\sum_{i \leq j \leq l} a_{ij} b_{jl} = 0$ pour tous i, l avec $i \leq l$. En particulier on a $b_{ii} = (a_{ii})^{-1}$, puis par récurrence descendante il n'y a aucune difficulté à déterminer de proche en proche tous les b_{jl} de manière unique, en commençant par $b_{n-1, n}$, puis ceux de la ligne $n-2$, etc... d'où la conclusion que M est inversible dans $\mathcal{T}_+(n, A)$.

(III) Il reste à prouver que $\mathcal{T}_+(n, K)$ est un sous- K -ev de dimension $\frac{n(n+1)}{2}$ de $\mathfrak{M}_n(K)$. Cela résulte du fait qu'il est engendré par les éléments $(E_{ij})_{i \leq j}$ de la base des matrices élémentaires, en nombre $\frac{n(n+1)}{2}$. ■

Il va sans dire qu'on a un théorème analogue au théorème XI.2.3 pour les matrices trigonales inférieures.

COROLLAIRE

|| $\mathcal{U}_+(n, A)$ (resp. $\mathcal{U}_-(n, A)$) est un sous-groupe du groupe des éléments inversibles de $\mathcal{T}_+(n, A)$ (resp. $\mathcal{T}_-(n, A)$).

Trace d'une matrice carrée

Nous supposerons ici pour simplifier que l'anneau de base A est un corps commutatif K .

DÉFINITION XI.2.3

Si $M = [a_{ij}] \in \mathfrak{M}_n(K)$, on appelle **trace** de M , et on note $\text{Tr}(M)$, le scalaire $\sum_{i=1}^n a_{ii}$.

Il est immédiat que l'application $\mathfrak{M}_n(K) \rightarrow K, M \mapsto \text{Tr}(M)$ est une *forme linéaire* sur $\mathfrak{M}_n(K)$ ($\text{Tr}(M) = a_{11} + a_{22} + \dots + a_{nn}$ montre que l'application trace est une fonction polynomiale homogène de degré 1 en les a_{ij}).

PROPOSITION XI.2.2

|| Pour $M \in \mathfrak{M}_n(K)$ et $N \in \mathfrak{M}_n(K)$, on a : $\text{Tr}(MN) = \text{Tr}(NM)$.

Démonstration :

Soit $M = [a_{ij}]$, $N = [b_{ij}]$. Le calcul de MN et NM montre que

$$\text{Tr}(MN) = \sum_{(i, k) \in \llbracket 1, n \rrbracket^2} a_{ik} b_{ki} = \sum_{(k, i) \in \llbracket 1, n \rrbracket^2} b_{ik} a_{ki} = \text{Tr}(NM). \quad \blacksquare$$

En particulier, si on a p matrices M_1, M_2, \dots, M_p dans $\mathfrak{M}_n(K)$ ($p \geq 2$), et si σ est une *puissance du cycle canonique* $c = \begin{pmatrix} 1 & 2 & \dots & p \\ 2 & 3 & \dots & 1 \end{pmatrix} \in \mathfrak{S}_p$, on a $\text{Tr}(M_1 M_2 \dots M_p) = \text{Tr}(M_{\sigma(1)} M_{\sigma(2)} \dots M_{\sigma(p)})$.

Une autre conséquence importante est :

COROLLAIRE

|| Soit $M \in \mathfrak{M}_n(K)$ et $P \in GL(n, K)$. Alors
||
|| $\text{Tr}(M) = \text{Tr}(P^{-1}MP)$.

Démonstration :

$$\text{Tr}(P^{-1}MP) = \text{Tr}(MPP^{-1}) = \text{Tr}(MI_n) = \text{Tr}(M). \quad \blacksquare$$

Exercice 1 : Le corps de base est \mathbb{R} . Soit $M = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \in GL(2, \mathbb{R})$ et $N = {}^tM$. Montrer : quels que soient $n \in \mathbb{N}^*$, et $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{Z}^n$, $(\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{Z}^n$, la relation $M^{\alpha_1} N^{\beta_1} M^{\alpha_2} N^{\beta_2} \dots M^{\alpha_n} N^{\beta_n} = I_n$ entraîne $\alpha_1 = \dots = \alpha_n = 0 = \beta_1 = \dots = \beta_n$.

Exercice 2 : a) K est un corps commutatif. Soit $n \in \mathbb{N}^*$ et $T: \mathfrak{M}_n(K) \rightarrow \mathfrak{M}_n(K)^*$ qui associe, à chaque matrice $M \in \mathfrak{M}_n(K)$, la forme linéaire $T_M: X \mapsto \text{Tr}(MX)$ sur $\mathfrak{M}_n(K)$. Démontrer que T est un isomorphisme de K -ev.

b) Soit $\varphi \in (\mathfrak{M}_n(K))^*$ telle que $\varphi(MN) = \varphi(NM)$ pour toutes matrices M, N de $\mathfrak{M}_n(K)$. On note M_0 la matrice élément de $\mathfrak{M}_n(K)$ telle que $\varphi = T_{M_0}$ (cf. a)). Démontrer que $(\forall M \in \mathfrak{M}_n(K)) M_0 M = M M_0$. En déduire que $\varphi = \lambda \text{Tr}$ pour un $\lambda \in K$. En déduire que le sous- K -ev de $\mathfrak{M}_n(K)$ engendré par les matrices $MN - NM$ ($M \in \mathfrak{M}_n(K), N \in \mathfrak{M}_n(K)$) est un hyperplan \mathcal{H} .

Exercice 3 : K est un corps commutatif, $n \in \mathbb{N}$, $n \geq 2$. Pour tout entier $k \geq 0$, soit \mathcal{T}_k le sous-ensemble de $\mathcal{T}_+(n, K)$ formé des matrices $M = [a_{ij}]$ telles que $a_{ij} = 0$ pour $i > j - k$.

a) Montrer : $\mathcal{T}_n = \{0\} \subset \mathcal{T}_{n-1} \subset \dots \subset \mathcal{T}_1 \subset \mathcal{T}_0 = \mathcal{T}_+(n, K)$.

Si $M \in \mathcal{T}_k$ et $N \in \mathcal{T}_l$, alors $MN \in \mathcal{T}_{k+l}$.

b) Montrer : si $M \in \mathcal{T}_+(n, K)$, il y a équivalence entre :

(I) $M \in \mathcal{T}_1$; (II) M est nilpotente ; (III) $M^n = 0$.

c) Si $k \in \llbracket 1, n-1 \rrbracket$, quelles sont les matrices $M \in \mathcal{T}_+(n, K)$ qui vérifient : $M^k = 0$?

d) Chaque \mathcal{T}_k est un idéal bilatère de $\mathcal{T}_+(n, K)$. Quelle est sa dimension comme K -ev ?

Exercice 4 : K est un corps commutatif, n un entier fixé ≥ 2 .

a) Trouver les idéaux bilatères \mathfrak{b} de la K -algèbre $\mathcal{T}_+(n, K)$ qui sont *maximaux* pour l'inclusion (i.e. tels que $\mathfrak{b} \subsetneq \mathcal{T}_+(n, K)$, et les seuls idéaux bilatères de $\mathcal{T}_+(n, K)$ contenant \mathfrak{b} sont \mathfrak{b} et $\mathcal{T}_+(n, K)$). Réponse : il y en a n .

b) Trouver dans $\mathcal{T}_+(n, K)$ un idéal bilatère *minimal* (i.e. non nul et ne contenant strictement aucun autre idéal bilatère non nul).

Exercice 5 : Le corps de base est \mathbb{R} . Dans le \mathbb{R} -ev $\mathfrak{M}_2(\mathbb{R})$ on considère l'ensemble \mathcal{E} des matrices M telles que $M^2 + M - 6I_2 = 0$.

a) Si $M \in \mathcal{E}$ et $M \neq 2I_2, M \neq -3I_2$, prouver que le \mathbb{R} -ev $\mathcal{V}(M)$ engendré par M et I_2 est de dimension 2.

b) On donne M comme en a). Trouver les $N \in \mathcal{V}(M)$ telles que $N^2 = N$ et $N \neq 0, N \neq I_2$. On en trouvera deux, notées A et B . Vérifier que $AB = BA = 0$. Exprimer A et B sur la base (I_2, M) de $\mathcal{V}(M)$. Prouver que (A, B) est une autre base de $\mathcal{V}(M)$. Montrer que $\mathcal{V}(M)$ est une sous- \mathbb{R} -algèbre de $\mathfrak{M}_2(\mathbb{R})$ et que si $N \in \mathcal{V}(M)$ est inversible dans $\mathfrak{M}_2(\mathbb{R})$, son inverse appartient à $\mathcal{V}(M)$.

c) Trouver toutes les matrices M de l'ensemble \mathcal{E} .

Exercice 6 : On donne l'entier $n \geq 2$ et des entiers n_1, n_2, \dots, n_r ($r \geq 2$) tels que $n_1 + n_2 + \dots + n_r = n$. On considère les matrices $M \in \mathfrak{M}_n(K)$ (où K est un corps commutatif) comme *matrices par blocs* associées aux suites (n_1, n_2, \dots, n_r) , (n_1, n_2, \dots, n_r) .

a) Soit \mathcal{D} l'ensemble des matrices *diagonales par blocs* :

$$M = \text{Diag} (B_{11}, \dots, B_{rr}) (B_{ii} \in \mathfrak{M}_{n_i}(K)).$$

Prouver que \mathcal{D} est une sous-algèbre de $\mathfrak{M}_n(K)$ et préciser quels sont ses éléments inversibles.

b) Soit \mathcal{E} l'ensemble des matrices *triangulaires supérieures par blocs*, c'est-à-dire les matrices $M \in \mathfrak{M}_n(K)$ de blocs $(B_{ij})_{(i,j) \in \llbracket 1, r \rrbracket^2}$ tels que $B_{i,j} = 0$ pour $i > j$. Montrer que \mathcal{E} est une sous- K -algèbre de $\mathfrak{M}_n(K)$ et que $M \in \mathcal{E}$ est inversible dans \mathcal{E} ssi elle est inversible dans $\mathfrak{M}_n(K)$, la C.N.S. pour cela étant que l'on ait : $B_{i,i} \in \text{GL}(n_i, K)$ pour tout $i \in \llbracket 1, r \rrbracket$.

c) Vérifier que \mathcal{D} est une sous-algèbre de \mathcal{E} .

Exercice 7 : Le corps de base K est un corps commutatif. On donne n entier ≥ 2 et on note J_n la matrice $[a_{ij}]$ telle que $a_{i,i+1} = 1_K$ pour $i \in \llbracket 1, n-1 \rrbracket$ et $a_{ij} = 0$ pour tous les autres couples (i, j) .

a) Calculer $(J_n)^k$ pour $k \in \mathbb{N}$ (par convention $(J_n)^0 = I_n$).

b) Pour $S \in K[[X]]$, $S = \sum_{p \geq 0} a_p X^p$, on note $S(J_n) = \sum_{p \geq 0} a_p (J_n)^p$. Montrer que $S(J_n)$ est bien

définie, et que $S \mapsto S(J_n)$ est un homomorphisme de K -algèbres de $K[[X]]$ dans $\mathfrak{M}_n(K)$.

c) On prend $K = \mathbb{C}$. Dédurre de b) l'inverse de chacune des matrices suivantes $M \in \mathfrak{M}_n(\mathbb{C})$:

$$M_1 = \begin{bmatrix} 1 & \binom{\alpha}{1} & \binom{\alpha}{2} & \dots & \binom{\alpha}{n-1} \\ 0 & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \binom{\alpha}{1} \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix} \quad \text{où } \alpha \in \mathbb{C},$$

$$M_2 = \begin{bmatrix} 1 & a & \frac{a^2}{2!} & \dots & \frac{a^{n-1}}{(n-1)!} \\ 0 & & & & \vdots \\ \vdots & & & & \frac{a^2}{2!} \\ \vdots & & & & a \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix} \quad \text{où } a \in \mathbb{C}$$

Exercice 8 (Matrices bistochastiques) :

a) Soit E et F deux ensembles finis non vides et $\varphi : E \rightarrow \mathcal{P}(F)$ une application qui satisfait l'hypothèse

$$(\mathcal{H}) \quad \forall X \subset E, \quad \text{card} \left(\bigcup_{x \in X} \varphi(x) \right) \geq \text{card}(X).$$

Montrer qu'il existe $f : E \rightarrow F$ *injective* telle que $f(x) \in \varphi(x)$ pour tout $x \in E$ (« Lemme des mariages » ⁽¹⁾).

(Indication : raisonner par récurrence sur $n = \text{card}(E)$; pour passer de $n-1$ ($n-1 \geq 1$) à n , introduire l'ensemble \mathcal{E} des parties X de E non vides et *distinctes* de E telles que $\text{card} \left(\bigcup_{x \in X} \varphi(x) \right) = \text{card}(X)$. Si $\mathcal{E} = \emptyset$, choisir $a \in E$, $b \in \varphi(a)$, et appliquer l'hypothèse

de récurrence à $\psi : E \setminus \{a\} \rightarrow \mathcal{P}(F \setminus \{b\})$, $x \mapsto \varphi(x) \setminus \{b\}$; si $\mathcal{E} \neq \emptyset$, choisir $A \in \mathcal{E}$ quelconque ; appliquer l'hypothèse de récurrence d'abord à $\varphi|_A : A \rightarrow \mathcal{P}(F)$ et obtenir ainsi

⁽¹⁾ Ainsi nommé parce qu'il constitue la condition nécessaire et suffisante pour que, si chacun des Messieurs $x \in E$ connaît l'ensemble $\varphi(x)$ de Dames, alors chacun d'eux puisse épouser une dame de sa connaissance.

$\alpha : A \rightarrow F$ injective telle que $\alpha(a) \in \varphi(a)$ pour tout $a \in A$; puis appliquer l'hypothèse de récurrence à $\psi : B \rightarrow \mathcal{P}(F \setminus \alpha(A))$, $x \mapsto \varphi(x) \setminus \alpha(A)$, où $B = E \setminus A$ pour obtenir $\beta : B \rightarrow F$ telle que $\beta(b) \in \psi(b)$ pour tout $b \in B$ et β injective, et conclure en considérant $f : E \rightarrow F$ définie par $f|_A = \alpha$ et $f|_B = \beta$.)

b) Soit $n \in \mathbb{N}$ ($n \geq 2$) ; on note S l'ensemble des matrices $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$ telles que :

$$(\forall (i, j) \in \llbracket 1, n \rrbracket^2) \quad \sum_{k=1}^n a_{ik} = 1, \quad \sum_{k=1}^n a_{kj} = 1 \quad \text{et} \quad a_{ij} \in \mathbb{R}_+.$$

Déduire de a) que si $M = [a_{ij}] \in S$, il existe $\sigma \in \mathfrak{S}_n$ telle que $a_{\sigma(i),i} > 0$ pour tout $i \in \llbracket 1, n \rrbracket$. (Indication : à chaque $i \in \llbracket 1, n \rrbracket$, associer l'ensemble

$$\varphi(i) = \{j \in \llbracket 1, n \rrbracket \mid a_{i,j} > 0\} \in \mathcal{P}(\llbracket 1, n \rrbracket) ;$$

puis montrer, lorsque $I \in \mathcal{P}(\llbracket 1, n \rrbracket)$ est de cardinal $p \geq 1$, que $\text{card} \left(\bigcup_{i \in I} \varphi(i) \right) \geq p$, en

calculant de deux manières $\sum_{(i,j) \in I \times \llbracket 1, n \rrbracket, j \in \varphi(i)} a_{i,j}$.

c) Un élément $M \in S$ est dit *extrémal* ssi les relations $M_1 \in S$, $M_2 \in S$ et $M = \frac{1}{2}(M_1 + M_2)$ entraînent $M_1 = M_2 = M$. Montrer que toute *matrice de permutation*, i.e. du type

$$M = [\delta_{\sigma^{-1}(i),j}]_{(i,j) \in \llbracket 1, n \rrbracket^2} = P_\sigma$$

pour une $\sigma \in \mathfrak{S}_n$, (δ : symbole de Kronecker) est élément extrémal de S .

d) Soit M un élément *extrémal* de S . Montrer que pour tout réel $\lambda > 1$, et toute matrice $N \in S$, $N \neq M$, on a : $\lambda M + (1 - \lambda)N \notin S$. En déduire que pour $\lambda > 1$ et $N \in S$, $N \neq M$, l'un au moins des coefficients de $\lambda M + (1 - \lambda)N$ est < 0 . En déduire que si P_σ ($\sigma \in \mathfrak{S}_n$) est une matrice de permutation distincte de M , posant $M = [a_{i,j}]$, alors il existe $(k, l) \in \llbracket 1, n \rrbracket^2$ tel que $a_{kl} = 0$ et $\sigma^{-1}(k) = l$. Comparer ce résultat avec b). Conclure enfin : les éléments extrémaux de S sont exactement les $(P_\sigma)_{\sigma \in \mathfrak{S}_n}$.

Exercice 9 (formule du binôme) : Le corps de base commutatif K est de caractéristique 0, n un entier ≥ 2 . On considère deux matrices A et B de $\mathfrak{M}_n(K)$ qui sont *permutables* (i.e. $AB = BA$, on dit encore que A et B commutent). Montrer que, pour tout entier $k \geq 1$, on a l'égalité

$$(A + B)^k = \sum_{i=0}^k \binom{k}{i} A^{k-i} B^i.$$

Application : On donne $M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Calculer M^k en posant $M = I + B$ même

question avec $M = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$.

Exercice 10 : Soit K un corps commutatif. Montrer que l'application de $\mathfrak{M}_2(K)$ dans $\mathfrak{M}_4(K)$ qui transforme chaque matrice $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ en $\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$ est un homomorphisme de K -algèbres.

Exercice 11 : Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice carrée d'ordre 2 à coefficients dans un anneau commutatif quelconque. Montrer qu'on a la relation : $A^2 - (a + d)A + (ad - bc)I_2 = 0$.

Exercice 12 : On prend $K = \mathbb{C}$ et n entier ≥ 2 . Une matrice $N \in \mathfrak{M}_n(\mathbb{C})$ est dite *nilpotente* s'il existe un entier $r \geq 1$ tel que $N^r = 0$.

a) Montrer que si N est nilpotente, on peut définir $\exp(N) = \sum_{p \geq 0} \frac{1}{p!} N^p$.

b) Vérifier que $N = \begin{bmatrix} 0 & a & d & f \\ 0 & 0 & b & e \\ 0 & 0 & 0 & c \\ 0 & 0 & 0 & 0 \end{bmatrix}$ est nilpotente.

c) Pour $t \in \mathbb{C}$ on considère la matrice

$$U(t) = \begin{bmatrix} 1 & t & 2t + 2t^2 & 3t + \frac{17}{2}t^2 + 4t^3 \\ 0 & 1 & 4t & 5t + 12t^3 \\ 0 & 0 & 1 & 6t \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Montrer que $U(t) = \exp(tN)$ où N est une matrice nilpotente que l'on calculera.

d) Montrer que l'ensemble des matrices $\{U(t)\}_{t \in \mathbb{C}}$ est un sous-groupe de $GL(4, \mathbb{C})$.

Exercice 13 : Soit K un corps commutatif. On se propose d'étudier des groupes de matrices pour la multiplication des matrices.

a) Montrer que les matrices éléments d'un tel groupe sont nécessairement carrées et de même ordre.

b) Vérifier qu'une matrice carrée *idempotente* (c'est-à-dire telle que $M^2 = M$) constitue un tel groupe à elle seule. En donner un exemple.

c) En déduire que l'élément neutre d'un groupe de matrices n'est pas nécessairement I_n (matrice unité d'ordre n).

d) Montrer que, soit aucune des matrices du groupe n'est inversible dans $\mathfrak{M}_n(K)$, soit elles le sont toutes (dans ce cas préciser l'unité du groupe et l'inverse dans le groupe d'une matrice donnée du groupe). (Voir aussi exercice 7 du § XI.6.)

Exercice 14 : On prend $K = \mathbb{R}$ et $n = 3$. Soit

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

données. On désigne par \mathcal{E} l'ensemble des matrices de la forme $xA + yB$, où $x \in \mathbb{R}$, $y \in \mathbb{R}$.

a) Calculer $(A + B)^k$ pour $k \in \mathbb{N}$.

b) Montrer que la matrice $xA + yB$ n'est pas inversible dans $\mathfrak{M}_3(\mathbb{R})$.

c) Montrer que l'addition et la multiplication des matrices font de \mathcal{E} un corps. Préciser l'élément unité et donner l'expression de l'inverse dans \mathcal{E} de la matrice $xA + yB$.

d) Montrer que \mathcal{E} est isomorphe à \mathbb{C} .

Exercice 15 : Une matrice $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$ est dite *magique* s'il existe $s \in \mathbb{C}$ tel que pour $i \in \llbracket 1, n \rrbracket$, $\sum_j a_{ij} = s$, pour $j \in \llbracket 1, n \rrbracket$, $\sum_i a_{ij} = s$, $\sum_i a_{ii} = s$ et $\sum_i a_{i, n+1-i} = s$. Les matrices magiques forment un sous- K -ev de $\mathfrak{M}_n(\mathbb{C})$ dont on demande la dimension et une base. Exemple : $n = 3$. (Voir aussi exercice 10, § XII.2.)

§ XI.3 MATRICES ET APPLICATIONS LINÉAIRES

Dans ce §, K désigne un corps commutatif fixé. Tous les K -ev considérés seront supposés de dimension finie et non nuls. Si E est un tel K -ev de dimension n , les seules bases considérées sur E seront celles indexées par $\llbracket 1, n \rrbracket$, c'est-à-dire du type $\mathcal{B} = (e_1, e_2, \dots, e_n)$.

DÉFINITION XI.3.1

Soit E et F deux K -ev, de dimensions respectives n et p . Soit $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_p)$ des bases respectives de E et F et $u \in \text{Hom}_K(E, F)$.

On appelle **matrice de u dans les bases \mathcal{B} et \mathcal{C}** , et on note $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$, la matrice $[a_{ij}] \in \mathfrak{M}_{p, n}(K)$ telle que

$$\forall j \in \llbracket 1, n \rrbracket \quad u(e_j) = \sum_{i=1}^p a_{ij} f_i$$

(autrement dit, a_{ij} est la i -ième coordonnée, dans la base \mathcal{C} , de $u(e_j)$). Si $E = F$, et $\mathcal{B} = \mathcal{C}$, on écrit $\text{Mat}_{\mathcal{B}, \mathcal{B}}(u) = \text{Mat}_{\mathcal{B}}(u)$ et on dit que cette matrice est la **matrice de u dans \mathcal{B}** .

Exemple 1 : Soit $E = F$ et $\mathcal{B} = \mathcal{C}$. Pour tout $\lambda \in K$ la matrice de l'homothétie λId_E dans la base \mathcal{B} est λI_n . On constate que cette matrice ne dépend pas du choix de \mathcal{B} , mais ce phénomène est exceptionnel. En général il est bien évident que $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$ dépend de \mathcal{B} et de \mathcal{C} , même si $\mathcal{B} = \mathcal{C}$.

THÉORÈME XI.3.1

Fixons $E, F, \mathcal{B}, \mathcal{C}$ dans la définition XI.3.1. Alors l'application $\psi : u \mapsto \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$ est un **isomorphisme de K -ev** de $\text{Hom}_K(E, F)$ sur $\mathfrak{M}_{p, n}(K)$.

Démonstration :

Tout d'abord ψ est K -linéaire. En effet, soit λ et μ dans K , et u et v dans $\text{Hom}_K(E, F)$. Posons

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = [a_{ij}], \quad \text{Mat}_{\mathcal{B}, \mathcal{C}}(v) = [b_{ij}].$$

Pour chaque (i, j) , la i -ième coordonnée dans \mathcal{C} du vecteur

$$(\lambda u + \mu v)(e_j) = \lambda u(e_j) + \mu v(e_j)$$

est $\lambda a_{ij} + \mu b_{ij}$, d'où

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(\lambda u + \mu v) = \lambda \text{Mat}_{\mathcal{B}, \mathcal{C}}(u) + \mu \text{Mat}_{\mathcal{B}, \mathcal{C}}(v).$$

Montrons ensuite que ψ est injective : si $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = 0$, alors $u(e_j) = 0$ pour tout $j \in \llbracket 1, n \rrbracket$, donc $u = 0$ puisque \mathcal{B} est une base de E , d'où $\text{Ker}(\psi) = \{0\}$, comme annoncé.

Pour prouver la surjectivité de ψ il suffit de comparer les dimensions de $\text{Hom}_K(E, F)$ et de $\mathfrak{M}_{p, n}(K)$ toutes deux égales à np (cf. proposition XI.1.1 et théorème IX.4.6) et d'appliquer le théorème IX.4.8. On peut aussi appliquer directement le théorème VI.3.2. ■

THÉORÈME XI.3.2

Soit E, F, G trois K -ev munis de bases notées respectivement $\mathcal{B} = (e_1, \dots, e_n)$, $\mathcal{C} = (f_1, \dots, f_p)$ et $\mathcal{D} = (g_1, \dots, g_q)$. Alors, pour toutes $u \in \text{Hom}_K(E, F)$ et $v \in \text{Hom}_K(F, G)$, on a :

(I) $\text{Mat}_{\mathcal{B}, \mathcal{D}}(v \circ u) = \text{Mat}_{\mathcal{C}, \mathcal{D}}(v) \times \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$.

Démonstration :

Posons $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = [a_{ij}]$ et $\text{Mat}_{\mathcal{C}, \mathcal{D}}(v) = [b_{ij}]$. On a, si $j \in \llbracket 1, n \rrbracket$: $u(e_j) = \sum_{k=1}^p a_{kj} f_k$, et si $k \in \llbracket 1, p \rrbracket$, $v(f_k) = \sum_{i=1}^q b_{ik} g_i$, d'où :

$$\begin{aligned} v \circ u(e_j) &= \sum_{k=1}^p a_{kj} \left(\sum_{i=1}^q b_{ik} g_i \right) = \sum_{i=1}^q \left(\sum_{k=1}^p a_{kj} b_{ik} \right) g_i = \\ &= (\text{puisque } K \text{ est commutatif}) = \sum_{i=1}^q \left(\sum_{k=1}^p b_{ik} a_{kj} \right) g_i. \end{aligned}$$

Or on reconnaît dans $\sum_{k=1}^p b_{ik} a_{kj}$ le terme général c_{ij} de la matrice produit $\text{Mat}_{\mathcal{C}, \mathcal{D}}(v) \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$, d'où le théorème. ■

COROLLAIRE 1

Si $E = F = G$ et $\mathcal{B} = \mathcal{C} = \mathcal{D}$, on a en particulier :

(II) $\text{Mat}_{\mathcal{B}}(v \circ u) = \text{Mat}_{\mathcal{B}}(v) \times \text{Mat}_{\mathcal{B}}(u)$.

COROLLAIRE 2

Si E et F sont deux K -ev, rapportés respectivement aux bases \mathcal{B} et \mathcal{C} , de même dimension n , si u désigne un isomorphisme de E sur F et v l'isomorphisme réciproque, on a :

$$\text{Mat}_{\mathcal{C}, \mathcal{B}}(v) = [\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)]^{-1}.$$

Compte tenu de l'exemple 1, les théorèmes XI.3.1 et XI.3.2 entraînent :

THÉORÈME XI.3.3

Soit E un K -ev et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . L'application $\psi_{\mathcal{B}} : \text{Hom}_K(E) \rightarrow \mathfrak{M}_n(K)$, $u \mapsto \text{Mat}_{\mathcal{B}}(u)$ est un **isomorphisme de K -algèbres** (donc aussi un **isomorphisme d'anneaux**).

En conséquence, pour n donné, toutes les K -algèbres $\text{Hom}_K(E)$, ($\dim_K(E) = n$) sont isomorphes entre elles, puisqu'isomorphes à la K -algèbre $\mathfrak{M}_n(K)$. Prenons $E = K^n$ et $\mathcal{B} =$ base canonique de l

Nous voyons que : $u \mapsto \text{Mat}_{\text{Can}}(u), \text{Hom}_K(K^n) \mapsto \mathfrak{M}_n(K)$ est un isomorphisme (dit *canonique*) de K -algèbres, ce qui prouve que la K -algèbre $\mathfrak{M}_n(K)$ est isomorphe à au moins une K -algèbre du type $\text{Hom}_K(E)$, $\dim_K(E) = n$. En passant aux éléments inversibles des anneaux $\text{Hom}_K(E)$ et $\mathfrak{M}_n(K)$, le théorème XI.3.3 entraîne :

COROLLAIRE 1

|| Avec les notations du théorème XI.3.3, l'application $\psi_{\mathcal{B}}$ définit un **isomorphisme du groupe linéaire** $\text{GL}_K(E)$ sur le groupe $\text{GL}(n, K)$ des matrices carrées inversibles d'ordre n sur K .

Donc, pour n donné, tous les groupes linéaires $\text{GL}_K(E)$ ($\dim(E) = n$) sont isomorphes entre eux, puisque isomorphes à $\text{GL}(n, K)$.

En prenant $E = K^n$ et $\mathcal{B} = \text{Can}$, nous voyons que les groupes $\text{GL}_K(K^n)$ et $\text{GL}(n, K)$ sont canoniquement isomorphes, ce qui prouve que $\text{GL}(n, K)$ est isomorphe à au moins un groupe linéaire $\text{GL}_K(E)$, où $\dim_K(E) = n$.

Exemple 2 : Nous savons que le *centre* du groupe linéaire $\text{GL}_K(E)$ est le groupe des homothéties de E . Compte tenu de l'exemple 1, on en déduit que le *centre du groupe* $\text{GL}(n, K)$ est le groupe des matrices scalaires non nulles λI_n ($\lambda \in K^*$), ce que l'on peut démontrer directement. Inversement, des méthodes purement matricielles permettent, à l'aide du corollaire ci-dessus, d'étudier la structure des groupes linéaires. On en verra plus loin des exemples.

Exemple 3 : Soit E un K -ev de dimension $n \geq 2$ muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. On donne un hyperplan H , d'équation $\sum_{i=1}^n a_i x_i = 0$ dans \mathcal{B} , et une droite vectorielle D supplémentaire de H , dirigée par $v = \sum_{i=1}^n v_i e_i \notin H$, d'où $s = \sum_{i=1}^n a_i v_i \neq 0$. Soit u (resp. v) le projecteur sur H parallèlement à D (resp. sur D parallèlement à H). Nous allons calculer $\text{Mat}_{\mathcal{B}}(u)$ et $\text{Mat}_{\mathcal{B}}(v)$.

Soit $x = \sum_{i=1}^n x_i e_i \in E$. Si $\lambda \in K$, la condition nécessaire et suffisante pour que $x + \lambda v \in H$ est

$$\lambda s + \sum_{i=1}^n a_i x_i = 0, \quad \text{i.e.} \quad \lambda = -\frac{1}{s} \sum_{j=1}^n a_j x_j.$$

Par suite $u(x) = x - \frac{1}{s} \left(\sum_{j=1}^n a_j x_j \right) v$. En particulier $u(e_j) = e_j - \frac{a_j}{s} v$, d'où

$$\text{Mat}_{\mathcal{B}}(u) = I_n - \frac{1}{s} [v_i a_j]_{(i,j) \in \llbracket 1, n \rrbracket^2}.$$

Comme $u + v = \text{Id}_E$, on a : $\text{Mat}_{\mathcal{B}}(u) + \text{Mat}_{\mathcal{B}}(v) = I_n$, d'où :

$$\text{Mat}_{\mathcal{B}}(v) = \frac{1}{s} [v_i a_j]_{(i,j) \in \llbracket 1, n \rrbracket^2}.$$

Exemple 4 : Soit n un entier ≥ 2 . Pour $a \in K$, soit

$$T_a : K_{n-1}[X] \rightarrow K_{n-1}[X], \quad P(X) \mapsto P(X+a).$$

Calculons $\text{Mat}_{\mathcal{C}}(T_a)$, où \mathcal{C} désigne la base canonique $(1, X, X^2, \dots, X^{n-1})$ du K -ev K^n . On a :

$$T_a(X^k) = (X+a)^k = \sum_{j=0}^k \binom{k}{j} a^{k-j} X^j,$$

d'où immédiatement : $\text{Mat}_{\mathcal{C}}(T_a) = \left[\binom{j-1}{i-1} a^{j-i} \right]_{(i,j) \in \llbracket 1, n \rrbracket^2}$, en convenant comme d'habitude que $\binom{k}{l} = 0$ pour k, l dans \mathbb{N} avec $k < l$.

En particulier, $M(a) = \text{Mat}_{\mathcal{C}}(T_a)$ est unipotente supérieure. Comme il est clair que $T_a \circ T_b = T_{a+b}$ pour $(a, b) \in K^2$, on en déduit sans calcul $M(a)M(b) = M(a+b)$ pour $(a, b) \in K^2$. Donc $a \mapsto M(a)$ est un homomorphisme du groupe additif K dans le groupe $\mathcal{U}_+(n, K)$.

Exemple 5 : Matrices et sommes directes de K -ev.

Supposons que $E = \bigoplus_{i=1}^r E_i$ et $F = \bigoplus_{j=1}^r F_j$ (avec $r \geq 1$ et les E_i et les F_j non nuls). Soit \mathcal{B} (resp. \mathcal{C}) une base de E (resp. F) obtenue en juxtaposant, dans cet ordre, des bases $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r$ de E_1, E_2, \dots, E_r (resp. des bases $\mathcal{C}_1, \dots, \mathcal{C}_r$ de F_1, \dots, F_r). Pour chaque $i \in \llbracket 1, r \rrbracket$, soit

$$u_i \in \text{Hom}_K(E_i, F_i) \quad \text{et} \quad M_i = \text{Mat}_{\mathcal{B}_i, \mathcal{C}_i}(u_i).$$

Notons u l'unique élément de $\text{Hom}_K(E)$ tel que $u\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n u_i(x_i)$ pour tout $(x_1, \dots, x_r) \in E_1 \times E_2 \times \dots \times E_r$ et $M = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$.

Considérons M comme *matrice par blocs* attachée aux suites (d_1, \dots, d_r) , (e_1, \dots, e_r) ($d_i = \dim_K(E_i)$, $e_i = \dim_K(F_i)$) (cf. la fin du § XI.1). Les blocs $\text{Bl}_{i,j}(M)$ sont évidemment les suivants : $\text{Bl}_{i,i}(M) = M_i$ pour tout i , et $\text{Bl}_{i,j}(M) = 0$ pour $i \neq j$. En particulier si $e_i = d_i$ pour tout i , alors M est *diagonale par blocs* : $M = \text{Diag}(M_1, \dots, M_r)$.

Exemple 6 : Soit E un K -ev muni d'une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$. A $\sigma \in \mathfrak{S}_n$, associons l'endomorphisme u_σ de E défini par :

$$(\forall i \in \llbracket 1, n \rrbracket) \quad u_\sigma(e_i) = e_{\sigma(i)}.$$

On a donc

$$u_\sigma \left(\sum_{i=1}^n x_i e_i \right) = \sum_{i=1}^n x_{\sigma^{-1}(i)} e_i \quad \text{pour } (x_1, \dots, x_n) \in K^n.$$

Alors la matrice $P_\sigma = \text{Mat}_{\mathcal{B}}(u_\sigma)$ est $[a_{ij}]$ telle que : $a_{ij} = 0$ si $i \neq \sigma(j)$ et $a_{\sigma(j),j} = 1$ pour tout j . Avec le symbole de Kronecker, on a donc :

$$P_\sigma = [\delta_{\sigma^{-1}(i),j}]_{(i,j) \in \llbracket 1,n \rrbracket^2} = (\sigma, \text{Id}) * I_n$$

(notations du § XI.1). Les matrices P_σ s'appellent *matrices de permutation*. L'expression $P_\sigma = (\sigma, \text{Id}) * I_n$ montre, compte tenu de l'étude du § XI.1, que $P_{\sigma\tau} = P_\sigma P_\tau$ pour tous $\sigma \in \mathfrak{S}_n$, $\tau \in \mathfrak{S}_n$.

Voici une importante conséquence du théorème XI.3.3 et de son corollaire 1, qui s'obtient en appliquant le théorème IX.6.1.

COROLLAIRE 2

Soit $n \in \mathbb{N}^*$ et $M \in \mathfrak{M}_n(K)$. Les assertions suivantes sont équivalentes :

- (I) $M \in \text{GL}(n, K)$
- (II) M est **inversible à gauche** dans l'anneau $\mathfrak{M}_n(K)$
- (III) M est **inversible à droite** dans l'anneau $\mathfrak{M}_n(K)$
- (IV) M est **régulière à gauche** dans l'anneau $\mathfrak{M}_n(K)$
- (V) M est **régulière à droite** dans l'anneau $\mathfrak{M}_n(K)$.

On remarquera l'extrême fécondité du lien qui unit matrices et applications linéaires. Ce lien joue dans les deux sens, par exemple l'associativité de la multiplication matricielle devient évidente quand on interprète chaque matrice comme la matrice d'une application linéaire, de même que pour une matrice carrée M , la seule existence de N telle que $MN = I_n$ prouve que M est inversible alors qu'une preuve purement matricielle ne serait pas facile. A l'inverse la commodité de l'outil matriciel peut rendre des services pour l'étude d'applications linéaires.

Utilisation de matrices unicolonnes

Soit E, F deux K -ev, munis de bases respectives $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_p)$, et soit $u \in \text{Hom}_K(E, F)$. Posons

$$M = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = [a_{ij}].$$

A chaque vecteur $x = \sum_{i=1}^n x_i e_i \in E$, associons la matrice *unicolonne* $X \in \mathfrak{M}_{n,1}(K)$ telle que ${}^tX = [x_1, \dots, x_n]$, et la matrice

$Y \in \mathfrak{M}_{p,1}(K)$ telle que ${}^tY = [y_1, \dots, y_p]$, où $u(x) = \sum_{i=1}^p y_i f_i$. Alors

$$u(x) = \sum_{j=1}^n x_j u(e_j) = \sum_{j=1}^n x_j \left(\sum_{i=1}^p a_{ij} f_i \right) = \sum_{i=1}^p \left(\sum_{j=1}^n a_{ij} x_j \right) f_i ,$$

d'où
$$(\forall i \in \llbracket 1, p \rrbracket) \quad y_i = \sum_{j=1}^n a_{ij} x_j .$$

Autrement dit :

$$(3) \quad \boxed{Y = MX} .$$

D'ailleurs M est la seule matrice $M_1 \in \mathfrak{M}_{p,n}(K)$ telle que, pour tout $x \in E$, on ait $Y = M_1 X$, comme on le voit en donnant à x les valeurs particulières $x = e_i$ ($1 \leq i \leq n$).

Matrices de passage

Soit E un K -ev de dimension n , muni d'une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$. Si $(f_1, f_2, \dots, f_n) \in E^n$, on a (théorème VI.3.2) une unique $u \in \text{Hom}_K(E)$ telle que $u(e_i) = f_i$ pour $1 \leq i \leq n$. La matrice $\text{Mat}_{\mathcal{B}}(u) = [a_{ij}]$ s'appelle **matrice des vecteurs colonnes f_i dans la base \mathcal{B}** et peut se noter $\text{Mat}_{\mathcal{B}}(f_1, \dots, f_n)$: c'est l'unique matrice telle que :

$$(4) \quad f_j = \sum_{i=1}^n a_{ij} e_i \quad \text{pour tout } j \in \llbracket 1, n \rrbracket .$$

DÉFINITION XI.3.2

$\left\{ \begin{array}{l} \text{Soit deux bases } \mathcal{B} = (e_1, \dots, e_n) \text{ et } \mathcal{C} = (f_1, \dots, f_n) \text{ d'un } K\text{-ev de} \\ \text{dimension } n. \text{ La matrice des vecteurs } (f_1, \dots, f_n) \text{ dans la base } \mathcal{B} \\ \text{s'appelle } \textbf{matrice de passage de } \mathcal{B} \text{ à } \mathcal{C}. \end{array} \right.$

Nous noterons $P_{\mathcal{B}, \mathcal{C}}$ cette matrice, et (a_{ij}) ses termes.

Voici quelques propriétés élémentaires, mais fondamentales, de ces matrices.

1) On a :

$$(5) \quad \boxed{P_{\mathcal{B}, \mathcal{C}} = \text{Mat}_{\mathcal{C}, \mathcal{B}}(\text{Id}_E)} .$$

C'est une conséquence évidente de (4) et de la définition de la matrice d'une application linéaire dans un système de bases.

2) La matrice $P_{\mathcal{B}, \mathcal{C}}$ est **inversible**, et on a :

$$(6) \quad \boxed{(P_{\mathcal{B}, \mathcal{C}})^{-1} = P_{\mathcal{C}, \mathcal{B}}}.$$

En effet, d'après le théorème XI.3.2, on a :

$$\text{Mat}_{\mathcal{C}, \mathcal{B}} (\text{Id}_E) \text{Mat}_{\mathcal{B}, \mathcal{C}} (\text{Id}_E) = \text{Mat}_{\mathcal{B}} (\text{Id}_E) = I_n$$

d'où $P_{\mathcal{B}, \mathcal{C}} P_{\mathcal{C}, \mathcal{B}} = I_n$, ce qui suffit pour avoir (6) (cf. corollaire 2 du théorème XI.3.3).

3) Soit g_1, \dots, g_n des vecteurs de E . On a :

$$(7) \quad \text{Mat}_{\mathcal{B}} (g_1, \dots, g_n) = P_{\mathcal{B}, \mathcal{C}} \times \text{Mat}_{\mathcal{C}} (g_1, \dots, g_n).$$

En effet, soit $P_{\mathcal{B}, \mathcal{C}} = [a_{ij}]$ et $\text{Mat}_{\mathcal{C}} (g_1, \dots, g_n) = b_{ij}$. On a, pour $1 \leq j \leq n$:

$$g_j = \sum_{k=1}^n b_{kj} f_k = \sum_{k=1}^n b_{kj} \left(\sum_{i=1}^n a_{ik} e_i \right) = \sum_{i=1}^n \left(\sum_{k=1}^n a_{ik} b_{kj} \right) e_i,$$

d'où (7).

4) En conséquence, si \mathcal{D} est une troisième base (g_1, \dots, g_n) de E , on a :

$$(8) \quad \boxed{P_{\mathcal{B}, \mathcal{D}} = P_{\mathcal{B}, \mathcal{C}} \times P_{\mathcal{C}, \mathcal{D}}}.$$

5) Soit $P = P_{\mathcal{B}, \mathcal{C}}$ et $x \in E$ de coordonnées (x_1, \dots, x_n) dans \mathcal{B} et (y_1, \dots, y_n) dans \mathcal{C} . Notons X (resp. Y) la matrice *unicolonne* ($X \in \mathfrak{M}_{n,1}(K)$) dont les termes sont (x_1, \dots, x_n) (resp. (y_1, \dots, y_n)). Alors

$$(9) \quad \boxed{X = PY} \quad (\text{formule de changement de coordonnées}).$$

En effet,

$$x = \sum_{i=1}^n x_i e_i = \sum_{j=1}^n y_j f_j = \sum_{j=1}^n y_j \left(\sum_{i=1}^n a_{ij} e_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} y_j \right) e_i,$$

d'où $(\forall i) \quad x_i = \sum_{j=1}^n a_{ij} y_j$, c'est-à-dire (9).

Changement de base

THÉORÈME XI.3.4

Soit E, F deux K -ev de dimensions respectives n et p , $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ (resp. $\mathcal{C} = (f_1, \dots, f_p)$ et $\mathcal{C}' = (f'_1, \dots, f'_p)$) deux bases de E (resp. de F), $P = P_{\mathcal{B}, \mathcal{B}'}$, $Q = Q_{\mathcal{C}, \mathcal{C}'}$, et soit $u \in \text{Hom}_K(E, F)$. Posons $M = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$, $M' = \text{Mat}_{\mathcal{B}', \mathcal{C}'}(u)$. Alors :

$$(10) \quad M' = Q^{-1}MP.$$

Démonstration :

Posons $M = [a_{ij}]$, $M' = [a'_{ij}]$, $P = [p_{ij}]$, $Q = [q_{ij}]$. Pour tout j , on a : $u(e'_j) = \sum_{k=1}^p a'_{kj} f'_k$, $u(e_j) = \sum_{k=1}^p a_{kj} f_k$, et $f'_k = \sum_{i=1}^p q_{ik} f_i$, $e'_j = \sum_{i=1}^n p_{ij} e_i$, d'où :

$$\sum_{k=1}^p a'_{kj} \left(\sum_{i=1}^p q_{ik} f_i \right) = \sum_{i=1}^n p_{ij} u(e_i) = \sum_{i=1}^n p_{ij} \left(\sum_{k=1}^p a_{ki} f_k \right),$$

soit

$$\sum_{i=1}^p \left(\sum_{k=1}^p q_{ik} a'_{kj} \right) f_i = \sum_{i=1}^n \left(\sum_{k=1}^p p_{kj} a_{ik} \right) f_i.$$

Donc, pour $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, on a : $\sum_{k=1}^p q_{ik} a'_{kj} = \sum_{k=1}^n a_{ik} p_{kj}$, d'où $QM' = MP$, d'où (10). ■

Exemple 7 : Supposons que $e'_i = e_{\sigma(i)}$ ($1 \leq i \leq n$) et $f'_j = f_{\tau(j)}$ ($1 \leq j \leq p$), avec $\sigma \in \mathfrak{S}_n$ et $\tau \in \mathfrak{S}_p$. Alors, avec les notations de l'exemple 6, $P = P_\sigma$, $Q = P_\tau$, d'où $Q^{-1} = P_{\tau^{-1}}$, et par suite : $M' = P_{\tau^{-1}} M P_\sigma$. Par ailleurs, avec les notations du § XI.1, il est clair que $M' = (\tau^{-1}, \sigma^{-1}) * M$, d'où finalement :

$$P_{\tau^{-1}} M P_\sigma = (\tau^{-1}, \sigma^{-1}) * M.$$

COROLLAIRE

Soit \mathcal{B} et \mathcal{B}' deux bases d'un même K -ev de dimension n , et soit $u \in \text{Hom}_K(E)$: si $M = \text{Mat}_{\mathcal{B}}(u)$, $M' = \text{Mat}_{\mathcal{B}'}(u)$, et $P = P_{\mathcal{B}, \mathcal{B}'}$, on a :

$$(11) \quad M' = P^{-1}MP \quad (\text{formule du changement de base}).$$

Application : trace d'un endomorphisme

THÉORÈME XI.3.5

Soit E un K -ev et $u \in \text{Hom}_K(E)$. Il existe un scalaire unique $\text{Tr}(u) \in K$ tel que, pour toute base \mathcal{B} de E , on ait : $\text{Tr}(u) = \text{Tr}(\text{Mat}_{\mathcal{B}}(u))$.
L'application $u \mapsto \text{Tr}(u)$ est une **forme linéaire** sur $\text{Hom}_K(E)$.

Démonstration :

E admettant au moins une base, $\text{Tr}(u)$ existe. L'unicité résulte de la formule (11) et du corollaire de la proposition XI.2.2. Fixons alors une base \mathcal{B} de E : l'application $\text{Hom}_K(E) \rightarrow \mathfrak{M}_n(K)$, $u \mapsto \text{Mat}_{\mathcal{B}}(u)$ est K -linéaire, et l'application $\mathfrak{M}_n(K) \rightarrow K$, $M \mapsto \text{Tr}(M)$ est K -linéaire, donc la composée des deux est K -linéaire, d'où la fin du théorème. ■

DÉFINITION XI.3.3

Avec les notations du théorème XI.3.5, pour $u \in \text{Hom}_K(E)$, le scalaire $\text{Tr}(u)$ s'appelle **trace de l'endomorphisme** u .

Exemple 8 : Soit u un projecteur de rang r dans E . Notons $F = \text{Im}(u)$ et $G = \text{Ker}(u)$. Dans une base $\mathcal{B} = (e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ de E telle que (e_1, \dots, e_r) soit une base de F et (e_{r+1}, \dots, e_n) une base de G , on a :

$$\text{Mat}_{\mathcal{B}}(u) = \text{Diag}(1_K, 1_K, \dots, 1_K, 0, \dots, 0), \quad \text{d'où} \quad \text{Tr}(u) = r \cdot 1_K.$$

En particulier, si K est de caractéristique 0, on a : $\text{Tr}(u) = r$.

Fonctions polynomiales sur un K -ev de dimension finie

Soit E un K -ev de dimension finie $n \geq 1$. A chaque base \mathcal{B} de E : $\mathcal{B} = (e_1, \dots, e_n)$, associons sa *base duale* $(\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}})$. Si \mathcal{B} et \mathcal{C} sont deux bases de E , et si $P = P_{\mathcal{B},\mathcal{C}} = [a_{ij}]$, la formule (9) signifie exactement

$$(12) \quad (\forall i \in \llbracket 1, n \rrbracket) \quad \varphi_{i,\mathcal{B}} = \sum_{j=1}^n a_{ij} \varphi_{j,\mathcal{C}}.$$

Soit $\mathcal{F}_{\mathcal{B}}$ la sous- K -algèbre de l'algèbre des fonctions $\mathcal{F}(E, K)$ définies sur E et à valeurs dans K , engendrée par $\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}}$. Ce qui précède montre que $\mathcal{F}_{\mathcal{B}} \subset \mathcal{F}_{\mathcal{C}}$, et en échangeant les rôles de \mathcal{B} et \mathcal{C} , on en déduit : $\mathcal{F}_{\mathcal{B}} = \mathcal{F}_{\mathcal{C}}$.

De plus, pour chaque base \mathcal{B} , les fonctions $\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}}$ sont **algébriquement libres sur K** , si le corps K est infini : en effet la démonstration du théorème X.1.2 peut être reproduite mot pour mot en remplaçant les projections canoniques $\varphi_1, \dots, \varphi_n$ de K^n dans K par $\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}}$. Ces considérations conduisent à la :

DÉFINITION XI.3.4

On appelle **algèbre des fonctions polynomiales** sur le K -ev E la sous- K -algèbre \mathcal{P} de $\mathcal{F}(E, K)$ telle que, pour toute base \mathcal{B} de E , on ait : $\mathcal{P} = \mathcal{F}_{\mathcal{B}} = K$ -algèbre engendrée par les formes coordonnées.

Et on peut donc énoncer :

THÉOREME XI.3.6

Soit \mathcal{P} la K -algèbre des fonctions polynomiales sur le K -ev E , le corps K étant supposé infini. Pour toute base \mathcal{B} de E , si $\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}}$ est la base duale, \mathcal{P} est l'algèbre de polynômes en les n lettres $\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}}$.

Les formules (12) montrent facilement que si P est polynomiale sur E , le degré et la valuation de P en les lettres $(\varphi_{1,\mathcal{B}}, \dots, \varphi_{n,\mathcal{B}})$ sont indépendantes du choix de \mathcal{B} . On les appelle degré et valuation de P .

Exercice 1 : Le corps de base est de caractéristique 0. Soit $\varphi_1, \varphi_2, \dots, \varphi_r$ des projecteurs d'un K -ev E de dimension finie $n \geq 1$ tels que $\varphi_1 + \varphi_2 + \dots + \varphi_r = \text{Id}_E$. Montrer que $E = \bigoplus_{i=1}^r \text{Im}(\varphi_i)$, et que les φ_i sont deux à deux permutables.

Exercice 2 : On donne n entier ≥ 2 . Le corps de base est commutatif quelconque. On ordonne $\llbracket 1, n \rrbracket^2 = I$ lexicographiquement : $I = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$, $N = n^2$. Soit \mathcal{B} la base $(\varepsilon_1, \dots, \varepsilon_N)$ du K -ev $\mathfrak{M}_n(K)$ obtenue avec cet ordre à partir de la base canonique $(E_{ij})_{(i,j) \in I}$ de $\mathfrak{M}_n(K)$.

a) Si $M = [a_{ij}]_{(i,j) \in I} \in \mathfrak{M}_n(K)$, trouver la matrice dans \mathcal{B} des applications linéaires de $\mathfrak{M}_n(K)$ dans $\mathfrak{M}_n(K) : X \mapsto MX ; X \mapsto XM$.

b) Si $M = [a_{ij}] \in \mathfrak{M}_n(K)$ et $N = [b_{ij}] \in \mathfrak{M}_n(K)$, trouver la matrice dans \mathcal{B} de l'application linéaire $X \mapsto MXN$.

Exercice 3 : Le corps de base est \mathbb{C} . Soit E un \mathbb{C} -ev de dimension $n \geq 2$. On considère une base $(\Phi_{ij})_{(i,j) \in \llbracket 1, n \rrbracket^2}$ de $\text{Hom}_{\mathbb{C}}(E)$ possédant les propriétés suivantes :

$$(I) \quad \sum_{i=1}^n \Phi_{ii} = \text{Id}_E \quad (II) \quad (\forall (i, j, k, l) \in \llbracket 1, n \rrbracket^4) \quad \Phi_{ij} \Phi_{kl} = \delta_{jk} \Phi_{il},$$

(δ = symbole de Kronecker).

Démontrer, en utilisant l'exercice 1, qu'il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que (Φ_{ij}) soit la base de $\text{Hom}_{\mathbb{C}}(E)$ canoniquement associée à \mathcal{B} (i.e. $\Phi_{ij}(e_k) = \delta_{kj} e_i$ pour tous i, j, k dans $\llbracket 1, n \rrbracket$).

Exercice 4 : Le corps de base est \mathbb{C} . On donne a et b distincts dans \mathbb{C} . Dans la base canonique $\mathcal{B} = (1, X, X^2, \dots, X^n)$ du \mathbb{C} -ev $\mathbb{C}_n[X]$ (n donné, $n \geq 1$) écrire la matrice de l'endomorphisme φ :

$$P(X) \mapsto (X - a)(X - b)P'(X) - 2n \left(X - \frac{a+b}{2} \right) P(X).$$

Exercice 5 : Soit K un corps commutatif de caractéristique 0 et E un K -ev de dimension finie $n \geq 1$. Soit $q \in \mathbb{N}^*$ et $u \in \text{Hom}_K(E)$ tel que $u^q = \text{Id}_E$. On pose $E_1 = \text{Ker}(u - \text{Id}_E)$. En utilisant $v = \frac{1}{q} \sum_{i=0}^{q-1} u^i$, démontrer : $\dim_K(E) = \frac{1}{q} \sum_{i=0}^{q-1} \text{Tr}(u^i)$. On prouvera d'abord que $E_1 = \text{Im}(v)$.

Exercice 6 : Soit E un K -ev de dimension finie $n \geq 1$. On considère une suite de sous- K -ev de E emboîtés strictement : $E_1 \subset E_2 \subset \dots \subset E_r = E$, avec $\dim_K(E_1) = n_1$, $\dim_K(E_2) = n_1 + n_2$, ..., $\dim_K(E_r) = n_1 + n_2 + \dots + n_r = n$.

a) Quelle est la forme de la matrice dans une base convenable d'un automorphisme u de E tel que $u(E_i) = E_i$ pour $i = 1, 2, \dots, r$? L'ensemble de ces automorphismes constitue un groupe. A quoi est-il isomorphe ?

b) Quelle est la forme de la matrice dans la même base d'un endomorphisme u de E tel que $u(E_i) \subset E_i$ pour tout $i \in \llbracket 1, r \rrbracket$. L'ensemble de ces endomorphismes constitue une K -algèbre. A quoi est-elle isomorphe ? (cf. exercice 6, § XI.2).

c) Cas particulier où $r = n$.

Exercice 7 : Soit K un corps commutatif et L une extension de K . On considère une matrice $M \in \mathfrak{M}_n(K)$ ($n \geq 1$). Montrer, en utilisant le corollaire 2 du théorème XI.3.3, que M est inversible dans $\mathfrak{M}_n(K)$ ssi elle est inversible dans $\mathfrak{M}_n(L)$, autrement dit qu'on a :

$$\mathrm{GL}(n, K) = \mathfrak{M}_n(K) \cap \mathrm{GL}(n, L).$$

Exercice 8 : Le corps de base est \mathbb{R} . On donne $n \in \mathbb{N}^*$, et on cherche les matrices $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$ inversibles et telles que, si $M^{-1} = [b_{ij}]$, on ait :

$$(\forall i, j) \quad a_{ij} \geq 0 \quad \text{et} \quad b_{ij} \geq 0.$$

a) Soit $\sigma \in \mathfrak{S}_n$ et c_1, \dots, c_n des réels > 0 . Montrer que $M = [c_j \delta_{\sigma^{-1}(i), j}]$ convient.

b) Soit M une matrice du type cherché. On lui associe $u \in \mathrm{GL}_{\mathbb{R}}(\mathbb{R}^n)$ de matrice M dans la base canonique $\mathscr{B} = (e_1, \dots, e_n)$. On note $\Omega = \mathbb{R}_+^n$. Montrer que $u(\Omega) = \Omega$. Soit $\mathscr{D}_k = \mathbb{R}e_k$; montrer que si $x \in \mathscr{D}_k \setminus \{0\}$, les relations $y \in \Omega$, $z \in \Omega$ et $x = \frac{1}{2}(y + z)$ entraînent $y \in \mathscr{D}_k$ et $z \in \mathscr{D}_k$, et que $\left(\bigcup_{k=1}^n \mathscr{D}_k\right) \setminus \{0\}$ est l'ensemble des $x \in \Omega$ possédant cette propriété. En déduire que u permute les \mathscr{D}_k , et que M est du type trouvé en a).

Exercice 9 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \mathrm{Hom}_K(E)$ ayant la même matrice dans toutes les bases de E . Montrer que u est une homothétie.

§ XI.4 RANG D'UNE MATRICE

Le corps commutatif de base K est fixé. Tous les K -ev considérés sont *non nuls* et de *dimension finie*.

Rappelons que si E et F sont deux K -ev, le **rang** de $u \in \mathrm{Hom}_K(E, F)$, noté $\mathrm{rg}(u)$, est l'entier $\dim_K(\mathrm{Im}(u))$. On a donc

$$\mathrm{rg}(u) \leq \min(\dim(E), \dim(F)).$$

THÉORÈME XI.4.1

|| Soit E et F deux K -ev, et u et v deux éléments de $\mathrm{Hom}_K(E, F)$. Pour que u et v aient même rang, il faut et il suffit qu'il existe $\alpha \in \mathrm{GL}_K(E)$ et $\beta \in \mathrm{GL}_K(F)$ tels que : $v = \beta \circ u \circ \alpha$.

Démonstration :

Puisqu'un automorphisme d'un K -ev conserve la dimension des sous- K -ev, il est clair que la condition indiquée est suffisante.

Montrons qu'elle est nécessaire. Soit donc $\mathrm{rg}(u) = \mathrm{rg}(v) = r \geq 1$. Notons $n = \dim_K(E)$, $p = \dim_K(F)$. Soit $N = \mathrm{Ker}(u)$, $I = \mathrm{Im}(u)$, $S =$ un supplémentaire de N dans E et $I' =$ un supplémentaire de I dans F . Soit de même $P = \mathrm{Ker}(v)$, $J = \mathrm{Im}(v)$, $T =$ un supplémentaire de P dans E , et $J' =$ un supplémentaire de J dans F . Nous savons que $u|_S^I: S \rightarrow I$ et $v|_T^{J'}: T \rightarrow J'$ sont des isomorphismes de K -ev (cf. théorème IX.1.4). Prenons une base (e_1, \dots, e_r) de S et une base (f_1, \dots, f_r) de T . So

éléments de $\text{Hom}_K(T, S)$ et $\text{Hom}_K(I, J)$ tels que $\varphi(f_i) = e_i$ et $\psi(u(e_i)) = v(f_i)$ pour $1 \leq i \leq r$, ζ (resp. η) un isomorphisme arbitraire de P sur N (resp. de I' sur J'), qui existent puisque $\dim(N) = \dim(P) = n - r$ et $\dim(I') = \dim(J') = p - r$. Alors $\alpha = \varphi \oplus \zeta$ et $\beta = \psi \oplus \eta$ sont respectivement dans $\text{GL}_K(E)$ et $\text{GL}_K(F)$, et on vérifie que : $v = \beta \circ u \circ \alpha$. ■

DÉFINITION XI.4.1

Soit $M \in \mathfrak{M}_{p,n}(K)$ ($p \geq 1, n \geq 1$). On appelle **rang** de M , et on note $\text{rg}(M)$, le rang de l'application linéaire $u_M \in \text{Hom}_K(K^n, K^p)$ dont la matrice dans les bases canoniques est M .

Soit $(e_1, \dots, e_n) = \mathcal{B}$ et $(f_1, \dots, f_p) = \mathcal{C}$ les bases canoniques de K^n et K^p . Les vecteurs $(u_M(e_j))_{1 \leq j \leq n}$ ne sont autres que les vecteurs-colonnes $(\mathcal{C}_j(M))_{1 \leq j \leq n}$ de M . Donc (cf. § IX.4), le rang de M n'est autre que le rang des vecteurs-colonnes de M . Le rang de la matrice nulle est bien sûr 0, et on a : $\text{rg}(M) \leq \min(n, p)$.

PROPOSITION XI.4.1

Le rang d'une matrice $M \in \mathfrak{M}_{p,n}(K)$ est égal au rang de toute application K -linéaire que M peut représenter.

Démonstration :

Conservons les notations ci-dessus, soit $\mathcal{B}' = (e'_1, \dots, e'_n)$ une base de E et $\mathcal{C}' = (f'_1, \dots, f'_p)$ une base de F . Notons v l'élément de $\text{Hom}_K(E, F)$ tel que $\text{Mat}_{\mathcal{B}', \mathcal{C}'}(v) = M$. Alors, on a des isomorphismes de K -ev : $\alpha : K^n \rightarrow E$ et $\beta : K^p \rightarrow F$ tels que $\alpha(e_i) = e'_i$ ($1 \leq i \leq n$) et $\beta(f_j) = f'_j$ ($1 \leq j \leq p$). On vérifie que $v = \beta \circ u_M \circ \alpha^{-1}$, d'où, puisque α et β sont des isomorphismes, $\text{rg}(v) = \text{rg}(u_M) = \text{rg}(M)$. ■

THÉORÈME XI.4.2

Soit $M \in \mathfrak{M}_{p,n}(K)$ ($p \geq 1, n \geq 1$) et $N \in \mathfrak{M}_{p,n}(K)$.
Pour qu'on ait $\text{rg}(M) = \text{rg}(N)$ il faut et il suffit qu'il existe $P \in \text{GL}(n, K)$ et $Q \in \text{GL}(p, K)$ telles que :
$$QMP = N.$$

Démonstration :

Soit \mathcal{B} et \mathcal{C} les bases canoniques des K -ev K^n et K^p . Notons u_M et u_N les éléments de $\text{Hom}_K(K^n, K^p)$ tels que $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u_M) = M$, et $\text{Mat}_{\mathcal{B}, \mathcal{C}}(u_N) = N$. De même, si $P \in \text{GL}(n, K)$ (resp. $Q \in \text{GL}(p, K)$), soit α_P (resp. β_Q) l'élément de $\text{GL}_K(K^n)$ (resp. $\text{GL}_K(K^p)$) tel que $\text{Mat}_{\mathcal{B}}(\alpha_P) = P$ (resp. $\text{Mat}_{\mathcal{C}}(\beta_Q) = Q$). On sait que $P \mapsto \alpha_P, \text{GL}(n, K) \rightarrow \text{GL}_K(K^n)$ et $Q \mapsto \beta_Q, \text{GL}(p, K) \rightarrow \text{GL}_K(K^p)$ sont bijectives. D'autre part, les relations $u_N = \beta_Q \circ u_M \circ \alpha_P^{-1}$

QMP sont équivalentes. Le théorème est donc conséquence du théorème XI.4.1. ■

En tenant compte de l'exemple 7 du § XI.3, on déduit l'importante conséquence de ce théorème XI.4.2 : pour toutes permutations $\sigma \in \mathfrak{S}_p$ et $\tau \in \mathfrak{S}_n$, on a :

$$(1) \quad \text{rg}(M) = \text{rg}((\sigma, \tau) * M).$$

Matrices carrées de rang maximum

Soit $n \in \mathbb{N}^*$, et $M = [a_{ij}] \in \mathfrak{M}_n(K)$. On note, comme ci-dessus, u_M l'élément de $\text{Hom}_K(K^n)$ tel que $\text{Mat}_{\text{Can}}(u_M) = M$. D'après ce qui précède, M est de rang n ssi ses vecteurs-colonnes $(\mathcal{C}_j(M))_{1 \leq j \leq n}$ sont linéairement indépendants. Pour cela, il faut et il suffit que u_M soit bijectif (cf. théorème IX.6.1), donc que M soit inversible.

Nous avons vu, au § XI.2, que $M \in \text{GL}(n, K)$ ssi ${}^tM \in \text{GL}(n, K)$. Nous obtenons donc :

PROPOSITION XI.4.2

|| Une matrice carrée $M \in \mathfrak{M}_n(K)$ est de rang n ssi sa transposée l'est aussi.

Matrices bordantes

Soit $M = [a_{ij}] \in \mathfrak{M}_{p,n}(K)$. Si $P = \mathcal{M}_{I,J}(M)$ et $P' = \mathcal{M}_{I',J'}(M)$ sont deux sous-matrices de M , nous dirons que P est incluse dans P' ssi $I \subset I'$ et $J \subset J'$, ce que nous écrirons $P \subseteq P'$. Cela entraîne que P est une sous-matrice de P' . On obtient ainsi une relation d'ordre partiel sur l'ensemble des sous-matrices de M .

Si $P \subseteq P'$, on dit que P' borde P , ou que P' est une **matrice bordante** de P , ssi $\text{card}(I') = 1 + \text{card}(I)$ et $\text{card}(J') = 1 + \text{card}(J)$.

Si P est carrée, toute matrice bordante de P l'est aussi.

THÉORÈME XI.4.3

|| Soit $M \in \mathfrak{M}_{p,n}(K)$ ($p \geq 1, n \geq 1$), $M \neq 0$. Supposons trouvée une sous-matrice **carrée** P de M , **inversible** d'ordre $r \geq 1$ telle que P n'admette aucune matrice bordante inversible dans M . Alors $\text{rg}(M) = r$.

Démonstration :

Notons φ l'application linéaire

$$K^p \rightarrow K^r, \quad (x_1, \dots, x_p) \mapsto (x_1, \dots, x_r).$$

En tenant compte de (1), nous pouvons toujours nous ramener au cas où $P = \mathcal{M}_{[1,r],[1,r]}(M)$, ce que nous supposons désormais

Les vecteurs-colonnes $(\mathcal{C}_j(P))_{1 \leq j \leq r}$ sont les images par φ des vecteurs-colonnes $(\mathcal{C}_j(M))_{1 \leq j \leq r}$ de M . Or puisque P est inversible, les $(\mathcal{C}_j(P))$ sont linéairement indépendants, donc *a fortiori* les colonnes $(\mathcal{C}_j(M))_{1 \leq j \leq r}$ le sont, d'où $\text{rg}(M) \geq r$.

Si $r = n$ ou si $r = p$, on est sûr que $\text{rg}(M) = r$, car de toute façon, $\text{rg}(M) \leq \min(n, p)$. D'ailleurs dans ce cas on ne peut plus border la matrice P .

Supposons donc $r < n$ et $r < p$. Nous allons montrer que toute colonne $\mathcal{C}_\mu(M)$, $\mu > r$, appartient à $\text{Vect}(\mathcal{C}_1(M), \dots, \mathcal{C}_r(M))$. Pour cela, fixons $\mu \in \llbracket r+1, n \rrbracket$, et soit $\lambda \in \llbracket r+1, p \rrbracket$. Posons $I' = \llbracket 1, r \rrbracket \cup \{\lambda\}$, $J' = \llbracket 1, r \rrbracket \cup \{\mu\}$, $P'_\lambda = \mathcal{M}_{I', J'}(M)$ et $P_0 = \mathcal{M}_{\llbracket 1, r \rrbracket, J'}(M)$. Par hypothèse, P'_λ est non inversible, et on a donc une relation linéaire

$$\alpha_1 \mathcal{C}_1(P'_\lambda) + \dots + \alpha_r \mathcal{C}_r(P'_\lambda) + \alpha_{r+1} \mathcal{C}_{r+1}(P'_\lambda) = 0,$$

avec des α_i non tous nuls, et en particulier $\alpha_{r+1} \neq 0$, sans quoi les r premières colonnes de P'_λ seraient liées, et donc aussi les r premières colonnes de P , contrairement à l'hypothèse que P est inversible. On en déduit, en multipliant par le scalaire α_{r+1}^{-1} :

$$(2) \quad \mathcal{C}_{r+1}(P'_\lambda) = \beta_1 \mathcal{C}_1(P'_\lambda) + \dots + \beta_r \mathcal{C}_r(P'_\lambda),$$

$$\text{d'où } a \text{ fortiori : } \mathcal{C}_{r+1}(P_0) = \beta_1 \mathcal{C}_1(P) + \dots + \beta_r \mathcal{C}_r(P).$$

Donc $(\beta_1, \dots, \beta_r)$ sont les coordonnées de la colonne $\mathcal{C}_{r+1}(P_0)$ sur la base $(\mathcal{C}_1(P), \dots, \mathcal{C}_r(P))$ de K' , et à ce titre, sont indépendants du choix de λ . Par suite, les relations (2) sont vraies avec ce système $(\beta_1, \dots, \beta_r) \in K'$, quel que soit $\lambda \in \llbracket r+1, p \rrbracket$. Cela signifie que :

$$\mathcal{C}_\mu(M) = \beta_1 \mathcal{C}_1(M) + \dots + \beta_r \mathcal{C}_r(M).$$

Cela étant vrai pour tout $\mu \in \llbracket r+1, n \rrbracket$, on voit que le rang des vecteurs-colonnes de M est $\leq r$, d'où $\text{rg}(M) \leq r$ et finalement $\text{rg}(M) = r$. ■

COROLLAIRE 1

|| Le rang d'une matrice $M \in \mathfrak{M}_{p,n}(K)$ est le maximum des ordres des sous-matrices carrées inversibles de M .

COROLLAIRE 2

|| Le rang d'une matrice $M \in \mathfrak{M}_{p,n}(K)$ est égal à celui de sa transposée.

Ce corollaire 2 se déduit immédiatement de la proposition XI.4.2.

Equivalence des matrices

Deux matrices M, N , éléments de $\mathfrak{M}_{p,n}(K)$ ($p \geq 1, n \geq 1$) sont dites équivalentes ssi elles ont même rang.

La relation binaire ainsi introduite sur $\mathfrak{M}_{p,n}(K)$ est évidemment une relation d'équivalence qui, d'après le théorème XI.4.2, peut encore s'exprimer de la façon suivante : deux (p, n) -matrices sont équivalentes ssi il existe $P \in GL(n, K)$ et $Q \in GL(p, K)$ telles que $N = QMP$.

Le nombre des classes d'équivalence est fini, puisque le rang de $M \in \mathfrak{M}_{p,n}(K)$ est $\leq \rho = \min(p, n)$. Si \mathcal{E} est l'une des classes, toutes les matrices $M \in \mathcal{E}$ ont même rang $r = r(\mathcal{E})$, et \mathcal{E} est formée de toutes les (p, n) -matrices de rang r . L'ensemble $\{0\}$ réduit à la matrice nulle de $\mathfrak{M}_{p,n}(K)$ est évidemment une classe. Pour les autres classes, on a $1 \leq r(\mathcal{E}) \leq \rho$. Pour chaque $r \in \llbracket 1, \rho \rrbracket$, notons L_r la matrice $[a_{ij}] \in \mathfrak{M}_{p,n}(K)$ telle que $a_{i,i} = 1$ si $i \in \llbracket 1, r \rrbracket$ et $a_{i,j} = 0$ pour tous les autres couples (i, j) . Il est clair que L_r est de rang r , par application du théorème XI.4.3. Donc il existe des matrices de tout rang r . En résumé :

PROPOSITION XI.4.3

Soit $p \geq 1, n \geq 1$ et $\rho = \min(p, n)$. Il y a exactement $\rho + 1$ classes de matrices équivalentes dans $\mathfrak{M}_{p,n}(K)$: la classe de 0, et les ρ classes \mathcal{E}_r ($1 \leq r \leq \rho$), où \mathcal{E}_r est l'ensemble des matrices de rang r , i.e. équivalentes à L_r .

Exercice 1 : a) Soit $M \in \mathfrak{M}_{p,n}(K)$, de rang r . Quel est le rang des applications linéaires

$$\mathfrak{M}_{n,m}(K) \rightarrow \mathfrak{M}_{p,n}(K), X \mapsto MX$$

et

$$\mathfrak{M}_{q,p}(K) \rightarrow \mathfrak{M}_{q,n}(K), Y \mapsto YM ?$$

(m et q donnés dans \mathbb{N}^*).

b) On fixe $n \in \mathbb{N}^*$, et deux matrices A et B dans $\mathfrak{M}_n(K)$, de rangs respectifs r et s . Quel est le rang de l'application linéaire : $\mathfrak{M}_n(K) \rightarrow \mathfrak{M}_n(K), M \mapsto AMB$?

Exercice 2 : a) Soit $n \in \mathbb{N}^*$. Montrer qu'une matrice $M = [a_{ij}] \in \mathfrak{M}_n(K)$ est de rang 1 ssi on a $(b_1, \dots, b_n) \in K^n$ et $(a_1, \dots, a_n) \in K^n$ non nuls tels que $(\forall i, j) \quad a_{ij} = b_i c_j$.

b) $M = [a_{ij}] \in \mathfrak{M}_n(K)$ est de rang 1 et symétrique ssi on a $\rho \in K^*$ et $(b_1, \dots, b_n) \in K^n \setminus \{0\}$ tels que $(\forall i, j) \quad a_{ij} = \rho b_i b_j$.

Application : $M = [\alpha^{i-j}]$, $\alpha \neq 0$.

c) Si $K = \mathbb{C}$, les matrices de rang 1 et symétriques dans $\mathfrak{M}_n(\mathbb{C})$ sont les $M(a_1, \dots, a_n) = [a_i a_j]$, où $(a_1, \dots, a_n) \in \mathbb{C}^n \setminus \{0\}$. Etudier les $(a_1, \dots, a_n) \in \mathbb{C}^n \setminus \{0\}$ tels que $M(a_1, \dots, a_n)$ soit une matrice symétrique de rang 1 donnée.

Exercice 3 : Soit K un corps commutatif et L une extension de K . On donne deux entiers $p \geq 1, n \geq 1$ et $M \in \mathfrak{M}_{p,n}(K)$. Montrer que le rang de M considérée comme élément de $\mathfrak{M}_{p,n}(K)$, et le rang de M considérée comme élément de $\mathfrak{M}_{p,n}(L)$, sont égaux (utiliser le corollaire 1 du théorème XI.4.2 et l'exercice 7, § XI.3).

Exercice 4 : Etudier le rang d'une matrice carrée diagonale par blocs, en fonction des rangs des divers blocs.

Exercice 5 : a) Soit $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$ ($n \geq 2$) telle que

$$(\forall i \in \llbracket 1, n \rrbracket) \quad |a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Prouver que M est inversible (*indication* : on pourra raisonner par l'absurde)

b) Soit $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$. On suppose :

$$(I) \quad \forall i \in \llbracket 1, n \rrbracket \quad |a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad (II) \quad \exists i \in \llbracket 1, n \rrbracket, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

$$(III) \quad \forall (i, j) \in \llbracket 1, n \rrbracket^2, \exists m \in \mathbb{N}^*, \exists (k_1, \dots, k_m) \in \llbracket 1, n \rrbracket^m, a_{ik_1} a_{k_1 k_2} \dots a_{k_m j} \neq 0.$$

Montrer que M est inversible (Hadamard).

Exercice 6 : On considère la matrice $C = [A : B]$ obtenue en juxtaposant une (m, n_1) -matrice A et une (m, n_2) -matrice B . Montrer que $\text{rg}(C) \leq \text{rg}(A) + \text{rg}(B)$.

Exercice 7 : Soit $A \in \mathfrak{M}_n(K)$, $B \in \mathfrak{M}_n(K)$ de rangs respectifs r_1 et r_2 . Montrer que le rang r de AB vérifie : $r_2 \geq r \geq r_1 + r_2 - n$ (inégalité de Sylvester).

Exercice 8 : Trouver toutes les matrices A de $\mathfrak{M}_3(\mathbb{C})$ telles que $A^2 = 0$.

Indication : si $A \neq 0$, $\text{rg}(A) = 1$ et $\text{Tr}(A) = 0$ (on pourra utiliser les exercices 2 et 7).

Exercice 9 : Rang de la matrice $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{R})$ telle que $a_{ij} = i + j - 1$.

Exercice 10 : Rang de la matrice $M = [a_{ij}] \in \mathfrak{M}_{p,n}(\mathbb{R})$ de terme général $a_{ij} = \cos(i + j - 1)\alpha$, où le réel α est donné.

§ XI.5 OPÉRATIONS ÉLÉMENTAIRES

Le corps commutatif de base K est fixé.

Etant donnée une matrice $M \in \mathfrak{M}_{p,n}(K)$ ($p \geq 1, n \geq 1$), on appelle **opération élémentaire** sur M toute opération de l'un des types suivants :

1° *addition à une colonne* (resp. *ligne*) de M , du produit par un scalaire d'une *autre* colonne (resp. *ligne*) de M ;

2° *multiplication d'une colonne* (resp. *ligne*) de M par un scalaire non nul ;

3° *permutation* sur les colonnes (resp. sur les lignes) de M .

Il est facile de vérifier directement que *toute opération élémentaire conserve le rang* d'une matrice (et cela sera confirmé, compte tenu des résultats du § XI.4, par la proposition XI.5.1 ci-dessous).

Nous avons déjà introduit les *matrices de permutation* au § XI.3. Afin de décrire complètement les opérations élémentaires, introduisons deux autres types de matrices carrées :

Si $m \in \mathbb{N}$, $m \geq 2$, notons $(E_{i,j})_{(i,j) \in \llbracket 1, m \rrbracket^2}$ la base canonique du K -ev $\mathfrak{M}_m(K)$ (formée des matrices élémentaires dont nous rappelons que la *table de multiplication* est :

$$(1) \quad (\forall (i, j, k, l) \in \llbracket 1, m \rrbracket^4) \quad E_{ij} E_{kl} = \delta_{jk} E_{il}.$$

Cela dit, soit i et j dans $\llbracket 1, m \rrbracket$, $i \neq j$. Posons, pour $\lambda \in K$:

$$T_{ij}(\lambda) = I_m + \lambda E_{ij}.$$

Lorsque $\lambda \neq 0$, $T_{ij}(\lambda)$ s'appelle une **matrice de transvection**.

Enfin, soit $i \in \llbracket 1, m \rrbracket$. Posons, pour $\lambda \in K^*$:

$$D_i(\lambda) = \lambda E_{ii} + \sum_{j \neq i} E_{jj} = \text{Diag}(1, 1, \dots, 1, \lambda, 1, \dots, 1),$$

la lettre λ figurant au rang i ; $D_i(\lambda)$ s'appelle une **matrice de dilatation** : une telle matrice est évidemment inversible. Il convient d'observer que toute matrice de transvection, en tant que matrice *unipotente*, est également *inversible*. Si (i, j) est fixé ($i \neq j$), on a en utilisant (1) :

$$(2) \quad \forall (\lambda, \mu) \in K^2 \quad T_{ij}(\lambda) T_{ij}(\mu) = I_n + (\lambda + \mu) E_{ij} = T_{ij}(\lambda + \mu).$$

Par suite on a un *homomorphisme de groupes*

$$(K, +) \rightarrow \text{GL}(m, K), \lambda \mapsto T_{ij}(\lambda).$$

De même, si $i \in \llbracket 1, m \rrbracket$ est fixé, on a un homomorphisme de groupes :

$$K^* \rightarrow \text{GL}(m, K), \lambda \mapsto D_i(\lambda).$$

A cause de (2), on a : $[T_{ij}(\lambda)]^{-1} = T_{ij}(-\lambda)$ pour tout $\lambda \in K$.

Il est évident d'autre part que $[D_i(\lambda)]^{-1} = D_i(\lambda^{-1})$. Si l'on se souvient que $(P_\sigma)^{-1} = P_{\sigma^{-1}}$ pour toute matrice de permutation, on voit que l'inverse de toutes ces matrices particulières s'obtient sans calculs.

En tenant compte de l'exemple 7, § XI.3, il suffit pour démontrer la proposition ci-dessous de quelques calculs simples de vérification. Elle permet de ramener les opérations élémentaires à des calculs matriciels.

PROPOSITION XI.5.1

Soit n et p deux entiers ≥ 2 , et soit $M \in \mathfrak{M}_{p,n}(K)$.

(I) Soit $(\sigma, \tau) \in \mathfrak{S}_p \times \mathfrak{S}_n$ et soit P_σ et P_τ les matrices de permutation associées. Alors : $(\sigma, \tau) * M = P_\sigma M P_\tau^{-1}$.

(II) Soit $T_{ij}(\lambda) \in \mathfrak{M}_n(K)$ ($i \neq j, \lambda \in K$). La matrice $MT_{ij}(\lambda)$ s'obtient en remplaçant, dans M , la colonne $\mathcal{C}_j(M)$ par $\mathcal{C}_j(M) + \lambda \mathcal{C}_i(M)$.

(III) Soit $T_{ij}(\lambda) \in \mathfrak{M}_p(K)$ ($i \neq j, \lambda \in K$). La matrice $T_{ij}(\lambda) M$ s'obtient en remplaçant, dans M , la ligne $\mathcal{L}_i(M)$ par $\mathcal{L}_i(M) + \lambda \mathcal{L}_j(M)$.

(IV) Soit $D_i(\lambda) \in \mathfrak{M}_n(K)$ ($1 \leq i \leq n, \lambda \in K^*$). La matrice $MD_i(\lambda)$ s'obtient en remplaçant, dans M , $\mathcal{C}_i(M)$ par $\lambda \mathcal{C}_i(M)$.

(V) Soit $D_i(\lambda) \in \mathfrak{M}_p(K)$ ($1 \leq i \leq p, \lambda \in K^*$). La matrice $D_i(\lambda) M$ s'obtient en remplaçant, dans M , $\mathcal{L}_i(M)$ par $\lambda \mathcal{L}_i(M)$.

Les opérations élémentaires permettent de passer d'une matrice donnée $M \in \mathfrak{M}_{p,n}(K)$ à n'importe quelle autre matrice de même rang. Compte tenu du théorème XI.4.2 et de la proposition XI.5.1 il suffit pour le voir, de démontrer :

THÉORÈME XI.5.1

|| Soit m un entier ≥ 1 . La réunion de l'ensemble des matrices de dilatation, et de l'ensemble (qui est vide pour $m = 1$) des matrices de transvection engendre le groupe $GL(m, K)$.

Démonstration :

Montrons, par récurrence sur m , que toute matrice $M \in GL(m, K)$ peut s'écrire sous la forme (3) $M = DT_1 T_2 \dots T_r$, où D est une matrice de dilatation du type $D_n(\lambda)$, ($\lambda \in K^*$) et où les T_i sont des matrices de transvection. Pour $m = 1$, les $D_1(\lambda)$, ($\lambda \in K^*$) suffisent à engendrer le groupe $GL(1, K)$ qui est isomorphe à K^* .

Supposons la propriété vraie à l'ordre $m - 1$, avec $m \geq 2$, et soit $M \in GL(m, K)$. Notons \mathcal{E} l'ensemble des matrices du type $MT_1 T_2 \dots T_r$, où les T_i sont des matrices de transvection en nombre fini. Puisque l'inverse d'une matrice de transvection en est une autre, il s'agit de prouver seulement que \mathcal{E} contient une matrice du type $D_n(\lambda)$. Posons $M = [a_{ij}]$. Les lignes de M sont indépendantes, donc $\mathcal{L}_1(M) \neq 0$. Soit $j \in \llbracket 1, m \rrbracket$ tel que $a_{1j} \neq 0$. La matrice $N = [b_{ij}]$ obtenue en remplaçant, dans M , $\mathcal{C}_1(M)$ par $\mathcal{C}_1(M) + \frac{1 - a_{11}}{a_{1j}} \mathcal{C}_j(M)$, appartient à \mathcal{E} d'après la proposi-

tion XI.5.1, et vérifie $b_{11} = 1_K$. De même, la matrice $P = [c_{ij}]$ obtenue en remplaçant, dans N , $\mathcal{C}_j(N)$ par $\mathcal{C}_j(N) - b_{1j} \mathcal{C}_1(N)$ pour $2 \leq j \leq m$ appartient à \mathcal{E} , et on a $c_{12} = \dots = c_{1m} = 0$.

Soit Q la sous-matrice $\mathcal{M}_{\llbracket 2, m \rrbracket, \llbracket 2, m \rrbracket}(P)$ de P : elle est carrée d'ordre $m - 1$ et ses colonnes sont linéairement indépendantes (car celles de P le sont), donc forment une base de K^{m-1} . On a donc $\alpha_2, \dots, \alpha_m$ dans K tels que $\alpha_2 \mathcal{C}_1(Q) + \dots + \alpha_m \mathcal{C}_{m-1}(Q)$ soit le vecteur (c_{21}, \dots, c_{m1}) . En remplaçant, dans P , la colonne $\mathcal{C}_1(P)$ par $\mathcal{C}_1(P) - \sum_{k=2}^m \alpha_k \mathcal{C}_k(P)$, on voit

que la matrice diagonale par blocs $R = \text{Diag}([1], Q)$ appartient à \mathcal{E} , et de plus on a vu que $Q \in GL(m - 1, K)$. Appliquons l'hypothèse de récurrence à Q , d'où $Q = \Delta U_1 \dots U_s$, avec

$$\Delta = \text{Diag}(1, \dots, 1, \lambda) \in GL(m - 1, K) \quad (\lambda \in K^*)$$

et des matrices de transvection $U_i \in GL(m - 1, K)$. Posons $D = \text{Diag}([1], \Delta)$ et $T_i = \text{Diag}([1], U_i)$, d'où $D = D_n(\lambda) \in GL(m, K)$ et les T_i sont des matrices de transvection dans $GL(m, K)$. Alors $R = DT_1 \dots T_s$, d'où $D = RT_s^{-1} \dots T_1^{-1} \in \mathcal{E}$ (car $R \in \mathcal{E}$ et les matrices T_i^{-1} sont de transvection). ■

Remarque 1 : En raisonnant sur les lignes, on prouverait de même que toute matrice $M \in GL(m, K)$ peut s'écrire sous la forme

$$(4) \quad M = U_1 U_2 \dots U_s D,$$

où les U_i sont des matrices de transvection, et où $D = D_m(\mu)$ ($\mu \in K^*$).

Remarque 2 : Nous verrons au prochain chapitre qu'en fait les valeurs de λ et de μ figurant dans (3) et (4) sont égales. Elles ne dépendent que de M et non de la décomposition choisie. Toutes deux valent $\det(M)$.

Exemple 1 : Prenons $K = \mathbb{C}$ et considérons la (4, 5)-matrice M que nous désirons ramener, par opérations élémentaires, à la forme réduite L_r :

$$M = \begin{bmatrix} 1 & -4 & -3 & -2 & 2 \\ 2 & -6 & -6 & -4 & 4 \\ -3 & 12 & 12 & 6 & -9 \\ 0 & 2 & 3 & 0 & -3 \end{bmatrix}.$$

1^{re} opération élémentaire : $\mathcal{C}_5 \rightsquigarrow \mathcal{C}_5 + \mathcal{C}_2$:

$$M \rightarrow \begin{bmatrix} 1 & -4 & -3 & -2 & -2 \\ 2 & -6 & -6 & -4 & -2 \\ -3 & 12 & 12 & +6 & 3 \\ 0 & 2 & 3 & 0 & -1 \end{bmatrix}$$

suivie d'une 2^e opération élémentaire qui fait apparaître une ligne de zéros : $\mathcal{L}_4 \rightsquigarrow \mathcal{L}_4 - (\mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3)$

$$\begin{aligned} &\rightarrow \begin{bmatrix} 1 & -4 & -3 & -2 & -2 \\ 2 & -6 & -6 & -4 & -2 \\ -3 & 12 & 12 & +6 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\mathcal{L}_3 \rightsquigarrow \mathcal{L}_3 + 3\mathcal{L}_1} \begin{bmatrix} 1 & -4 & -3 & -2 & -2 \\ 2 & -6 & -6 & -4 & -2 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &\xrightarrow{\mathcal{L}_2 \rightsquigarrow \mathcal{L}_2 - 2\mathcal{L}_1} \begin{bmatrix} 1 & -4 & -3 & -2 & -2 \\ 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &\xrightarrow{\mathcal{C}_5 \rightsquigarrow \mathcal{C}_5 - \mathcal{C}_2 - \mathcal{C}_3} \begin{bmatrix} 1 & -4 & -3 & -2 & -3 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Les opérations simples :

$$\mathcal{C}_2 \rightsquigarrow \mathcal{C}_2 + 4\mathcal{C}_1, \mathcal{C}_3 \rightsquigarrow \mathcal{C}_3 + 3\mathcal{C}_1, \mathcal{C}_4 \rightsquigarrow \mathcal{C}_4 + 2\mathcal{C}_1 \text{ et } \mathcal{C}_5 \rightsquigarrow \mathcal{C}_5 + 3\mathcal{C}_1$$

ramènent finalement la matrice à la forme $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, et par

dilatations on arrive à $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = L_3$. Il est clair que la matrice M

est de rang 3 et l'on trouvera sans difficulté une sous-matrice carrée de M d'ordre 3 inversible.

THÉORÈME XI.5.2

Soit $m \in \mathbb{N}$, $m \geq 2$. Les matrices de transvection $T_{ij}(\lambda)$ ($1 \leq i < j \leq m$, $\lambda \in K^*$) engendrent le groupe $\mathcal{U}_+(m, K)$ des matrices unipotentes supérieures.

Démonstration :

Soit I l'ensemble $\{(i, j) \in \llbracket 1, m \rrbracket^2 \mid i < j\}$. Pour $a = (a_{ij})_{(i,j) \in I}$ élément de K^I , définissons la matrice $\Phi(a) \in \mathcal{U}_+(m, K)$ de la manière suivante :

$$(5) \quad \Phi(a) = M_{m-1} M_{m-2} \dots M_1,$$

avec, pour $1 \leq i \leq m-1$: $M_i = T_{im}(a_{im}) T_{i, m-1}(a_{i, m-1}) \dots T_{i, i+1}(a_{i, i+1})$. Un calcul de proche en proche, en partant de M_1 , puis $M_1 M_2$, etc... prouve que : $\Phi(a) = I_m + \sum_{(i,j) \in I} a_{ij} E_{ij}$. Par suite, Φ est une bijection de K^I sur $\mathcal{U}_+(m, K)$, ce qui entraîne évidemment le théorème. ■

Remarque 3 : L'expression (5) ci-dessus donne un calcul de l'inverse de $\Phi(a)$:

$$[\Phi(a)]^{-1} = M_1^{-1} M_2^{-1} \dots M_{m-1}^{-1}$$

avec $M_i^{-1} = T_{i, i+1}(-a_{i, i+1}) \dots T_{i, m}(-a_{i, m})$.

Exercice 1 : Le corps de base est \mathbb{C} . Déterminer le rang des matrices suivantes :

$$\begin{aligned} \text{a) } M &= \begin{bmatrix} 1 & 2 & -4 & -2 & -1 \\ 0 & -2 & 4 & 2 & 0 \\ 1 & 1 & -2 & -1 & 1 \end{bmatrix}; \quad \text{b) } M = \begin{bmatrix} 0 & -1 & 2 & -2 \\ -7 & -7 & 2 & -8 \\ 0 & 4 & -6 & 6 \\ +2 & -2 & 0 & -2 \end{bmatrix}; \\ \text{c) } M &= \begin{bmatrix} 1 & 7 & 2 & 5 \\ -2 & 1 & 1 & 5 \\ -1 & 2 & 1 & 4 \\ 1 & 4 & 1 & 2 \end{bmatrix}; \quad \text{d) } M = \begin{bmatrix} 1 & 4 & -1 & 2 & 4 \\ 2 & 0 & -3 & -1 & 7 \\ -2 & 3 & 2 & 1 & 4 \end{bmatrix}. \end{aligned}$$

Exercice 2 : Etendre la proposition XI.5.1 (II), (III), (IV), (V), avec des définitions adéquates, à des matrices par blocs.

Exercice 3 : Soit K un corps commutatif, $n \in \mathbb{N}^*$ et \mathcal{H} un hyperplan du K -ev $\mathfrak{M}_n(K)$. Montrer que \mathcal{H} rencontre $\text{GL}(n, K)$.

Indication : On pourra utiliser l'exercice 2a) du XI.2 et penser à une opération élémentaire.

Exercice 4 : K est un corps commutatif quelconque. Soit r, p et n des entiers vérifiant $p > r \geq 1$, $n > r$. On considère l'ensemble \mathcal{E}_r des $M \in \mathfrak{M}_{p, n}(K)$ telles que la matrice $M_r = \mathcal{M}_{\llbracket 1, r \rrbracket, \llbracket 1, r \rrbracket}(M)$ soit inversible. Construire une bijection de $\text{GL}(r, K) \times K^{r(p+n-2r)}$ sur \mathcal{E}_r .

Indication : Pour construire $M \in \mathcal{E}_r$, choisir M_r ; puis compléter les colonnes $\mathcal{C}_1(M), \dots, \mathcal{C}_r(M)$; enfin définir les $\mathcal{C}_j(M)$ ($j > r$) comme combinaisons linéaires arbitraires des $\mathcal{C}_i(r)$, $1 \leq i \leq r$.

Exercice 5 : Le corps commutatif de base K est supposé *fini*, de cardinal q . On donne trois entiers r, p, n tels que $1 \leq r \leq \min(p, n)$. Calculer le nombre de matrices de rang r dans $\mathfrak{M}_{p,n}(K)$.

Indication : Chercher le nombre d'applications K -linéaires $K^n \rightarrow K^p$ qui sont de rang r , en s'aidant de l'exercice 1 du § IX.6.

§ XI.6 SIMILITUDE D'ENDOMORPHISMES OU DE MATRICES

Le corps commutatif de base K est fixé. Les K -ev considérés sont de *dimension finie*.

Endomorphismes semblables

DÉFINITION XI.6.1

Soit E un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$, $v \in \text{Hom}_K(E)$. On dit que u et v sont **semblables** ssi il existe $\alpha \in \text{GL}_K(E)$ tel que

$$(1) \quad v = \alpha u \alpha^{-1}.$$

Il est clair que, sur $\text{Hom}_K(E)$, la relation binaire « u et v sont semblables » est **d'équivalence**. Les classes d'équivalence sont appelées les **classes de similitude** de $\text{Hom}_K(E)$.

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $\alpha \in \text{GL}_K(E)$; si $M = [a_{ij}]$ est la matrice de $u \in \text{Hom}_K(E)$ dans \mathcal{B} , c'est aussi la matrice de $\alpha u \alpha^{-1}$ dans la base $(\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n))$. On en déduit que *deux éléments u et v de $\text{Hom}_K(E)$ sont semblables ssi il existe des bases \mathcal{B} et \mathcal{C} de E telles que :* $\text{Mat}_{\mathcal{B}}(u) = \text{Mat}_{\mathcal{C}}(v)$.

D'après (1) on voit que, si u et v sont semblables, ils sont inversibles ou non en même temps. Par conséquent, le groupe $\text{GL}_K(E)$ est réunion de classes de similitude. Avec la terminologie du § V.6 (exemple 4), on voit que les **classes de similitude des automorphismes** de E ne sont autres que les **classes de conjugaison du groupe $\text{GL}_K(E)$** .

Remarque 1 : On a une *action à gauche* du groupe $\text{GL}_K(E)$ sur $\text{Hom}_K(E)$ par la loi $\text{GL}_K(E) \times \text{Hom}_K(E) \rightarrow \text{Hom}_K(E)$, $(\alpha, u) \mapsto \alpha u \alpha^{-1}$. Les classes de similitude de $\text{Hom}_K(E)$ sont les $\text{GL}_K(E)$ -orbites pour cette action.

Matrices semblables

DÉFINITION XI.6.2

Deux matrices $M \in \mathfrak{M}_n(K)$, $M' \in \mathfrak{M}_n(K)$ ($n \geq 1$) sont dites **semblables** ssi il existe $P \in \text{GL}(n, K)$ telle que

$$(2) \quad M' = PMP^{-1}.$$

Soit u_M et $u_{M'}$ les endomorphismes du K -ev K^n représentés par M et M' dans la base canonique. Alors M et M' sont semblables ssi u_M et $u_{M'}$ le sont, au sens de la définition XI.6.1. Donc la relation « les matrices M et M' sont semblables » est **d'équivalence** sur $\mathfrak{M}_n(K)$ (ce qu'il est facile de vérifier directement).

Les classes d'équivalence s'appellent **classes de similitude** dans $\mathfrak{M}_n(K)$. Le groupe $\text{GL}(n, K)$ est union de classes de similitude qui ne sont autres que les classes de conjugaison du groupe. Les classes de similitude sont les $\text{GL}(n, K)$ -orbites de $\mathfrak{M}_n(K)$ pour l'action $(P, M) \mapsto PMP^{-1}$.

Mais il existe une autre interprétation fructueuse de la similitude de deux matrices : soit E un K -ev de dimension n ; d'après la relation (11) du § XI.3, les matrices M et M' sont semblables ssi il existe des bases \mathcal{B} et \mathcal{B}' de E telles que M et M' représentent dans \mathcal{B} et \mathcal{B}' le **même endomorphisme**.

L'étude des classes de similitude dans $\text{Hom}_K(E)$ (ou dans $\mathfrak{M}_n(K)$, ce qui revient au même d'après ce qui précède) sera l'objet du dernier chapitre de cet ouvrage (Chap. XVI).

Exercice 1 : a) Soit n un entier ≥ 2 et $u \in \text{Hom}_K(K^n)$. Si u n'est pas une homothétie, on a e_1 et e_2 dans K^n tels que $u(e_1) = e_2$ et (e_1, e_2) linéairement indépendants.

b) Utiliser a) pour prouver : si $M \in \mathfrak{M}_n(K)$ est de trace nulle, et si K est de caractéristique 0, alors M est semblable à une matrice $[a_{ij}]$ telle que $a_{ii} = 0$ pour $1 \leq i \leq n$.

Exercice 2 : Le corps de base K est commutatif quelconque. Soit n un entier ≥ 2 . Montrer que la matrice $J_n = [a_{ij}]$, où $a_{i,i+1} = 1_K$ si $i \in \llbracket 1, n-1 \rrbracket$ et où $a_{ij} = 0$ pour tous les autres couples (i, j) est semblable à tJ_n .

Exercice 3 : Soit E un K -ev de dimension $n \geq 1$. Si deux endomorphismes u et v de E sont semblables, et si $v = \alpha u \alpha^{-1}$ avec $\alpha \in \text{GL}_K(E)$, étudier l'action de α sur les sous- K -ev u -stables de E .

Exercice 4 : Soit E un K -ev de dimension $n \geq 2$. On dit que u est un endomorphisme *nilpotent* s'il existe un entier p tel que $u^{p-1} \neq 0$ et $u^p = 0$.

a) Montrer qu'il existe un vecteur x de E tel que $(x, u(x), \dots, u^{p-1}(x))$ soit libre, et donc que $p \leq n$.

b) On suppose $p = n$. Montrer que la matrice de u dans une base quelconque est semblable à la matrice tJ_n de l'exercice 2.

c) Montrer que $u - \text{Id}_E$ est un automorphisme de E et calculer son inverse.

Exercice 5 : a) Soit A et B deux matrices semblables de $\mathfrak{M}_n(K)$. Montrer que $\forall a \in K$, les matrices $A - aI_n$ et $B - aI_n$ sont semblables.

b) Montrer que, pour un a bien choisi les matrices $M = \begin{bmatrix} 4 & -1 & 0 \\ 0 & 2 & 2 \\ 4 & -2 & 0 \end{bmatrix}$ et $\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}$ de $\mathfrak{M}_3(\mathbb{C})$ sont semblables.

Exercice 6 (Application de l'exercice 7, du § IX.4).

Soit $M \in \mathfrak{M}_n(K)$ ($n \geq 2$); on suppose M non inversible et non nilpotente (donc $M^k \neq 0$ pour tout $k \in \mathbb{N}$).

Démontrer que M est semblable à une matrice du type suivant (écriture par blocs): $\begin{bmatrix} G & O \\ O & N \end{bmatrix}$, avec $G \in \text{GL}_s(K)$ pour s convenable ($1 \leq s \leq n-1$) et $N \in \mathfrak{M}_{n-s}(K)$, N nilpotente.

Exercice 7 (suite de l'exercice 13 du § XI.2).

On donne un entier $n \geq 2$; on appellera *groupe de matrices* dans $\mathfrak{M}_n(K)$ tout sous-ensemble non vide de matrices qui, pour le produit des matrices, est un groupe. L'élément neutre d'un tel groupe n'est pas forcément I_n .

a) Si un groupe de matrices contient une matrice inversible dans $\mathfrak{M}_n(K)$, alors ce groupe est un sous-groupe de $\text{GL}(n, K)$ (et donc son élément neutre est I_n).

b) Soit \mathcal{G} un groupe de matrices non inversibles et J son élément neutre. Montrer que tous les éléments de \mathcal{G} ont le même rang r . (Indication : pour $M \in \mathcal{G}$, $MJ = JM = M$). Soit φ l'élément de $\text{Hom}_K(K^n)$ de matrice J dans la base canonique de K^n . Montrer que :

$$E = \text{Im}(\varphi) \oplus \text{Ker}(\varphi), \text{ et que } \text{Im}(\varphi) = \text{Ker}(\text{Id}_E - \varphi).$$

c) Soit $r \in \llbracket 1, n-1 \rrbracket$; montrer que l'ensemble Γ_r , formé des matrices du type $\begin{bmatrix} G & O \\ O & O \end{bmatrix} \in \mathfrak{M}_n(K)$, avec $G \in \text{GL}(r, K)$, est un groupe de matrices, et en préciser l'élément neutre J_r .

d) On reprend un groupe \mathcal{G} comme en b) et on suppose le rang commun r des matrices de \mathcal{G} tel que $1 \leq r \leq n-1$. Montrer que J est semblable à J_r .

Si $p \in \text{GL}(n, K)$ est tel que $J = PJ_r P^{-1}$, démontrer que l'application $M \mapsto P^{-1}MP$ définit un homomorphisme injectif de \mathcal{G} dans le groupe Γ_r . Conclure en décrivant tous les groupes de matrices dans $\mathfrak{M}_n(K)$.

Exercice 8 : Soit p et q deux entiers ≥ 1 , et $n = p + q$.

On considère la matrice $M \in \mathfrak{M}_n(K)$ suivante (écrite par blocs) :

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

avec $A \in \text{GL}(p, K)$; $B \in \mathfrak{M}_{p,q}(K)$; $C \in \mathfrak{M}_{q,p}(K)$; $D \in \mathfrak{M}_{q,q}(K)$.

On note T la matrice $CA^{-1}B - D$ ($T \in \mathfrak{M}_{q,q}(K)$).

a) Montrer que : $\text{rg}(M) = p + \text{rg}(T)$.

b) En déduire : $\text{rg}(M) = p$ ssi $T = 0$.

Exercice 9 : On donne un entier $n \geq 2$; on appellera *algèbre de matrices* dans $\mathfrak{M}_n(K)$ tout sous-ensemble non vide \mathcal{A} de $\mathfrak{M}_n(K)$ qui est un sous- K -ev stable pour le produit et qui, muni de cette structure de K -ev et de ce produit, est une K -algèbre.

a) Soit $r \in \llbracket 1, n-1 \rrbracket$; montrer que l'ensemble \mathcal{A}_r , formé des matrices du type $\begin{bmatrix} M & O \\ O & O \end{bmatrix} \in \mathfrak{M}_n(K)$, avec $M \in \mathfrak{M}_r(K)$ est une algèbre non nulle de matrices, préciser son élément neutre J_r ($J_r = \text{unité de } \mathcal{A}_r$).

b) Soit \mathcal{A} une algèbre non nulle de matrices et E son élément unité; montrer que $\text{rg}(M) \leq r = \text{rg}(E)$ pour toute $M \in \mathcal{A}$.

Montrer que si $r = n$, alors $E = I_n$ et \mathcal{A} est une sous-algèbre de $\mathfrak{M}_n(K)$.

c) Avec les notations du b), on suppose $r \in \llbracket 1, n-1 \rrbracket$. En reprenant la méthode de l'exercice 7b, montrer que E est semblable à J_r , et que, si $P \in \text{GL}(n, K)$ vérifie $E = PJ_r P^{-1}$, l'application $M \mapsto P^{-1}MP$ définit un homomorphisme de K -algèbres injectif de \mathcal{A} dans \mathcal{A}_r . Conclure en décrivant toutes les algèbres de matrices dans $\mathfrak{M}_n(K)$.

Chapitre XII

DUALITÉ. ESPACES VECTORIELS QUOTIENTS

Dans tout ce chapitre, K désigne un corps commutatif fixé une fois pour toutes.

§ XII.1 DUAL ; FORME BILINÉAIRE CANONIQUE

Soit E un K -ev. Nous avons défini au § IX.4 le **dual algébrique** de E : c'est le K -ev $\text{Hom}_K(E, K)$, noté E^* ⁽¹⁾, dont les éléments sont appelés **formes linéaires** sur E .

Transposée

Soit E et F deux K -ev, et $u \in \text{Hom}_K(E, F)$. Pour toute forme linéaire φ sur F , $\varphi \circ u$ est une forme linéaire sur E (composée de deux applications linéaires). On pose :

DÉFINITION XII.1.1

⎧ Avec les notations ci-dessus, l'application $F^* \rightarrow E^*$, $\varphi \mapsto \varphi \circ u$
⎧ s'appelle **transposée** de u , et se note ${}^t u$.

Il est facile de vérifier les résultats rassemblés dans la proposition ci-dessous :

PROPOSITION XII.1.1

Soit E, F, G trois K -ev :

(I) Pour $u \in \text{Hom}_K(E, F)$, on a : ${}^t u \in \text{Hom}_K(F^*, E^*)$, et l'application $\text{Hom}_K(E, F) \rightarrow \text{Hom}_K(F^*, E^*)$, $u \mapsto {}^t u$, est K -linéaire.

(II) On a : ${}^t \text{Id}_E = \text{Id}_{E^*}$.

(III) Pour $u \in \text{Hom}_K(E, F)$ et $v \in \text{Hom}_K(F, G)$, on a :

$${}^t(v \circ u) = {}^t u \circ {}^t v.$$

⁽¹⁾ Certains auteurs le notent \check{E} au lieu de E^* .

En conséquence, supposons E et F isomorphes, et soit $u : E \rightarrow F$ un isomorphisme de K -ev, et $v = u^{-1}$; alors $v \circ u = \text{Id}_E$ et $u \circ v = \text{Id}_F$, d'où, en utilisant les propriétés (II) et (III) ci-dessus :

$${}^t u \circ {}^t v = \text{Id}_{E^*} \quad \text{et} \quad {}^t v \circ {}^t u = \text{Id}_{F^*},$$

donc ${}^t u$ et ${}^t v$ sont des isomorphismes réciproques l'un de l'autre. En particulier, si E est fixé, l'application $u \mapsto {}^t u$, $\text{Hom}_K(E) \rightarrow \text{Hom}_K(E^*)$ envoie $\text{GL}_K(E)$ dans $\text{GL}_K(E^*)$. Fixons E . L'application $E \times E^* \rightarrow K$, $(x, \varphi) \mapsto \langle x, \varphi \rangle = \varphi(x)$ est évidemment linéaire en x (resp. en φ) pour φ fixé (resp. pour x fixé) : on dit qu'elle est **bilinéaire** sur le K -ev produit $E \times E^*$, et on l'appelle **forme bilinéaire canonique sur $E \times E^*$** .

Avec cette convention d'écriture, si E et F sont deux K -ev, et si $u \in \text{Hom}_K(E, F)$, la transposée ${}^t u$ vérifie :

$$(1) \quad \forall (x, \varphi) \in E \times F^* \quad \langle x, {}^t u(\varphi) \rangle = \langle u(x), \varphi \rangle.$$

Les définitions prouvent d'ailleurs que ${}^t u$ est la seule application $v : F^* \rightarrow E^*$ vérifiant

$$\forall (x, \varphi) \in E \times F^* \quad \langle x, v(\varphi) \rangle = \langle u(x), \varphi \rangle.$$

DÉFINITION XII.1.2

Soit E un K -ev, $A \subset E$ et $B \subset E^*$. On appelle **orthogonal de A dans E^*** , et on note A^0 ou A^\perp l'ensemble $\{\varphi \in E^* \mid A \subset \text{Ker } \varphi\}$.
On appelle **orthogonal de B dans E** , et on note ${}^0 B$ ou ${}^\perp B$, l'ensemble

$$\{x \in E \mid \forall \varphi \in B, \varphi(x) = 0\} = \bigcap_{\varphi \in B} \text{Ker } (\varphi).$$

Remarque 1 : La précision « orthogonal de B dans E » est importante : on peut aussi considérer l'orthogonal B^0 de B dans $(E^*)^*$, qu'il ne faut pas confondre avec ${}^0 B$.

On convient que si $A = \emptyset$ (resp. $B = \emptyset$), alors

$$A^0 = E^* \quad (\text{resp. } {}^0 B = E).$$

Des vérifications élémentaires montrent :

PROPOSITION XII.1.2

Avec les notations de la définition XII.1.2 :

- (I) A^0 est un sous- K -ev de E^* , et ${}^0 B$ est un sous- K -ev de E .
- (II) $A^0 = (\text{Vect } A)^0$ et ${}^0 B = {}^0(\text{Vect } B)$.
- (III) Si $A_1 \subset A_2 \subset E$, alors $A_2^0 \subset A_1^0$; si $B_1 \subset B_2 \subset E^*$, alors ${}^0 B_2 \subset {}^0 B_1$.

En particulier, pour toute partie A de E (resp. B de E^*), on a : $A \subset {}^0(A^0)$ et $B \subset ({}^0B)^0$, ces inclusions étant *strictes* en général.

Précisons cependant que lorsque A et B sont des sous- K -ev, on a en général $B \subsetneq ({}^0B)^0$ mais qu'on a toujours $A = {}^0(A^0)$ ⁽¹⁾.

THÉOREME XII.1.1

|| Soit E et F deux K -ev et $u \in \text{Hom}_K(E, F)$. Alors $\text{Ker } ({}^t u) = (\text{Im } (u))^0$.

Démonstration :

$$\begin{aligned} \text{Par définition, } \text{Ker } ({}^t u) &= \{ \varphi \in F^* \mid \varphi \circ u = 0 \} = \\ &= \{ \varphi \in F^* \mid \text{Im } (u) \subset \text{Ker } (\varphi) \} = (\text{Im } (u))^0. \quad \blacksquare \end{aligned}$$

Duaux itérés

A partir d'un K -ev E , on peut former la suite de K -ev $E^{(*)n}$ définie par $E^{(*)0} = E$ et $(\forall n \geq 0) E^{(*)n+1} = [E^{(*)n}]^*$. Cette suite s'appelle suite des **duaux itérés** de E ; $E^{(*)1}$ est le dual E^* ; $E^{(*)2}$ est le K -ev $(E^*)^*$, appelé **bidual** de E , et noté E^{**} ; $E^{(*)3}$ est appelé le *tridual*, et pour n quelconque, $E^{(*)n}$ est le n -ième dual de E . Nous verrons au § XII.2 que, si E est de dimension finie, cette suite, moyennant des identifications naturelles, est *périodique*, les espaces $E^{(*)2n+1}$ s'identifiant à E^* , et les $E^{(*)2n}$ s'identifiant à E . En revanche, pour E de dimension infinie, la suite $E^{(*)n}$ est formée de K -ev tous distincts, et qui « grandissent » extrêmement vite.

Soit E un K -ev. Si $x \in E$, on a vu que l'application $\delta_x : E^* \rightarrow K$, $\varphi \mapsto \langle x, \varphi \rangle = \varphi(x)$, est K -linéaire. C'est donc une *forme linéaire* sur E^* , et on vérifie que l'application $J_E : E \rightarrow E^{**}$, $x \mapsto \delta_x$, est K -linéaire. On l'appelle **l'application linéaire canonique de E dans E^{**}** . On peut montrer que J_E est toujours *injective* (mais la démonstration utilise le fait que tout espace vectoriel a au moins une base, ce qui fait intervenir l'axiome du choix, et dépasse donc le niveau de ce livre) (cf. exercice 3).

Soit E et F deux K -ev et $u \in \text{Hom}_K(E, F)$. On peut former ${}^t(u)$, que l'on note ${}^{''}u$ pour abréger : c'est une application linéaire de E^{**} dans F^{**} , et on a :

$$(2) \quad {}^{''}u \circ J_E = J_F \circ u.$$

En effet, soit $x \in E$ et $\varphi \in F^*$. On a : $({}^{''}u \circ J_E)(x) = \delta_x \circ {}^t u$, donc $\langle \varphi, ({}^{''}u \circ J_E)(x) \rangle = \langle \varphi, \delta_x \circ {}^t u \rangle = \delta_x(\varphi \circ u) = \varphi(u(x)) = \langle u(x), \varphi \rangle = \langle \varphi, \delta_{u(x)} \rangle = \langle \varphi, J_F(u(x)) \rangle$, d'où $({}^{''}u \circ J_E)(x) = (J_F \circ u)(x)$ pour tout $x \in E$, d'où (2).

⁽¹⁾ La démonstration de ce dernier résultat sort du cadre imparti à cet ouvrage dans le cas où E est de dimension infinie, mais elle sera donnée plus loin pour E de dim

Bidual et orthogonalité

Soit E un K -ev et V un sous- K -ev de E^* . Les conventions ci-dessus conduisent à considérer 0V (orthogonal de V dans E) et V^0 (orthogonal de V dans E^{**}). Si $x \in E$, on a : $x \in {}^0V$ ssi $(\forall \varphi \in V), \varphi(x) = 0$, i.e. $\langle \varphi, J_E(x) \rangle = 0$, soit encore : $J_E(x) \in V^0$. Par suite :

$$(3) \quad {}^0V = J_E^{-1}(V^0).$$

Exercice 1 : Si E et F sont deux K -ev, et si $u \in \text{Hom}_K(E, F)$ est surjectif, alors ${}^u u$ est injectif.

Exercice 2 : Soit I un ensemble non vide, et E le K -ev K^I , F son sous- K -ev $K^{(I)}$. On note $(e_i)_{i \in I}$ la base canonique de $K^{(I)}$. Soit, pour $i \in I$, φ_i la forme linéaire sur F telle que $(\forall j) \varphi_i(e_j) = \delta_{ij} 1_K$ (δ symbole de Kronecker).

a) Prouver que l'application $F^* \rightarrow E, \varphi \mapsto (\varphi(e_i))_{i \in I}$ est un isomorphisme de K -ev ; quel est le sous- K -ev de F^* engendré par les $(\varphi_i)_{i \in I}$?

b) On prend $K = \mathbb{Q}$ et $I = \mathbb{N}$. Montrer que $\mathbb{Q}^{(\mathbb{N})}$ est dénombrable, que $\mathbb{Q}^{\mathbb{N}}$ ne l'est pas, et que le \mathbb{Q} -ev $\mathbb{Q}^{(\mathbb{N})}$ n'est pas de dimension finie. En déduire que $\mathbb{Q}^{(\mathbb{N})}$ n'est isomorphe au dual d'aucun \mathbb{Q} -ev.

Exercice 3 : Soit E un K -ev, supposé muni d'une base $\mathcal{B} = (e_i)_{i \in I}$. Démontrer que l'application canonique $J_E : E \rightarrow E^{**}$ est injective (lorsque I est infini, on peut prouver que J_E n'est pas surjective, mais cela déborde le cadre de cet ouvrage).

Exercice 4 : On reprend les notations de l'exercice 2, en supposant I infini. Montrer que si $W = \text{Vect}(\varphi_i)_{i \in I}$, on a : $W \subsetneq ({}^0W)^0$.

Exercice 5 : Soit E un K -ev et V_1, V_2 (resp. W_1, W_2) deux sous- K -ev de E (resp. de E^*). Comparer $(V_1 + V_2)^0$ et $V_1^0 \cap V_2^0$ ainsi que : $(V_1 \cap V_2)^0$ et $V_1^0 + V_2^0$; puis ${}^0(W_1 + W_2)$ et ${}^0W_1 \cap {}^0W_2$, et enfin ${}^0(W_1 \cap W_2)$ et ${}^0W_1 + {}^0W_2$.

Exercice 6 : Soit E un ensemble non vide muni d'une loi interne notée $+$, et d'une loi externe de domaine K notée $(\lambda, x) \mapsto \lambda \cdot x, K \times E \rightarrow E$. On note F l'ensemble des $\varphi \in \mathcal{F}(E, K)$ telles que $\forall (x, y) \in E^2, \forall \lambda \in K \quad \varphi(x + y) = \varphi(x) + \varphi(y)$ et $\varphi(\lambda x) = \lambda \varphi(x)$. On suppose que $(\forall (x, y) \in E^2), (x \neq y) \Rightarrow \exists \varphi \in F \mid \varphi(x) \neq \varphi(y)$. Montrer que $(E, +, \cdot)$ est un K -ev (ENS de Sèvres, oral).

Exercice 7 : Soit E un K -ev et E_1, \dots, E_n des sous- K -ev tels que $E = \bigoplus_{i=1}^n E_i$. Montrer que E^* et $\bigoplus_{i=1}^n E_i^*$ sont canoniquement isomorphes. Plus généralement, si $E = \bigoplus_{i \in I} E_i$ avec I non vide arbitraire, E^* et $\prod_{i \in I} E_i^*$ sont canoniquement isomorphes (cf. théorème IX.2.2).

§ XII.2 DUALITÉ EN DIMENSION FINIE

Rappelons (théorème IX.4.7) que le dual E^* d'un K -ev E de dimension finie $n \geq 1$ est aussi de dimension finie n .

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , une base de E^* est constituée des formes $(e_i^*)_{1 \leq i \leq n}$ telles que $(\forall (i, j)) \langle e_i, e_j^* \rangle = \delta_{ij} \cdot 1_K$. La base $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ est appelée la **base duale** de \mathcal{B} . C'est (cf. théorème VI.3.2) l'unique suite $(\varphi_1, \dots, \varphi_n) \in (E^*)^n$ telle que

$$(\forall (i, j)) \langle e_i, \varphi_j \rangle = \delta_{ij} \cdot 1_K.$$

En outre, pour tout vecteur $x \in E$, on a

$$(1) \quad x = \sum_{i=1}^n \langle x, e_i^* \rangle e_i.$$

LEMME 1

|| Soit E un K -ev de dimension finie $n \geq 1$, et $x \in E \setminus \{0\}$. Il existe au moins une forme linéaire $\varphi \in E^*$ telle que $\varphi(x) \neq 0$.

Démonstration :

Le théorème de la base incomplète permet de construire une base $\mathcal{B} = (x, e_2, \dots, e_n)$. Si $\mathcal{B}^* = (e_i^*)$ est sa base duale, on a $e_1^*(x) = 1_K \neq 0$, donc $\varphi = e_1^*$ convient. ■

THÉORÈME XII.2.1

|| Soit E un K -ev et F un K -ev de dimension finie. L'application $\tau : \text{Hom}_K(E, F) \rightarrow \text{Hom}_K(F^*, E^*)$, $u \mapsto {}^t u$ est injective, et si E et F sont tous deux de dimension finie, c'est un isomorphisme.

Démonstration :

Le noyau N de τ est

$$N = \{u \in \text{Hom}_K(E, F) \mid \forall \varphi \in F^*, \varphi \circ u = 0\}.$$

Or, si, pour un $x \in E$, on avait $u(x) \neq 0$, le lemme 1 fournirait $\varphi \in F^*$ tel que $\varphi(u(x)) \neq 0$, d'où $\varphi \circ u \neq 0$, donc $u \notin N$, donc $N = \{0\}$.

Supposons en plus que E est de dimension finie. Alors les K -ev $\text{Hom}_K(E, F)$ et $\text{Hom}_K(F^*, E^*)$ ont même dimension, car

$$\dim E = \dim E^* \quad \text{et} \quad \dim F = \dim F^*$$

(cf. théorèmes IX.4.6 et IX.4.7), et l'application linéaire τ , qui est injective, est donc bijective. ■

THÉORÈME XII.2.2

|| Soit E et F deux K -ev non nuls de dimensions finies n et p , munis de bases $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_p)$, et soit \mathcal{B}^* et \mathcal{C}^* leurs bases duales. Alors, si $u \in \text{Hom}_K(E, F)$, on a :

$$\text{Mat}_{\mathcal{C}^*, \mathcal{B}^*}({}^t u) = {}^t [\text{Mat}_{\mathcal{B}, \mathcal{C}}(u)].$$

Démonstration :

Posons $M = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = [a_{ij}]$, d'où $u(e_j) = \sum_{i=1}^p a_{ij} f_i$ ($1 \leq j \leq n$). On a ${}^t u(f_l^*) = f_l^* \circ u$ ($1 \leq l \leq p$), donc $\langle e_j, {}^t u(f_l^*) \rangle = f_l^*(u(e_j)) = \sum_{k=1}^p a_{kj} f_l^*(f_k) = a_{lj}$. Si nous posons :

$$\text{Mat}_{\mathcal{C}^*, \mathcal{B}^*}({}^t u) = [b_{ij}], \quad {}^t u(f_l^*) = \sum_{k=1}^n b_{kl} e_k^*.$$

Alors
$$\langle e_j, {}^t u(f_l^*) \rangle = \sum_{k=1}^n b_{kl} \langle e_j, e_k^* \rangle = b_{jl}.$$

Cela prouve : $b_{jl} = a_{lj}$ pour tous $(l, j) \in \llbracket 1, p \rrbracket \times \llbracket 1, n \rrbracket$, autrement dit $\text{Mat}_{\mathcal{C}^*, \mathcal{B}^*}({}^t u) = {}^t M$. ■

Ce théorème XII.2.2 fournit évidemment une deuxième démonstration de la dernière assertion du théorème XII.2.1 (on choisit des bases \mathcal{B} et \mathcal{C} de E et F et l'assertion se ramène au fait que $M \mapsto {}^t M$ est un isomorphisme de $K\text{-ev}$ de $\mathfrak{M}_{p,n}(K)$ sur $\mathfrak{M}_{n,p}(K)$).

THÉORÈME XII.2.3

Soit E un $K\text{-ev}$ quelconque et $\varphi_1, \varphi_2, \dots, \varphi_p$ des formes linéaires sur E ($p \geq 1$). Notons Φ l'application linéaire :

$$E \rightarrow K^p, \quad x \mapsto (\varphi_1(x), \varphi_2(x), \dots, \varphi_p(x)).$$

Alors Φ est **surjective** ssi les formes linéaires $\varphi_1, \varphi_2, \dots, \varphi_p$ sont **linéairement indépendantes**.

Démonstration :

A chaque $\lambda = (\lambda_1, \dots, \lambda_p) \in K^p$, associons la forme linéaire ζ_λ sur K^p définie par $\zeta_\lambda(y_1, \dots, y_p) = \sum_{k=1}^p \lambda_k y_k$. On obtient ainsi un isomorphisme de $K\text{-ev}$ de K^p sur $(K^p)^*$. (Si (e_1, \dots, e_p) est la base canonique de K^p , et si (e_i^*) est sa base duale, cet isomorphisme associe e_i^* à e_i pour $1 \leq i \leq p$.)

Soit $I = \text{Im}(\Phi) \subset K^p$. Pour qu'on ait $I \neq K^p$, il faut et il suffit que I soit contenu dans au moins un hyperplan de K^p (corollaire 2 du théorème IX.5.1), donc (théorème IX.5.2) qu'il existe une forme linéaire $\zeta \in (K^p)^*$, *non nulle*, et nulle sur I . Compte tenu de la description ci-dessus de $(K^p)^*$, cette condition $I \neq K^p$ équivaut donc à

$$\text{« il existe } (\lambda_1, \dots, \lambda_p) \in K^p \setminus \{0\} \text{ tel que } \sum_{k=1}^p \lambda_k y_k = 0$$

pour tout $(y_1, \dots, y_p) \in I$ », c'est-à-dire à :

« il existe $(\lambda_1, \dots, \lambda_p) \in K^p \setminus \{0\}$ tel que $(\forall x \in E) \sum_{k=1}^p \lambda_k \varphi_k(x) = 0$ »,

ou encore à :

« il existe $(\lambda_1, \dots, \lambda_p) \in K^p \setminus \{0\}$ tel que $\sum_{k=1}^p \lambda_k \varphi_k = 0$ »,

ce qui signifie que les formes linéaires $\varphi_1, \varphi_2, \dots, \varphi_p$ sont liées. ■

Bidual en dimension finie

THÉORÈME XII.2.4

|| Si E est un K -ev de dimension finie, l'application linéaire canonique $J_E: E \rightarrow E^{**}$ est bijective.

Démonstration :

On sait déjà que $\dim(E) = \dim(E^*) = \dim(E^{**})$. Il suffit donc de prouver que J_E est injective, c'est-à-dire que $x \in E$ et $x \neq 0_E$ entraîne $J_E(x) \neq 0_{E^{**}}$. Or soit $x \in E \setminus \{0\}$. Il existe $\varphi \in E^*$ tel que $\varphi(x) \neq 0$ (cf. lemme 1). Alors $(J_E(x))(\varphi) = \varphi(x) \neq 0$. Donc $J_E(x) \neq 0$. ■

Si l'on identifie, à l'aide de J_E , chaque K -ev E de dimension finie avec son bidual E^{**} , la relation (2) du § XII.1 montre que toute application linéaire u sera identifiée à sa bitransposée " u ".

COROLLAIRE 1

|| Si E est un K -ev de dimension finie, pour tout sous- K -ev V de E^* , on a : $V^0 = J_E({}^0V)$; les espaces V^0 et 0V sont donc canoniquement isomorphes.

(C'est une conséquence immédiate de la relation (3) du § XII.1 et du fait que J_E est un isomorphisme.) Si l'on identifie E et E^{**} à l'aide de J_E , on voit que V^0 et 0V deviennent un même sous- K -ev de E .

COROLLAIRE 2

|| Soit E un K -ev de dimension finie $n \geq 1$. Pour toute base $(\varphi_1, \varphi_2, \dots, \varphi_n)$ de E^* , il existe une et une seule base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $(\varphi_1, \dots, \varphi_n)$ soit la base duale de \mathcal{B} . Cette base \mathcal{B} est l'unique suite de vecteurs $(v_1, \dots, v_n) \in E^n$ telle que

(4) $\forall (i, j) \in \llbracket 1, n \rrbracket^2 \quad \varphi_i(v_j) = \delta_{ij} \cdot 1_K.$

Démonstration :

L'existence et l'unicité d'une suite $(w_1, \dots, w_n) \in (E^{**})^n$ telle que (5) $\langle \varphi_i, w_i \rangle = \delta_{ij} \cdot 1_K$ résulte du théorème VI.3.2. Notons $(\varepsilon_1, \dots, \varepsilon_n)$ cette suite : elle n'est autre que la *base duale* de $(\varphi_1, \dots, \varphi_n)$ dans E^{**} . Pour $(w_1, \dots, w_n) \in (E^{**})^n$, posons :

$$(v_1, \dots, v_n) = (J_E^{\langle -1 \rangle}(w_1), \dots, J_E^{\langle -1 \rangle}(w_n)).$$

Les relations (5) *équivalem* à (4), d'où l'existence et l'unicité d'une suite (v_i) vérifiant (4). De plus cette suite est $(J_E^{\langle -1 \rangle}(\varepsilon_1), \dots, J_E^{\langle -1 \rangle}(\varepsilon_n))$. Notons-la (e_1, \dots, e_n) : puisque J_E est bijectif, (e_1, \dots, e_n) est une base de E . ■

Ainsi, avec ces hypothèses, la correspondance qui associe, à *chaque* base (e_1, \dots, e_n) de E sa base duale, établit une *bijection* entre les bases indexées par $\llbracket 1, n \rrbracket$ de E et les bases indexées par $\llbracket 1, n \rrbracket$ de E^* .

Finalement, tous ces résultats montrent que, lorsque E est de dimension finie, les couples (E, E^*) et $(E^*, E^{**} \cong E)$ jouent des rôles symétriques.

Exemple 1 : Supposons le corps K de caractéristique 0. Soit $n \in \mathbb{N}^*$ et E le K -ev $K_{n-1}[X]$. Notons φ_k la forme linéaire

$$P \mapsto P^{(k)}(0) \quad \text{sur } E \quad (0 \leq k \leq n-1).$$

La suite $(\varphi_0, \varphi_1, \dots, \varphi_{n-1})$ est *libre*, car si $\sum_{k=0}^{n-1} \lambda_k \varphi_k = 0$ ($\lambda_i \in K$), en appliquant cette relation aux polynômes $1, X, \dots, X^{n-1}$, on trouve $\lambda_k k! = 0$ pour $0 \leq k \leq n-1$, d'où $\lambda_k = 0$. Puisque $\dim(E) = n$, $(\varphi_0, \varphi_1, \dots, \varphi_{n-1})$ est une base de E^* . Notant $e_k = X^k$ ($0 \leq k \leq n-1$), on a

$$\langle e_i, \varphi_j \rangle = i! \delta_{ij} \quad (i \in \llbracket 0, n-1 \rrbracket, j \in \llbracket 0, n-1 \rrbracket).$$

Donc la base de E dont $(\varphi_0, \dots, \varphi_{n-1})$ est la base duale est

$$\left(1, X, \dots, \frac{X^k}{k!}, \dots, \frac{X^{n-1}}{(n-1)!} \right).$$

Orthogonalité en dimension finie

THÉORÈME XII.2.5

Soit E un K -ev de dimension finie $n \geq 1$. Pour tout sous- K -ev V de E , on a :

$$(6) \quad \dim(V) + \dim(V^0) = n.$$

Pour tout sous- K -ev W de E , on a :

$$(7) \quad \dim(W) + \dim({}^0W) = n.$$

Démonstration :

Notons dès l'abord que (7) se déduit de (6) en remplaçant le couple (E, E^*) par (E^*, E^{**}) et en appliquant le corollaire 1 du théorème XII.2.4. Montrons (6). Si $V = E$, il est évident que $V^0 = \{0\}$; si $V = \{0\}$ il est non moins évident que $V^0 = E^*$. Il reste à examiner le cas où $\dim(V) = p \in \llbracket 1, n-1 \rrbracket$ avec $n \geq 2$. Choisissons alors une base (e_1, \dots, e_p) de V , et complétons-la en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E , dont nous notons $(e_1^*, \dots, e_n^*) = \mathcal{B}^*$ la base duale.

Une forme linéaire $\varphi = \sum_{i=1}^n \lambda_i e_i^*$ appartient à V^0 ssi $\langle e_k, \varphi \rangle = 0$ pour

$1 \leq k \leq p$, c'est-à-dire : $\sum_{i=1}^n \lambda_i \langle e_k, e_i^* \rangle = 0$, ou encore $\lambda_k = 0$ pour

$1 \leq k \leq p$. Donc $V^0 = \text{Vect}(e_{p+1}^*, \dots, e_n^*)$ est bien un K -ev de dimension $n-p$, admettant la base $(e_{p+1}^*, \dots, e_n^*)$. ■

Dans ce qui suit, pour tout K -ev \mathcal{E} et tout entier $p \geq 0$, on note $\mathcal{G}_p(\mathcal{E})$ l'ensemble des sous- K -ev de dimension p de \mathcal{E} .

COROLLAIRE 1

Soit E un K -ev de dimension finie $n \geq 1$. Pour tout $p \in \llbracket 0, n \rrbracket$, l'application $\mathcal{G}_p(E) \rightarrow \mathcal{G}_{n-p}(E^*)$, $V \mapsto V^0$ est bijective, et sa bijection réciproque est $W \mapsto {}^0W$. En particulier, pour tout sous- K -ev V de E (resp. W de E^*) on a :

$$(8) \quad V = {}^0(V^0) \quad \text{et} \quad (9) \quad W = ({}^0W)^0.$$

Démonstration :

Remarquons d'abord que (9) se déduit de (8) en remplaçant (E, E^*) par (E^*, E^{**}) et en appliquant le corollaire 1 du théorème XII.2.4. Pour prouver (8), notons d'abord que $V \subset {}^0(V^0)$. Puis, par application de (6) à V et de (7) à V^0 , on voit que $\dim({}^0(V^0)) = \dim(V)$, d'où $V = {}^0(V^0)$ (cf. théorème IX.4.2). ■

Remarque 1 : Compte tenu du théorème IX.5.2, la relation (8) signifie que V est l'intersection des hyperplans qui le contiennent. On retrouve donc d'une autre façon le corollaire 2 du théorème IX.5.1.

Remarque 2 : En conséquence de ce corollaire 1, si V est un sous- K -ev de E , les relations $V = \{0\}$ et $V^0 = E^*$ sont équivalentes. Cette assertion, qui équivaut au lemme 1, constitue un puissant outil pour mettre en évidence des parties génératrices de E^* (cf. exemple 2 ci-dessous).

De même $V = E$ ssi $V^0 = \{0\}$, et cela permet de découvrir des parties génératrices de E (cf. exemple 3 ci-dessous).

COROLLAIRE 2

|| Soit E et F deux K -ev de dimensions finies n et p , et $u \in \text{Hom}_K(E, F)$. Alors $\text{rg}({}^t u) = \text{rg}(u)$.

Démonstration :

On sait que $\text{Ker}({}^t u) = (\text{Im}(u))^0$; d'où, par application du théorème XII.2.5 :

$$\text{rg}(u) = \dim(F) - \dim[\text{Ker}({}^t u)] .$$

Mais la formule du rang donne $\text{rg}({}^t u) = \dim(F^*) - \dim[\text{Ker}({}^t u)]$, et l'on en conclut que $\text{rg}({}^t u) = \text{rg}(u)$ puisqu'on sait que $\dim(F) = \dim(F^*)$. ■

Bien entendu, ce corollaire 2 aurait aussi pu se déduire du fait, prouvé au chapitre XI (corollaire 2 du théorème XI.4.2), que le rang d'une matrice est égal au rang de sa transposée, combiné avec le théorème XII.2.2.

Exemple 2 : Soit n un entier ≥ 1 , et soit E le K -ev $K_{n-1}[X]$. Donnons-nous a_1, a_2, \dots, a_n , éléments distincts dans K (ce qui sous-entend que $\text{card}(K) \geq n$) et notons φ_i la forme linéaire $P \mapsto P(a_i)$ sur E ($1 \leq i \leq n$). Soit $W = \text{Vect}(\varphi_1, \dots, \varphi_n)$, alors

$${}^0 W = \{P \in E \mid P(a_1) = P(a_2) = \dots = P(a_n) = 0\} = \{0\}$$

à cause du théorème VII.4.2. Donc $W = E^*$, et comme

$$\dim(E) = \dim(E^*) = n ,$$

on a prouvé que $(\varphi_1, \varphi_2, \dots, \varphi_n)$ est une base de E^* , ce qui n'était pas évident *a priori*. En outre, l'unique base (e_1, \dots, e_n) de E dont $(\varphi_1, \dots, \varphi_n)$ est la base duale est définie par les relations $\varphi_i(e_j) = \delta_{ij}$ pour $(i, j) \in \llbracket 1, n \rrbracket^2$, ce qui donne immédiatement $e_i = \frac{1}{A_i} \prod_{j \neq i} (X - a_j)$ (cf. corollaire 1 du théorème VII.4.2), avec $A_i = \prod_{j \neq i} (a_i - a_j)$.

Soit $P \in E$. On a : $P = \sum_{i=1}^n \lambda_i e_i$, d'où $\forall k \in \llbracket 1, n \rrbracket$:

$$P(a_k) = \varphi_k(P) = \sum_{i=1}^n \lambda_i \varphi_k(e_i) = \lambda_k , \quad \text{d'où} \quad P = \sum_{i=1}^n P(a_i) e_i .$$

On retrouve ainsi la *formule d'interpolation de Lagrange* vue au § VII.4.

Considérons maintenant l'application linéaire $\psi : E \rightarrow K^n$,

$$P \mapsto (P(a_1), \dots, P(a_n)) = (\varphi_1(P), \dots, \varphi_n(P))$$

et notons \mathcal{B} la base $(1, X, \dots, X^{n-1})$ de E et $\mathcal{C} = (\varepsilon_1, \dots, \varepsilon_n)$ la base canonique de K^n . L'application ψ est surjective (théorème XII.2.3) car $\varphi_1, \dots, \varphi_n$ sont indépendants, donc bijective car $\dim E = \dim K^n = n$. On a :

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(\psi) = [(a_i)^j]_{(i,j) \in \llbracket 1, n \rrbracket^2},$$

soit sous forme développée :

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(\psi) = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ \vdots & & & & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{bmatrix}.$$

Cette matrice s'appelle *matrice de Vandermonde* de a_1, \dots, a_n , et peut se noter $\text{Vand}(a_1, \dots, a_n)$. Ce qui précède nous prouve sans calculs que la *matrice* $\text{Vand}(a_1, \dots, a_n)$ est *invertible* (sous l'hypothèse faite au départ que a_1, \dots, a_n sont *distincts*).

D'ailleurs ψ^{-1} associe, à $(b_1, \dots, b_n) \in K^n$, l'unique polynôme $P \in E$ tel que $P(a_i) = b_i$ pour tout i . En particulier :

$$\psi^{-1}(\varepsilon_j) = e_j = \frac{1}{A_j} \prod_{k \neq j} (X - a_k) = \frac{1}{A_j} \left(X^{n-1} + \sum_{k=1}^{n-1} (-1)^k \sigma_{j,k} X^{n-1-k} \right),$$

en notant $(\sigma_{j,k})_{1 \leq k \leq n-1}$ les fonctions symétriques élémentaires des termes $(a_l)_{l \in \llbracket 1, n \rrbracket, l \neq j}$. On en déduit la matrice inverse V^{-1} de la matrice

$$V = \text{Vand}(a_1, \dots, a_n) : V^{-1} = \text{Mat}_{\mathcal{C}, \mathcal{B}}(\psi^{-1}) = \text{Mat}_{(1, X, \dots, X^{n-1})}(e_1, e_2, \dots, e_n),$$

soit :

$$V^{-1} = \begin{bmatrix} \frac{(-1)^{n-1}}{A_1} \sigma_{1, n-1} & \dots & \frac{(-1)^{n-1}}{A_n} \sigma_{n, n-1} \\ \frac{(-1)^{n-2}}{A_1} \sigma_{1, n-2} & & \frac{(-1)^{n-2}}{A_n} \sigma_{n, n-2} \\ \vdots & & \vdots \\ \frac{1}{A_1} & \dots & \frac{1}{A_n} \end{bmatrix}.$$

Exemple 3 : Supposons le corps K de caractéristique 0. Soit $n \in \mathbb{N}^*$, et E le K -ev $K_n[X]$. Donnons-nous a_1, \dots, a_{n+1} , éléments distincts dans K et posons

$$P_i = (X + a_i)^n \quad (1 \leq i \leq n+1).$$

Nous allons montrer que (P_1, \dots, P_{n+1}) est une *base* de E . Pour cela, il suffit de montrer que $V = \text{Vect}(P_1, \dots, P_{n+1}) = E$ (car $\dim(E) = n+1$), donc que $V^0 = \{0\}$. Soit donc $\varphi \in E^*$, nulle sur

$$P_i = \sum_{k=0}^n \binom{n}{k} a_i^k X^{n-k}, \quad \varphi(P_i) = 0 \text{ s'écrit :}$$

$$(10) \quad \sum_{k=0}^n \binom{n}{k} a_i^k A_{n-k} = 0, \quad \text{en posant } A_j = \varphi(X^j) \text{ pour } 0 \leq j \leq n.$$

Introduisons le polynôme $Q = \sum_{k=0}^n \binom{n}{k} A_{n-k} X^k \in E$. Les relations (10) signifient que $Q(a_i) = 0$ pour $1 \leq i \leq n+1$, donc entraînent (cf. théorème VII.4.2) $Q = 0$, d'où $\binom{n}{k} A_{n-k} = 0$ pour tout k , d'où, puisque la caractéristique est nulle, $A_j = 0$ pour tout j , et finalement $\varphi = 0$. On a donc bien $V^0 = \{0\}$, et (P_1, \dots, P_{n+1}) est bien une base de E .

Ces exemples montrent combien les théorèmes XII.2.4 et XII.2.5, bien compris, peuvent être performants.

Exercice 1 : On suppose $K = \mathbb{C}$. On donne sur \mathbb{C}^3 les formes linéaires $\varphi_1, \varphi_2, \varphi_3$ définies, pour $(x, y, z) \in \mathbb{C}^3$, par : $\varphi_1(x, y, z) = x + 2y - 3z$, $\varphi_2(x, y, z) = 5x - 3y$ et $\varphi_3(x, y, z) = 2x - y - z$. Vérifier que $(\varphi_1, \varphi_2, \varphi_3)$ est une base de $(\mathbb{C}^3)^*$ et déterminer la base \mathcal{B} de \mathbb{C}^3 dont $(\varphi_1, \varphi_2, \varphi_3)$ est la base duale.

Exercice 2 : Soit E un K -ev de dimension finie $n \geq 1$. Si $x \in E$ et $y \in E$, avec $x \neq y$, il existe $\varphi \in E^*$ telle que $\varphi(x) \neq \varphi(y)$ (on dit que E^* *sépare* les points de E).

Réciproquement, soit $V \subset E^*$ un sous- K -ev de E^* qui *sépare* les points de E , i.e. $\forall (x, y) \in E^2, (x \neq y) \Rightarrow \exists \varphi \in V \mid \varphi(x) \neq \varphi(y)$. Montrer que $V = E^*$.

Exercice 3 : Soit E et F deux K -ev de dimensions finies $n \geq 1$ et $p \geq 1$. Pour chaque couple $(\varphi, x) \in E^* \times F$, on définit $u = \varphi \otimes x \in \text{Hom}_K(E, F)$ par la condition :

$$(\forall z \in E) \quad u(z) = \varphi(z)x.$$

a) Montrer que l'application $\zeta : E^* \times F \rightarrow \text{Hom}_K(E, F), (\varphi, x) \mapsto \varphi \otimes x$, est bilinéaire, et que, si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , de base duale $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$, et si (f_1, \dots, f_p) est une base de F , alors $(e_i^* \otimes f_j)_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket}$ est une base de $\text{Hom}_K(E, F)$. Reconnaître cette base.

b) Montrer : pour tout K -ev G , et toute application bilinéaire $f : E^* \times F \rightarrow G$, il existe une, et une seule, application linéaire $\bar{f} : \text{Hom}_K(E, F) \rightarrow G$ telle que $\bar{f} \circ \zeta = f$.

c) Soit en particulier $g : E^* \times E \rightarrow K, (\varphi, x) \mapsto \varphi(x) = \langle x, \varphi \rangle$. Montrer que $\bar{g}(u) = \text{Tr}(u)$ pour tout $u \in \text{Hom}_K(E)$ (on peut ainsi définir la trace d'un endomorphisme d'une façon intrinsèque).

Exercice 4 : Soit E la K -algèbre $\mathcal{F}(K, K)$. On considère des éléments f_1, \dots, f_n de E linéairement indépendants. En raisonnant par dualité, montrer qu'il existe a_1, \dots, a_n dans K tels que la matrice $[f_i(a_j)]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ soit inversible.

Indication : Soit V le K -ev $\text{Vect}(f_1, \dots, f_n)$. Étudier les formes linéaires sur $V : \varphi_1, \dots, \varphi_n$ telles que $\varphi_i(f) = f(a_i)$ pour $f \in \text{Vect}(f_j)_{1 \leq j \leq n}$.

Exercice 5 : Soit $n \in \mathbb{N}^*$. Pour chaque $d \in \mathbb{N}^*$, on note \mathcal{H}_d le \mathbb{C} -ev des polynômes homogènes de degré d à n lettres X_1, \dots, X_n sur \mathbb{C} . On fixe $p \geq 1$. Montrer que l'ensemble $\{\varphi^p\}_{\varphi \in \mathcal{H}_1}$ engendre le \mathbb{C} -ev \mathcal{H}_p , en étudiant son orthogonal dans \mathcal{H}_p^* .

Exercice 6 : On suppose K infini. Soit n et p des entiers ≥ 1 . On note E le K -ev des polynômes $P \in K[X_1, \dots, X_n]$ tels que $\deg(P) \leq p$.

a) Calculer $N = \dim_K(E)$.

b) Si $a = (a_1, \dots, a_n) \in K^n$, soit $\delta_a \in E^*$ tel que $\delta_a(F) = F(a)$ pour $F \in E$. Montrer qu'il existe $A_{(1)}, \dots, A_{(N)}$ dans K^n tels que $(\delta_{A_{(1)}}, \dots, \delta_{A_{(N)}})$ soit une base de E^* (raisonner par dualité en étudiant l'orthogonal des $(\delta_a)_{a \in K^n}$ dans E).

c) Soit $A_{(1)}, \dots, A_{(N)}$ dans K^n tels que $(\delta_{A_{(i)}})_{i \leq N}$ soit une base de E^* . On note (P_1, \dots, P_N) la base de E dont $(\delta_{A_{(i)}})$ est la base duale. Si $P \in E$, calculer les coordonnées de P sur (P_1, \dots, P_N) . Montrer enfin que les (P_k) sont tous de degré p .

Exercice 7 : Soit E un K -ev de dimension finie $n \geq 1$.

a) Si $u \in \text{Hom}_K(E)$, soit $T_u \in (\text{Hom}_K(E))^*$ tel que $T_u(v) = \text{Tr}(uv)$ pour $v \in \text{Hom}_K(E)$. Alors $u \mapsto T_u, \text{Hom}_K(E) \rightarrow (\text{Hom}_K(E))^*$ est un isomorphisme, et T_{Id_E} engendre le K -ev des $\varphi \in (\text{Hom}_K(E))^*$ telles que $\varphi(uv - vu) = 0$ pour tous u et $v \in \text{Hom}_K(E)$. (cf. exercice 2 du § XI.2).

$$\mathcal{C}_\alpha = \{u \in \text{Hom}_K(E) \mid \alpha u = u \alpha\}.$$

Montrer, pour α, β et γ dans $\text{Hom}_K(E)$: $\text{Tr}([\alpha, \beta] \gamma) = 0$.

c) Soit α et u dans $\text{Hom}_K(E)$ tels que T_u s'annule sur \mathcal{C}_α . Montrer qu'il existe $v \in \text{Hom}_K(E)$ tel que $u = [\alpha, v]$.

Exercice 8 : Soit E un K -ev de dimension finie $n \geq 2$ et $\lambda_1, \dots, \lambda_n$ distincts dans K . On considère une base $\mathcal{B} = (e_1, \dots, e_n)$ de E , et $\alpha \in \text{Hom}_K(E)$ tel que

$$\text{Mat}_{\mathcal{B}}(\alpha) = \text{Diag}(\lambda_1, \dots, \lambda_n).$$

a) Montrer que \mathcal{C}_α est un sous- K -ev de dimension n de $\text{Hom}_K(E)$, dont $(\text{Id}, \alpha, \dots, \alpha^{n-1})$ est une base. (On reprend les notations de l'exercice 7).

b) En déduire : si $v \in \text{Hom}_K(E)$, pour qu'il existe $u \in \text{Hom}_K(E)$ tel que $[\alpha, u] = v$, il faut et il suffit que $\text{Tr}(v\alpha^k) = 0$ pour $k \in \llbracket 0, n-1 \rrbracket$.

c) Soit $v \in \text{Hom}_K(E)$ et $[a_{ij}]$ sa matrice dans \mathcal{B} . Pour qu'il existe $u \in \text{Hom}_K(E)$ tel que $[\alpha, u] = v$, il faut et il suffit que $a_{ii} = 0$ pour tout i .

Exercice 9 : On suppose K de caractéristique nulle. Soit $E = K_{n-1}[X]$ avec $n \in \mathbb{N}^*$. On note Δ l'opérateur de différence de E : $\Delta(P(X)) = P(X+1) - P(X)$ soit $\varphi_k \in E^*$ défini par $\varphi_k(P) = (\Delta^k(P))(0)$. Montrer que $(\varphi_k)_{0 \leq k \leq n-1}$ est une base de E^* , et trouver la base de E dont elle est la base duale.

Exercice 10 (matrices magiques)

Pour toute matrice carrée $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$ (où $n \geq 2$), on définit

$$\lambda_i(M) = \sum_{j=1}^n a_{ij} \quad (1 \leq i \leq n), \quad \gamma_j(M) = \sum_{i=1}^n a_{ij} \quad (1 \leq j \leq n),$$

$$\delta(M) = \text{Tr}(M) = \sum_{i=1}^n a_{ii} \quad \text{et enfin} \quad d(M) = \sum_{k=1}^n a_{k, n+1-k}.$$

On note \mathcal{M} le sous- \mathbb{C} -ev des matrices M telles que

$$\lambda_1(M) = \lambda_2(M) = \dots = \lambda_n(M) = \gamma_1(M) = \dots = \gamma_n(M) = \delta(M) = d(M),$$

et si $M \in \mathcal{M}$ on appelle $\varphi(M)$ la valeur commune de ces nombres (somme magique). On désigne en particulier par \mathcal{M}_0 l'ensemble des $M \in \mathcal{M}$ tels que $\varphi(M) = 0$.

a) Par dualité, calculer $\dim_{\mathbb{C}}(\mathcal{M}_0)$. Montrer que $\mathcal{M} = \mathcal{M}_0 \oplus \Delta$, où Δ est la droite engendrée par la matrice $C = (c_{ij})$ telle que $c_{ij} = 1$ pour tous i, j .

b) Soit $\mathcal{S}_0 = \mathcal{M}_0 \cap \text{Sym}(n, \mathbb{C})$ et $\mathcal{A}_0 = \mathcal{M}_0 \cap \text{Asym}(n, \mathbb{C})$. Calculer de même $\dim_{\mathbb{C}}(\mathcal{S}_0)$ et $\dim_{\mathbb{C}}(\mathcal{A}_0)$.

c) Pour \mathcal{S}_0 et \mathcal{A}_0 , donner une base formée de matrices dont les coefficients appartiennent à $\{0, \pm 1, \pm 2\}$. Ecrire complètement de telles bases pour $n = 3$, $n = 4$ ou

d) Pour $n = 4$, décomposer sur la base trouvée la matrice

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

qui figure sur une célèbre eau-forte de Albrecht Dürer (1471-1528) intitulée « Melencolia ».

N.B. Pour construire un « carré magique » on s'impose la condition draconienne que les a_{ij} doivent constituer l'ensemble $\llbracket 1, n^2 \rrbracket$. Le plus ancien carré magique connu est le « lo-shu »

$$\begin{bmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{bmatrix}$$

qui apparaît dans le « livre des permutations » attribué à Confucius

(VI^e siècle av. J.-C.) mais aurait été aperçu sur le dos d'une tortue 2000 ans plus tôt !

Exercice 11 : Donner une autre preuve des assertions de l'exemple 3 en utilisant le fait que la matrice de Vandermonde est inversible (cf. exemple 2).

Exercice 12 : Soit E et F deux K -ev de dimension finie.

a) A tout couple $(x, y) \in E \times F$, on associe $u = \alpha(x, y) = x \otimes y \in \text{Hom}_K(E^*, F)$ ainsi défini :

$$(\forall \varphi \in E^*) \quad u(\varphi) = \varphi(x)y.$$

Vérifier que l'application : $\alpha : E \times F \rightarrow \text{Hom}_K(E^*, F)$ est bilinéaire, et que $\text{Im}(\alpha)$ engendre le K -ev $\text{Hom}_K(E^*, F)$.

b) Soit G sur K -ev arbitraire. Montrer que, pour toute application bilinéaire $\beta : E \times F \rightarrow G$, il existe une, et une seule, application linéaire $\bar{\beta} \in \text{Hom}_K(E^*, F)$ telle que $\bar{\beta} \circ \alpha = \beta$ (comparer avec l'exercice 3).

Exercice 13 : Soit (e_1, \dots, e_n) une base du K -ev E . Etudier comment est modifiée la base duale (e_1^*, \dots, e_n^*) lorsqu'on fait subir des opérations élémentaires aux vecteurs de la base (e_i) .

Exercice 14 : On considère dans $\mathcal{M}_3(\mathbb{Q})$ les matrices M dont l'inverse est égal à la transposée.

a) Montrer que l'ensemble de ces matrices forme un sous-groupe de $\text{GL}(3, \mathbb{Q})$.

b) Trouver celles de ces matrices qui sont à coefficients dans \mathbb{Z} .

§ XII.3 QUOTIENTS D'ESPACES VECTORIELS

Soit E un K -ev et F un sous- K -ev. Notons provisoirement \hat{E} et \hat{F} les groupes additifs de E et F . Cela nous permet (cf. Chap. V, § 7) de définir le **groupe quotient** \hat{E}/\hat{F} : c'est l'ensemble des classes d'équivalence pour la relation d'équivalence \mathcal{R} définie sur \hat{E} par :

$$(1) \quad x \mathcal{R} y \quad \text{ssi} \quad x - y \in \hat{F} \quad (x \in \hat{E}, y \in \hat{E}),$$

muni de la loi de composition suivante notée additivement, et qui est une loi de groupe abélien :

$$(2) \quad (\forall X \in \hat{E}/\hat{F}, \quad \forall Y \in \hat{E}/\hat{F})$$

$X + Y =$ la classe commune des $x + y \bmod (\mathcal{R})$ pour $(x,$

L'application canonique $\varphi : \hat{E} \rightarrow \hat{E}/\hat{F}$ est un homomorphisme de groupes surjectif, donc le noyau est F . L'élément nul de \hat{E}/\hat{F} est \hat{F} . Les classes de \mathcal{R} sont appelées **classes de E modulo F** : ce sont les ensembles du type $a + F$, avec $a \in E$.

PROPOSITION XII.3.1

|| Avec les notations ci-dessus, il existe une et une seule loi externe de domaine K sur \hat{E}/\hat{F} , $(\lambda, X) \mapsto \lambda X$ telle que $\varphi(\lambda X) = \lambda \varphi(X)$ pour tous $\lambda \in K$ et $x \in \hat{E}$. Muni de cette loi externe, le groupe abélien \hat{E}/\hat{F} devient un K -ev. Pour cette structure de K -ev, φ est K -linéaire, et on a : $\text{Im}(\varphi) = \hat{E}/\hat{F}$ et $\text{Ker}(\varphi) = \hat{F}$.

Démonstration :

Si une telle loi externe existe, on a nécessairement $\lambda X =$ classe commune mod (F) aux λx pour $x \in X$. Réciproquement, si $x \in X$ et $y \in X$, on a : $x - y \in F$, d'où $\lambda(x - y) \in F$ puisque F est un sous- K -ev de E . Donc la classe mod (F) des $(\lambda x)_{x \in X}$ ne dépend que de λ et de X et on peut la noter $\lambda \cdot X$. Il reste à vérifier, et cela ne présente aucune difficulté, que la loi ainsi définie satisfait bien toutes les assertions de la proposition XII.3.1. ■

La loi externe définie dans la proposition XII.3.1 sera dite *multiplication naturelle par les scalaires*.

DÉFINITION XII.3.1

~ Avec les notations de la proposition XII.3.1, le K -ev obtenu en munissant le groupe abélien \hat{E}/\hat{F} de la multiplication naturelle par les scalaires s'appelle **K -ev quotient de E par F** , et se note E/F .
~ L'application linéaire $\varphi : E \rightarrow E/F$ est dite **canonique**.

Les trois théorèmes qui vont suivre permettent une utilisation quasi mécanique du quotient d'espaces vectoriels.

THÉORÈME XII.3.1 (« Propriété universelle du quotient »)

|| Soit E et G deux K -ev, et F un sous- K -ev de E . Soit $u \in \text{Hom}_K(E, G)$ tel que $F \subset \text{Ker}(u)$. Il existe une unique application $\bar{u} : E/F \rightarrow G$ telle que $\bar{u} \circ \varphi = u$, en notant $\varphi : E \rightarrow E/F$ l'application canonique. Cette application \bar{u} est K -linéaire.

Démonstration :

Si \bar{u} existe, soit $X \in E/F$. On a nécessairement $\bar{u}(X) = u(x)$ pour tout $x \in X$. Réciproquement, si $x \in X$ et $y \in X$, on a : $x - y \in F \subset \text{Ker}(u)$, d'où $u(x) = u(y)$, donc la valeur de u es

X . Notons $\bar{u}(X)$ cette valeur constante. On vérifie sans peine que $\bar{u} \in \text{Hom}(E/F, G)$ et on a : $\bar{u} \circ \varphi = u$ par définition même de \bar{u} . ■

On dit que « u se factorise en \bar{u} dans le quotient E/F ».

THÉORÈME XII.3.2

(Décomposition canonique d'une application linéaire)
 Soit E et F deux K -ev, $u \in \text{Hom}_K(E, F)$, $\varphi : E \rightarrow E/\text{Ker}(u)$ et $\psi : \text{Im}(u) \rightarrow F$ les applications canoniques.
 Il existe une et une seule application $\bar{u} : E/\text{Ker}(u) \rightarrow \text{Im}(u)$ telle que $\psi \circ \bar{u} \circ \varphi = u$, et cette application est un **isomorphisme** de K -ev entre $E/\text{Ker}(u)$ et $\text{Im}(u)$.

(L'application \bar{u} est dite obtenue à partir de u par passage au quotient.)

Démonstration :

Soit $v : E/\text{Ker}(u) \rightarrow F$ l'unique application telle que $v \circ \varphi = u$ (cf. théorème XII.3.1). Si \bar{u} existe, ce ne peut être que $v|_{\text{Im}(u)}$. Réciproquement, $\bar{u} = v|_{\text{Im}(u)}$ est K -linéaire et vérifie $\psi \circ \bar{u} \circ \varphi = u$. \bar{u} est surjective par construction. Il reste à montrer qu'elle est injective. Or

$$\begin{aligned} \text{Ker}(\bar{u}) &= \{X \in E/\text{Ker}(u) \mid u(x) = 0 \text{ pour } x \in X\} \\ &= \{\text{Ker}(u)\} = \{0_{E/\text{Ker}(u)}\}, \end{aligned}$$

donc finalement \bar{u} est un isomorphisme de K -ev. ■

THÉORÈME XII.3.3

Soit E un K -ev, F un sous- K -ev et $\varphi : E \rightarrow E/F$ l'application canonique. L'application $\alpha : V \mapsto \varphi(V)$ définit une bijection de l'ensemble $\mathcal{G}_F(E)$ des sous- K -ev de E contenant F sur l'ensemble $\mathcal{G}(E/F)$ des sous- K -ev de E/F . La bijection réciproque est $\beta : W \mapsto \varphi^{-1}(W)$.

Démonstration :

Tout d'abord, si $W \in \mathcal{G}(E/F)$, $\varphi^{-1}(W)$ est un sous- K -ev de E contenant $\text{Ker}(\varphi) = F$, donc β est bien une application de $\mathcal{G}(E/F)$ dans $\mathcal{G}_F(E)$. Si $W \in \mathcal{G}(E/F)$, on a $W = \varphi(\varphi^{-1}(W))$ parce que φ est surjective, d'où $\alpha \circ \beta = \text{Id}_{\mathcal{G}(E/F)}$. Si $V \in \mathcal{G}_F(E)$, on a $V \subset \varphi^{-1}(\varphi(V))$, et aussi $F \subset \varphi^{-1}(\varphi(V))$. Soit $x \in \varphi^{-1}(\varphi(V))$ et $y = \varphi(x)$, d'où $y = \varphi(x')$, avec $x' \in V$; alors $\varphi(x - x') = 0$, d'où $x = x' + z$, avec $z \in F$, d'où $x \in V$ et cela prouve que $\varphi^{-1}(\varphi(V)) = V$, d'où $\beta \circ \alpha = \text{Id}_{\mathcal{G}_F(E)}$. ■

(En fait le raisonnement ci-dessus prouve que, pour tout soi

on a $\varphi^{-1}(V) = V + F$.) Il convient de noter que, dans la bijection α , E correspond à E/F , et F correspond à $\{0_{E/F}\}$.

Application : théorèmes d'isomorphisme

THÉORÈME XII.3.4 (Premier théorème d'isomorphisme de Noether).

|| Soit E un K -ev et F, G des sous- K -ev de E tels que $G \subset F \subset E$. Alors F/G est un sous- K -ev de E/G et les K -ev E/F et $(E/G)/(F/G)$ sont canoniquement isomorphes.

Démonstration :

Une classe de $F \bmod (G)$ est un ensemble du type $a + G$, avec $a \in F$, donc c'est bien une classe de $E \bmod (G)$, d'où $F/G \subset E/G$, et on voit facilement que F/G est un sous- K -ev de E/G .

Soit $\varphi : E \rightarrow E/G$ et $\psi : E/G \rightarrow (E/G)/(F/G)$ les applications canoniques. Alors $\zeta = \psi \circ \varphi : E \rightarrow (E/G)/(F/G)$ est linéaire et surjective. On a : $\text{Ker}(\zeta) = \varphi^{-1}(\text{Ker}(\psi)) = \varphi^{-1}(F/G) = F$ (cf. théorème XII.3.3). Le théorème XII.3.2 fournit donc un isomorphisme de K -ev :

$$\bar{\zeta} : E/F \rightarrow (E/G)/(F/G) . \quad \blacksquare$$

THÉORÈME XII.3.5 (Deuxième théorème d'isomorphisme de Noether).

|| Soit E un K -ev et F et G deux sous- K -ev de E . Les K -ev $(F+G)/G$ et $F/F \cap G$ sont canoniquement isomorphes.

Démonstration :

Notons $\varphi : (F+G) \rightarrow (F+G)/G$ l'application canonique et $\psi : F \rightarrow F/F \cap G$ l'injection canonique $x \mapsto x$.

Alors $\zeta = \varphi \circ \psi : F \rightarrow (F+G)/G$ est linéaire.

On a : $\text{Ker}(\zeta) = \psi^{-1}(\text{Ker}(\varphi)) = \psi^{-1}(G) = F \cap G$.

Montrons que ζ est surjective : soit $z = x + y \in F + G$ avec $x \in F$ et $y \in G$. Alors $\varphi(z) = \varphi(x) = \zeta(x)$, donc ζ est bien surjective et c'est un isomorphisme. \blacksquare

Codimension

Nous avons vu au § IX.1 que tous les supplémentaires dans E d'un sous- K -ev d'un K -ev E sont isomorphes entre eux. Cette propriété est confirmée par le résultat suivant :

THÉORÈME XII.3.6

|| Si F est un sous- K -ev du K -ev E , tout supplémentaire de F dans E est isomorphe à E/F .

Démonstration :

Soit G un supplémentaire de F dans E . Notons $\varphi : E \rightarrow E/F$ l'application canonique, et $\psi = \varphi|_G$. Alors $\text{Ker}(\psi) = G \cap \text{Ker}(\varphi) = G \cap F = \{0\}$. De plus ψ est surjectif

$z \in E/F$, on a : $z = \varphi(t)$ avec $t \in E$, d'où $t = x + y$ ($x \in F, y \in G$), d'où $z = \varphi(x) + \varphi(y) = \varphi(y) = \psi(y)$. Finalement ψ est bien un isomorphisme. ■

DÉFINITION XII.3.2

Soit E un K -ev et F un sous- K -ev de E . Si l'espace quotient E/F n'est pas de dimension finie, on dit que F est **de codimension infinie dans E** . Si E/F est de dimension finie, on dit que F est **de codimension finie dans E** ; l'entier $\dim_K(E/F)$ s'appelle **codimension de F dans E** , et se note $\text{Codim}_E(F)$.

En combinant le théorème XII.3.6 et le corollaire 1 du théorème IX.4.4, on obtient :

PROPOSITION XII.3.2

Avec les notations de la définition XII.3.2, si E est de dimension finie, on a, pour tout sous- K -ev F de E ,

$$\text{Codim}_E(F) = \dim(E) - \dim(F).$$

En particulier, pour E de dimension finie, les hyperplans de E sont les sous- K -ev de codimension 1 (cf. corollaire 1 du théorème IX.5.1), et cette propriété se généralise :

PROPOSITION XII.3.3

Dans un K -ev quelconque non nul, les hyperplans sont les sous- K -ev de codimension 1.

Démonstration :

La définition d'un hyperplan, et le théorème XII.3.3, montrent qu'un sous- K -ev H de E est un hyperplan ssi l'espace quotient $Q = E/H$ est non nul et ne possède pour sous-espaces que $\{0_Q\}$ et Q , ce qui signifie que Q est de dimension 1. ■

Supplémentaires d'un sous-espace de codimension finie

PROPOSITION XII.3.4

Soit E un K -ev et F un sous- K -ev de E de codimension finie égale à $p \geq 1$. Pour toute base (V_1, \dots, V_p) de E/F et pour tout système $(x_1, \dots, x_p) \in E^p$ tel que $\varphi(x_i) = V_i$ (où $\varphi : E \rightarrow E/F$ est l'application canonique), le sous- K -ev $G = \text{Vect}(x_1, \dots, x_p)$ est un supplémentaire de F dans E .

Démonstration :

Soit $x \in E$. On a $\varphi(x) = \lambda_1 V_1 + \dots + \lambda_p V_p$ avec $(\lambda_1, \dots, \lambda_p) \in K^p$; d'où $\varphi\left(x - \sum_{i=1}^p \lambda_i x_i\right) = 0$ et par suite :

$$x = \left(\sum_{i=1}^p \lambda_i x_i\right) + y, \quad \text{avec } y \in F, \text{ donc } E = F + G.$$

Soit $x \in F \cap G$, d'où $x = \sum_{i=1}^p \lambda_i x_i$ avec $(\lambda_1, \dots, \lambda_p) \in K^p$ et $\varphi(x) = \sum_{i=1}^p \lambda_i V_i = 0$, d'où $\lambda_1 = \dots = \lambda_p = 0$. Alors $x = 0$, d'où $E = F \oplus G$. ■

Cette proposition établit donc l'existence d'un supplémentaire pour tout sous- K -ev de codimension finie (le cas particulier des hyperplans avait déjà été vu au § IX.5). En réalité, dans tout K -ev E , tout sous- K -ev admet toujours un supplémentaire, mais la démonstration générale sort du cadre de ce livre.

Exercice 1 : Soit E et E' deux K -ev et F, F' des sous- K -ev de codimension finie respectivement de E et de E' . On suppose $\text{Codim}_E(F) = \text{Codim}_{E'}(F')$. Montrer que $F \times E'$ et $E \times F'$ sont isomorphes.

Exercice 2 : Soit E, F, G trois K -ev, $u \in \text{Hom}_K(E, F)$ et $v \in \text{Hom}_K(F, G)$. On suppose F de dimension finie. Démontrer : $\text{rg}(v) = \text{rg}(v \circ u) + \text{Codim}_F(\text{Im}(u) + \text{Ker}(v))$.

Exercice 3 : Reprendre l'exercice 5 du § XII.1 en y supposant E de dimension finie.

Exercice 4 : Si E est un K -ev de dimension finie $n \geq 1$, l'application $u \mapsto {}^t u^{-1}$, $\text{GL}_K(E) \rightarrow \text{GL}_K(E^*)$ est un isomorphisme de groupes.

Exercice 5 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. On pose $\mathcal{F} = \{v \in \text{Hom}_K(E) \mid v \circ u = u \circ v = 0\}$.

- \mathcal{F} est un sous- K -ev de $\text{Hom}_K(E)$ isomorphe à $\text{Hom}_K(E/\text{Im}(u), \text{Ker}(u))$.
- Calculer $\dim_K(\mathcal{F})$ en fonction de n et de $r = \text{rg}(u)$.

Exercice 6 : Soit F_1 (resp. F_2) un sous- K -ev du K -ev E_1 (resp. E_2) et

$$\varphi_i : E_i \rightarrow E_i/F_i \quad (i \in \{1, 2\}) \quad \text{l'application canonique.}$$

On note \mathcal{F} l'ensemble $\{u \in \text{Hom}_K(E_1, E_2) \mid u(F_1) \subset F_2\}$.

- \mathcal{F} est un sous- K -ev de $\text{Hom}_K(E_1, E_2)$.
- Si $u \in \mathcal{F}$, montrer qu'il existe un unique $\bar{u} \in \text{Hom}_K(E_1/F_1, E_2/F_2)$ tel que $\bar{u} \circ \varphi_1 = \varphi_2 \circ u$. Étudier l'application linéaire $u \mapsto \bar{u}$ (noyau, image).
- Si E_1 et E_2 sont de dimension finie, calculer $\dim(\mathcal{F})$ en fonction des dimensions de E_1, E_2, F_1 et F_2 .

Exercice 7 : Soit F un sous- K -ev d'un K -ev E , et $\varphi : E^* \rightarrow F^*$, $\varphi \mapsto \varphi/F$. Vérifier que $\text{Ker}(\rho) = F^0$. Montrer que si on suppose connu un supplémentaire G de F dans E , alors ρ est surjective, et ρ , par passage au quotient, définit un isomorphisme entre E^*/F^0 et F^* . Conclure si F est de codimension finie dans E .

Exercice 8 : Soit E un K -ev de dimension finie $n \geq 1$, et V_1, \dots, V_r des sous- K -ev de codimensions μ_1, \dots, μ_r ($r \leq n$).

a) Montrer d'abord que $\text{Codim}(\cap V_i) \leq \sum \mu_i$.

b) Prouver ensuite l'équivalence des quatre conditions suivantes :

$$(I) \quad \forall i \in \llbracket 1, r \rrbracket, \quad \text{Codim}_E \left(\bigcap_{j \in I} V_j \right) = \sum_{j \in I} \mu_j.$$

(II) Les $(V_j^0)_{j \in \llbracket 1, r \rrbracket}$ sont linéairement indépendants.

(III) Il existe une décomposition $E = F \oplus W_1 \oplus \dots \oplus W_r$ telle que $(\forall j) V_j = F \oplus \left(\bigoplus_{k \neq j} W_k \right)$.

$$(IV) \quad \text{Si } F = \bigcap_{j=1}^r V_j, \text{ alors } \text{Codim}_E(F) = \sum_{j=1}^r \mu_j.$$

c) Cas où les V_j sont des hyperplans ?

Exercice 9 : Soit E, F, G trois K -ev. On donne $u \in \text{Hom}_K(E, F)$ et $v \in \text{Hom}_K(F, G)$ tels que $\text{Im}(u)$ et $\text{Im}(v)$ soient de codimension finie dans F et G respectivement. Montrer que $\text{Im}(v \circ u)$ est de codimension finie dans G et que :

$$\text{Codim}_G(\text{Im}(v \circ u)) \leq \text{Codim}_F(\text{Im}(u)) + \text{Codim}_G(\text{Im}(v)).$$

Exercice 10 : Soit E et F deux K -ev et $u \in \text{Hom}_K(E, F)$. On suppose connu un supplémentaire de $\text{Im}(u)$ dans F . Montrer que $\text{Im}(u) = \text{Ker}(u)^0$. Conclure si $\text{Im}(u)$ est de codimension finie dans F .

Exercice 11 : (cas particulier du « lemme de Zassenhaus »).

Soit V, V', W, W' quatre sous- K -ev d'un K -ev E , tels que $V' \subset V, W' \subset W$. Démontrer que les trois K -ev quotients

$$Q = V \cap W / (V \cap W' + V' \cap W), \quad Q' = (V' + V \cap W) / (V' + V \cap W')$$

et $Q'' = (W' + V \cap W) / (W' + V' \cap W)$ sont canoniquement isomorphes.

Indication : Pour construire par exemple un isomorphisme de Q sur Q' , soit $\varphi' : V' + V \cap W \rightarrow Q'$ l'application canonique, $\psi' : V \cap W \rightarrow V' + V \cap W'$ l'injection canonique, et $\zeta' = \varphi' \circ \psi'$. Montrer que ζ' est surjective, que son noyau est $V' + V \cap W'$, et passer au quotient.

Exercice 12 : Soit E et F deux K -ev et $u \in \text{Hom}_K(E, F)$. On dit que u est *fini* ssi $\text{Ker}(u)$ et $F/\text{Im}(u)$ sont de dimension finie. On pose alors

$$D(u) = \dim(\text{Ker}(u)) - \text{Codim}_F(\text{Im}(u)).$$

Soit G un autre K -ev et $v \in \text{Hom}_K(F, G)$. Si deux des applications $u, v, v \circ u$ sont finies, la troisième l'est aussi, et alors on a : $D(v \circ u) = D(u) + D(v)$.

Indication : On prouvera que E peut s'écrire $E = \text{Ker}(u) \oplus V \oplus W$, avec

$$\text{Ker}(v \circ u) = V \oplus \text{Ker}(u) \quad \text{et} \quad \text{Im}(v \circ u) = v(u(W)).$$

On admettra que tout sous- K -ev d'un K -ev admet au moins un supplémentaire.

Exercice 13 : Donner un exemple, à l'aide d'un K -ev E convenable :

a) d'un endomorphisme u de E tel que $\text{Im}(u)$ soit de codimension infinie dans E bien que u soit injectif ;

b) d'un endomorphisme *surjectif* u de E dont le noyau est de dimension infinie.

Exercice 14 : Soit V, E deux K -ev, et F un sous- K -ev de E . On note $\varphi : E \rightarrow E/F$ l'application canonique. On suppose connu un supplémentaire de F dans E . En étudiant l'application

$$\text{Hom}_K(V, E) \rightarrow \text{Hom}_K(V, E/F), \quad u \mapsto \varphi \circ u,$$

montrer que les K -ev $\text{Hom}_K(V, E)/\text{Hom}_K(V, F)$ et $\text{Hom}_K(V, E/F)$ sont isomorphes. Conclure si F est de codimension finie dans E .

Exercice 15 : On reprend les hypothèses et notations de l'exercice 14. Soit

$$\rho : \text{Hom}_K(E, V) \rightarrow \text{Hom}_K(F, V), \quad u \mapsto u|_F.$$

Montrer qu'il y a un isomorphisme naturel entre $\text{Ker}(\rho)$ et $\text{Hom}_K(E/F, V)$. Conclure si F est de codimension finie dans E .

Exercice 16 : Soit E et E' deux K -ev et F (resp. F') un sous- K -ev de E (resp. E'). On note $\varphi' : E' \rightarrow E'/F'$ l'application canonique. On suppose connus des supplémentaires de F dans E et de F' dans E' .

Soit $\alpha : \text{Hom}_K(E, E') \rightarrow \text{Hom}_K(F, E'/F')$, $u \mapsto \varphi' \circ (u|_F)$. Étudier le noyau et l'image de α . Montrer que $\text{Ker}(\alpha)$ est isomorphe à $\mathcal{E}/\text{Hom}_K(E, F')$, où

$$\mathcal{E} = \{u \in \text{Hom}_K(E, E') \mid u(F) \subset F'\} \quad (\text{cf. exercice 6}).$$

Conclure si F et F' sont de codimension finie dans E et E' respectivement.

Exercice 17 : Soit E un K -ev et F un sous- K -ev de E de codimension finie dans E . On donne $u \in \text{GL}_K(E)$ tel que $u(F) \subset F$. Montrer que $u(F) = F$. Donner un exemple où cette propriété tombe en défaut lorsque F n'est pas de codimension finie.

Exercice 18 : Soit E un K -ev de dimension finie $n \geq 2$ et F un sous- K -ev strict et non nul de E . On pose $p = \dim(F)$ ($1 \leq p \leq n-1$) et $q = n-p$.

a) L'ensemble $\mathcal{L}_F = \{u \in \text{Hom}_K(E) \mid u(F) \subset F\}$ est une sous- K -algèbre de $\text{Hom}_K(E)$.

b) Soit $\varphi : E \rightarrow E/F$ l'application canonique. Pour $u \in \mathcal{L}_F$, on a un

$$\bar{u} \in \text{Hom}_K(E/F)$$

et un seul tel que $\bar{u} \circ \varphi = \varphi \circ u$ (cf. exercice 6). L'application $\rho : u \mapsto \bar{u}$ est un homomorphisme de K -algèbres $\mathcal{L}_F \rightarrow \text{Hom}_K(E/F)$, dont le noyau est $\text{Hom}_K(E, F) = \mathfrak{b}$ (c'est donc un idéal bilatère de \mathcal{L}_F). De plus, ρ est surjectif.

c) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E telle que (e_1, \dots, e_p) soit une base de F . On pose $V_j = \varphi(e_{p+j})$ ($1 \leq j \leq q$). Vérifier que $\mathcal{C} = (V_1, \dots, V_q)$ est une base de E/F . Si $u \in \mathcal{L}_F$ admet pour matrice $M = [a_{ij}]$ dans la base \mathcal{B} , montrer que le bloc $\mathcal{M}_{[p+1, n], [p+1, n]}(M)$ est la matrice de \bar{u} dans la base \mathcal{C} . (On notera que le bloc $\mathcal{M}_{[1, p], [1, p]}(M)$ n'est autre que $\text{Mat}_{(e_1, \dots, e_p)}(u|_F)$.)

d) Le groupe G des éléments inversibles de l'algèbre \mathcal{L}_F est $\mathcal{L}_F \cap \text{GL}_K(E)$. À l'aide de ce qui précède, montrer qu'à chaque base \mathcal{B} de E du type envisagé en c) est associée une bijection

$$f_{\mathcal{B}} : \text{Ker}(\tau) \rightarrow \text{GL}(p, K) \times K^{pq},$$

τ étant l'application de G dans $\text{GL}_E(E/F)$ telle que $\tau(u) = \bar{u}$ dont on vérifiera que c'est un homomorphisme surjectif.

§ XII.4 QUOTIENTS, PRODUITS ET SOMMES DIRECTES

Nous allons étudier ici le comportement du quotient d'espaces vectoriels relativement à la formation des produits et sommes directes externes de K -ev (revoir le § IX.2).

THÉORÈME XII.4.1

Soit $(E_i)_{i \in I}$ une famille non vide de K -ev et pour chaque $i \in I$, soit F_i un sous- K -ev de E_i . On a alors des isomorphismes naturels de K -ev entre :

$$\left\| \begin{array}{l} \text{(I)} \quad \prod_{i \in I} (E_i/F_i) \quad \text{et} \quad \left(\prod_{i \in I} E_i \right) / \left(\prod_{i \in I} F_i \right) \\ \text{(II)} \quad \coprod_{i \in I} (E_i/F_i) \quad \text{et} \quad \left(\coprod_{i \in I} E_i \right) / \left(\coprod_{i \in I} F_i \right). \end{array} \right.$$

Démonstration :

Notons pour abréger $P = \prod_{i \in I} E_i$, $Q = \prod_{i \in I} F_i$, $S = \coprod_{i \in I} E_i$ et $T = \coprod_{i \in I} F_i$. Soit également $\varphi_i : E_i \rightarrow E_i/F_i$ l'application canonique ($i \in I$).

Définissons $f : P \rightarrow \prod_{i \in I} (E_i/F_i)$, $(x_i)_{i \in I} \mapsto (\varphi_i(x_i))_{i \in I}$: f est linéaire, surjective, et son noyau est Q , donc par passage au quotient f définit un isomorphisme de P/Q sur $\prod_{i \in I} (E_i/F_i)$, d'où l'assertion (I).

Soit $g_1 = f|_S$; alors g_1 est à valeurs dans $\prod_{i \in I} (E_i/F_i) = V$, et $g = g_1|_T^V$ est surjective. On a d'autre part :

$$\text{Ker}(g) = \text{Ker}(f) \cap S = T,$$

d'où g définit, par passage au quotient, un isomorphisme de S/T sur V . ■

En revenant au théorème IX.2.3, (I), on en déduit notamment :

COROLLAIRE

*Supposons que le K -ev E soit **somme directe interne** d'une famille $(E_i)_{i \in I}$ de sous- K -ev : $E = \bigoplus_{i \in I} E_i$. Pour chaque $i \in I$, soit F_i un sous- K -ev de E_i , alors les espaces*

$$E / \bigoplus_{i \in I} F_i \quad \text{et} \quad \prod_{i \in I} (E_i/F_i)$$

sont canoniquement isomorphes.

Le théorème XII.4.1 et son corollaire seront, par la suite, principalement utilisés quand I est fini. Si par exemple $I = \llbracket 1, n \rrbracket$, ($n \geq 1$), il donne les résultats suivants :

• Si E_1, \dots, E_n sont des K -ev et F_1, \dots, F_n des sous- K -ev respectifs, alors les espaces $(E_1 \times E_2 \times \dots \times E_n) / (F_1 \times F_2 \times \dots \times F_n)$ et $E_1/F_1 \times E_2/F_2 \times \dots \times E_n/F_n$ sont canoniquement isomorphes.

- Si $E = \bigoplus_{i=1}^n E_i$, et si, pour tout i , F_i est un sous- K -ev de E_i , alors

$E / \bigoplus_{i=1}^n F_i$ et $E_1/F_1 \times \cdots \times E_n/F_n$ sont canoniquement isomorphes.

Application à un théorème de dualité

LEMME 1

|| Soit E un K -ev et F un sous- K -ev de E . Les espaces $(E/F)^*$ et F^0 sont canoniquement isomorphes.

Démonstration :

Soit $\varphi : E \rightarrow E/F$ l'application canonique. L'application

$${}'\varphi : (E/F)^* \rightarrow E^*, \quad \alpha \mapsto \alpha \circ \varphi$$

est injective, car φ est surjective, et son image est F^0 à cause du théorème XII.3.1. Donc $'\varphi$ définit un isomorphisme de $(E/F)^*$ sur F^0 . ■

THÉORÈME XII.4.2

|| Soit E un K -ev non nul, et soit H un sous- K -ev de E^* de dimension finie égale à $p \geq 1$. Alors

(I) 0H est de codimension finie dans E , et on a :

$$\text{Codim}_E ({}^0H) = p .$$

(II) On a : $({}^0H)^0 = H$.

Démonstration :

Soit $(\varphi_1, \dots, \varphi_p)$ une base de H . Définissons $\Phi : E \rightarrow K^p$, $x \mapsto (\varphi_1(x), \dots, \varphi_p(x))$. Alors, Φ est surjective (théorème XII.2.3), et

$$\text{Ker} (\Phi) = \bigcap_{i=1}^p \text{Ker} (\varphi_i) = {}^0H .$$

Donc, par passage au quotient, Φ définit un isomorphisme $\Psi : E/{}^0H \rightarrow K^p$, ce qui prouve (I). D'après le lemme 1, $({}^0H)^0$ est isomorphe à $(E/{}^0H)^*$, donc à $(K^p)^*$, donc

$$\dim ({}^0H)^0 = p = \dim (H) ;$$

et comme par ailleurs $H \subset ({}^0H)^0$, il s'ensuit que $({}^0H)^0 = H$, d'où (II). ■

Le théorème XII.4.2 n'apporte du nouveau que si E est de dimension infinie. Concrètement, on l'utilise sous la forme suivante : si une forme linéaire Ψ sur un K -ev E s'annule en tout zéro commun à n formes linéaires ψ_1, \dots, ψ_n sur E , alors $\Psi \in \text{Vect} (\psi_1, \dots, \psi_n)$.

L'exercice 4 du § XII.1 montre un exemple où le théorème XII.4.2 est en défaut lorsque H n'est pas de dimension finie.

Exercice 1 : Soit E, E' deux K -ev, et des sous- K -ev $(E_i)_{1 \leq i \leq n}, (F_i)_{1 \leq i \leq n}$, (resp. $(E'_j)_{1 \leq j \leq p}, (F'_j)_{1 \leq j \leq p}$) de E (resp. de E'), tels que $(\forall i) F_i \subset E_i, (\forall j) F'_j \subset E'_j$, $E = \bigoplus_{i=1}^n E_i$ et $E' = \bigoplus_{j=1}^p E'_j$. On posera $F = \bigoplus_{i=1}^n F_i$ et $F' = \bigoplus_{j=1}^p F'_j$.

a) Démontrer que le K -ev $\text{Hom}_K(E/F, E'/F')$ est canoniquement isomorphe à $\bigoplus_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket} \text{Hom}_K(E_i/F_i, E'_j/F'_j)$.

b) On suppose $n = p$. Soit

$$\mathcal{E} = \{u \in \text{Hom}_K(E, E') \mid (\forall i) u(E_i) \subset E'_i \text{ et } u(F_i) \subset F'_i\}.$$

Vérifier que \mathcal{E} est un sous- K -ev de $\text{Hom}_K(E, E')$. Montrer, à l'aide de l'exercice 6 du § XII.3, qu'on a une application linéaire naturelle $\alpha : \mathcal{E} \rightarrow \bigtimes_{i=1}^n \text{Hom}_K(E_i/F_i, E'_i/F'_i)$. Étudier

$\text{Im}(\alpha)$ et $\text{Ker}(\alpha)$.

Exercice 2 : Soit E un K -ev, et deux sous- K -ev V et W de E . On note $\varphi : E/V \cap W \rightarrow E/V$ et $\psi : E/V \cap W \rightarrow E/W$ les applications naturelles (obtenues à partir des applications canoniques $E \rightarrow E/V$ et $E \rightarrow E/W$ par application du théorème XII.3.1). On définit de la même manière les applications canoniques $\alpha : E/V \rightarrow E/V + W$ et $\beta : E/W \rightarrow E/V + W$. Soit

$$u : E/V \cap W \rightarrow E/V \times E/W, \quad x \mapsto (\varphi(x), \psi(x))$$

et
$$v : E/V \times E/W \rightarrow E/V + W, \quad (y, z) \mapsto \alpha(y) - \beta(z).$$

Démontrer : u est injective, v est surjective, et $\text{Im}(u) = \text{Ker}(v)$. En déduire que, si V et W sont de codimension finie, $V \cap W$ l'est encore (ainsi, bien sûr, que $V + W$), et qu'on a :

$$\text{Codim}_E(V) + \text{Codim}_E(W) = \text{Codim}_E(V \cap W) + \text{Codim}_E(V + W).$$

Chapitre XIII

DÉTERMINANTS

Dans tout le chapitre, K désigne un corps commutatif.

§ XIII.1 APPLICATIONS MULTILINÉAIRES

DÉFINITION XIII.1.1

Soit E_1, \dots, E_n et F des K -ev ($n \geq 1$). Une application $f : E_1 \times \dots \times E_n \rightarrow F$ est dite **multilinéaire** ssi elle est linéaire par rapport à chaque facteur, autrement dit, pour tout $i \in \llbracket 1, n \rrbracket$, et tout $a = (a_j)_{j \in \llbracket 1, n \rrbracket \setminus \{i\}}$, l'application $E_i \rightarrow F$, $x_i \mapsto f(t_1, \dots, t_n)$ (où $t_j = a_j$ pour $j \neq i$, et $t_i = x_i$) est K -linéaire. Lorsque $F = K$, on parle de **forme multilinéaire**.

Pour $n = 1$, on retrouve la notion d'application linéaire ; pour $n = 2$, on parle d'application bilinéaire ; pour $n = 3$, trilinéaire ; pour n quelconque, n -linéaire.

Les propriétés suivantes se prouvent par une simple vérification, qui sera laissée au lecteur.

(ML1) L'ensemble (que nous noterons $ML(E_1, \dots, E_n; F)$) des applications multilinéaires $E_1 \times \dots \times E_n \rightarrow F$ forme un sous- K -ev du K -ev $\mathcal{F}(E_1 \times \dots \times E_n, F)$ des fonctions de $E_1 \times \dots \times E_n$ dans F .

Lorsque $E_1 = \dots = E_n = E$, ce K -ev sera noté en abrégé $ML_n(E, F)$ et si de plus $F = K$, l'ensemble des formes n -linéaires sur E sera noté $ML_n(E)$.

(ML2) Si $f \in ML(E_1, \dots, E_n; F)$ et si $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ est tel que $\exists i \mid x_i = 0_{E_i}$, alors $f(x_1, \dots, x_n) = 0$.

En conséquence, si l'un des K -ev E_i est nul, $ML(E_1, \dots, E_n; F)$ est nul.

(ML3) Si $F = F_1 \times F_2 \times \dots \times F_p$, pour tout système (f_1, f_2, \dots, f_p) , avec $f_i \in ML(E_1, \dots, E_n; F_i)$ pour tout i , l'application $f : E_1 \times \dots \times E_n \rightarrow F$, $x \mapsto (f_1(x), \dots, f_p(x))$ appartient à $ML(E_1, \dots, E_n; F)$, et l'application $(f_1, \dots, f_p) \mapsto f$ est un isomorphisme de K -ev entre

$$\bigtimes_{i=1}^p ML(E_1, \dots, E_n; F_i) \quad \text{et} \quad ML(E_1, \dots, E_n; F)$$

(ML4) Soit G un K -ev. Si $f \in \text{ML}(E_1, \dots, E_n; F)$ et si $u \in \text{Hom}_K(F, G)$, alors $u \circ f \in \text{ML}(E_1, \dots, E_n; G)$.

(ML5) Soit D_1, \dots, D_n des K -ev et, pour tout $i \in \llbracket 1, n \rrbracket$, soit $u_i \in \text{Hom}_K(D_i, E_i)$. Si $f \in \text{ML}(E_1, \dots, E_n; F)$, l'application

$$D_1 \times \dots \times D_n \rightarrow F, \quad (t_1, \dots, t_n) \mapsto f(u_1(t_1), \dots, u_n(t_n))$$

est multilinéaire.

Supposons F de dimension finie $p \geq 1$. Munissons-le d'une base $\mathcal{B} = (e_1, \dots, e_p)$ qui définit un isomorphisme de K -ev $u : F \rightarrow K^p$, $y \mapsto (y_1, \dots, y_p)$ (système des coordonnées du vecteur y dans la base \mathcal{B}). En appliquant (ML4) avec u , puis (ML3), on voit que \mathcal{B} définit un isomorphisme entre $\text{ML}(E_1, \dots, E_n; F)$ et $\bigtimes_{i=1}^p \text{ML}(E_1, \dots, E_n; K)$. Ainsi,

l'étude des applications multilinéaires à valeurs dans un K -ev de dimension finie se ramène à celle des formes multilinéaires. C'est pourquoi nous nous limiterons à l'étude des *formes multilinéaires*.

THÉORÈME XIII.1.1

|| Supposons chaque K -ev E_i de dimension finie d_i ($n \geq 1$). Alors le K -ev $\text{ML}(E_1, \dots, E_n; K)$ est de dimension finie ; sa dimension est $d_1 d_2 \dots d_n$.

Démonstration :

Choisissons une base $\mathcal{B}_i = (e_{i1}, \dots, e_{id_i})$ pour chaque E_i ($1 \leq i \leq n$) et cherchons la forme nécessaire de $f \in \text{ML}(E_1, \dots, E_n; K)$. Fixons une telle f . Pour tous vecteurs $x_1 \in E_1, \dots, x_n \in E_n$, écrivons :

$$x_i = \sum_{k=1}^{d_i} x_{ki} e_{ik}, \quad \text{où les } (x_{ki})_{1 \leq k \leq d_i} \text{ sont les coordonnées de } x_i \text{ dans } \mathcal{B}_i.$$

La linéarité de f en chaque x_i nous permet de développer $f(x_1, \dots, x_n)$ par multilinéarité :

$$(1) \quad f(x_1, \dots, x_n) = \sum_{(k_1, \dots, k_n) \in \prod_{i=1}^n \llbracket 1, d_i \rrbracket} x_{k_1, 1} x_{k_1, 2} \dots x_{k_n, n} f(e_{1, k_1}, \dots, e_{n, k_n}).$$

Pour chaque élément $\mathbf{k} = (k_1, \dots, k_n) \in I = \prod_{i=1}^n \llbracket 1, d_i \rrbracket$, notons $g_{\mathbf{k}}$ la

fonction $E_1 \times \dots \times E_n \rightarrow K$, $(x_1, \dots, x_n) \mapsto x_{k_1, 1} \dots x_{k_n, n} = \prod_{i=1}^n x_{k_i, i}$.

La famille $(g_{\mathbf{k}})_{\mathbf{k} \in I}$ engendre un sous- K -ev \mathcal{V} du K -ev de toutes les fonctions de $E_1 \times \dots \times E_n$ dans K . La relation (1) nous prouve l'inclusion $\text{ML}(E_1, \dots, E_n; K) \subset \mathcal{V}$.

Maintenant étudions les g_k : on vérifie facilement que chaque g_k est multilinéaire, d'où en fait l'égalité :

$$(2) \quad \text{ML}(E_1, \dots, E_n; K) = \mathcal{V},$$

et la famille (g_k) est *génératrice* du K -ev $\text{ML}(E_1, \dots, E_n; K)$.

Pour prouver le théorème il suffit donc d'établir que $(g_k)_{k \in I}$ est une famille *libre*, puisque $\text{card}(I) = d_1 d_2 \dots d_n$.

Or, soit $(\lambda_k) \in K^I$ tel que

$$(3) \quad \sum_{k \in I} \lambda_k g_k = 0.$$

Fixons $l = (l_1, \dots, l_n) \in I$. On contrôle que, pour tout $k = (k_1, \dots, k_n) \in I$, on a :

$$g_k(e_{1,l_1}, \dots, e_{n,l_n}) = \delta_{k_1,l_1} \delta_{k_2,l_2} \dots \delta_{k_n,l_n} \cdot 1_K$$

(où δ est le symbole de Kronecker). Donc, prenant la valeur du premier membre de (3) sur la suite $(e_{1,l_1}, \dots, e_{n,l_n})$, on obtient $\lambda_l = 0$, et c'est vrai pour tout l . Finalement $(g_k)_{k \in I}$ est une base du K -ev

$$\text{ML}(E_1, \dots, E_n; K). \quad \blacksquare$$

Naturellement, pour $n = 1$, on retrouve le théorème IX.4.6.

Remarque 1 : En juxtaposant les bases \mathcal{B}_i , on obtient une base \mathcal{B} de $E_1 \times \dots \times E_n$. L'examen de la relation (1) et la preuve ci-dessus montrent que toute forme multilinéaire sur $E_1 \times \dots \times E_n$ est une *fonction polynomiale particulière* sur cet espace produit.

Dérivation d'une forme multilinéaire

Conservons les notations précédentes, en supposant les E_i de dimension finie, et K de *caractéristique nulle*. Soit I une partie non vide de K et, pour chaque $i \in \llbracket 1, n \rrbracket$, $\varphi_i : I \rightarrow E_i$ une fonction *rationnelle* (c.-à-d. dont les coordonnées dans une base de E_i sont rationnelles, cette notion ne dépendant pas du choix de la base). Notons φ'_i la dérivée de φ_i (c'est la fonction dont les coordonnées dans une base de E_i sont définies par les dérivées dans $K(X)$ de celles de φ_i , cette fonction étant, elle aussi, indépendante du choix de la base). Enfin, soit $f \in \text{ML}(E_1, \dots, E_n; K)$. Nous allons montrer :

THÉORÈME XIII.1.2

Avec les notations ci-dessus, K désignant un corps commutatif de caractéristique 0, la fonction

$$\psi : I \rightarrow K, \quad t \mapsto f(\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t))$$

est rationnelle, et on a :

$$\psi'(t) = \sum_{k=1}^n f(\varphi_1(t), \dots, \varphi_{k-1}(t), \varphi'_k(t), \varphi_{k+1}(t), \dots, \varphi_n(t)) .$$

Démonstration :

On peut écrire $\varphi_i(t) = \frac{1}{D(t)} \Phi_i(t)$, où Φ_i est polynomiale pour tout i , et où D est polynomiale (Φ_i à valeurs dans E_i , D à valeurs dans K). D'où

$$\psi(t) = \frac{1}{(D(t))^n} \Psi(t) , \quad \text{avec} \quad \Psi(t) = f(\Phi_1(t), \dots, \Phi_n(t)) .$$

On en déduit immédiatement qu'il suffit de prouver le théorème avec des φ_i polynomiales (*i.e.* dont les coordonnées dans les diverses bases de E_i sont polynomiales), et on peut alors supposer les φ_i définies sur K , ce que nous ferons désormais. La remarque 1 montre d'abord que ψ est polynomiale. Fixons $a \in K$. On a donc

$$\psi(t) = \psi(a) + (t-a) \psi'(a) + (t-a)^2 P(t) ,$$

où P est polynomiale, d'après la formule de Taylor. De même, pour chaque i ,

$$\varphi_i(t) = \varphi_i(a) + (t-a) \varphi'_i(a) + (t-a)^2 P_i(t) ,$$

avec P_i polynomiale, d'où en développant par n -linéarité :

$$\begin{aligned} f[\varphi_1(t), \dots, \varphi_n(t)] &= f[\varphi_1(a), \dots, \varphi_n(a)] + (t-a) \times \\ &\times \sum_{k=1}^n f[\varphi_1(a), \dots, \varphi_{k-1}(a), \varphi'_k(a), \varphi_{k+1}(a), \dots, \varphi_n(a)] + (t-a)^2 g(t) , \end{aligned}$$

avec g polynomiale. Par identification, on en déduit :

$$\psi'(a) = \sum_{k=1}^n f[\varphi_1(a), \dots, \varphi_{k-1}(a), \varphi'_k(a), \varphi_{k+1}(a), \dots, \varphi_n(a)] ,$$

et c'est vrai pour tout $a \in K$, d'où le théorème. ■

Action de \mathfrak{S}_n sur $\text{ML}_n(E)$

Fixons $n \in \mathbb{N}^*$ et soit E un K -ev. Pour chaque $f \in \text{ML}_n(E)$ et chaque $\sigma \in \mathfrak{S}_n$, l'application $\sigma * f : E^n \rightarrow K$, $(x_1, \dots, x_n) \mapsto f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ est évidemment n -linéaire. Il est clair que $\text{Id} * f = f$. De plus, soit $\sigma \in \mathfrak{S}_n$, $\tau \in \mathfrak{S}_n$ et $(x_1, \dots, x_n) \in E^n$. On a :

$$[(\tau\sigma) * f](x_1, \dots, x_n) = f(x_{\tau\sigma(1)}, \dots, x_{\tau\sigma(n)}) ,$$

et :

$$[\tau * (\sigma * f)](x_1, \dots, x_n) = (\sigma * f)(x_{\tau(1)}, \dots, x_{\tau(n)}) = f(x_{\tau\sigma(1)}, \dots, x_{\tau\sigma(n)}),$$

d'où

$$(4) \quad \tau * (\sigma * f) = (\tau\sigma) * f.$$

De plus, pour $\sigma \in \mathfrak{S}_n$ donné, $f \mapsto \sigma * f$, $\text{ML}_n(E) \rightarrow \text{ML}_n(E)$ est K -linéaire. En résumé, on a prouvé :

PROPOSITION XIII.1.1

|| L'application $\mathfrak{S}_n \times \text{ML}_n(E) \rightarrow \text{ML}_n(E)$, $(\sigma, f) \mapsto \sigma * f$ est une action à gauche de \mathfrak{S}_n sur $\text{ML}_n(E)$. Cette action est K -linéaire, i.e. pour tout $\sigma \in \mathfrak{S}_n$, la bijection $\text{ML}_n(E) \rightarrow \text{ML}_n(E)$, $f \mapsto \sigma * f$ est linéaire.

Les types les plus simples, et les plus intéressants, de formes n -linéaires sur E sont ceux donnés par la définition suivante :

DÉFINITION XIII.1.2

⎧ Soit E un K -ev et n un entier ≥ 1 . Une forme n -linéaire
 ⎧ $f \in \text{ML}_n(E)$ est dite :
 ⎧ **symétrique** ssi $\sigma * f = f$ pour tout $\sigma \in \mathfrak{S}_n$,
 ⎧ **antisymétrique** ssi $\sigma * f = \varepsilon(\sigma) f$ pour tout $\sigma \in \mathfrak{S}_n$,
 ⎧ **alternée** ssi $f(x_1, \dots, x_n) = 0$ pour toute suite $(x_1, \dots, x_n) \in E^n$ telle
 ⎧ que l'application $i \mapsto x_i$ soit **non injective**.

Nous noterons $S_n(E)$ (resp. $A_n(E)$, $\Lambda_n^*(E)$) l'ensemble des formes n -linéaires sur E qui sont symétriques (resp. antisymétriques, alternées). En raison de la proposition XIII.1.1, ces ensembles sont des sous- K -ev de $\text{ML}_n(E)$.

Exemple 1 : Prenons $E = K^2$ et $n = 2$. L'application $E^2 \rightarrow K$, $((x_1, y_1), (x_2, y_2)) \mapsto x_1 y_2 + x_2 y_1$ est bilinéaire symétrique. L'application $((x_1, y_1), (x_2, y_2)) \mapsto x_1 y_2 - x_2 y_1$ est bilinéaire alternée.

Formes antisymétriques et formes alternées

Prenons d'abord une forme n -linéaire *alternée* f sur un K -ev E ($n \geq 1$). Soit $i \in \llbracket 1, n \rrbracket$ et $(x_1, \dots, x_n) \in E^n$. Notons (y_1, \dots, y_n) la suite telle que $y_j = x_j$ si $j \neq i$ et $y_i = x_i + \sum_{j \neq i} c_j x_j$, où les scalaires c_j sont quelconques.

En développant $f(y_1, \dots, y_n)$ par rapport au i -ième facteur, et en tenant compte que f est alternée, on voit que $f(y_1, \dots, y_n) = f(x_1, \dots, x_n)$.

PROPOSITION XIII.1.2

Soit E un K -ev, $n \in \mathbb{N}^*$ et $f \in \Lambda_n^*(E)$. Si l'on rajoute à l'un des vecteurs x_i de la suite $(x_1, \dots, x_n) \in E^n$ une combinaison linéaire arbitraire des $(x_j)_{j \neq i}$, la valeur de f sur la nouvelle suite obtenue est $f(x_1, \dots, x_n)$. En conséquence, si les vecteurs x_i sont liés, on a : $f(x_1, \dots, x_n) = 0$.

(La dernière assertion vient du fait que si les x_i sont liés, l'un d'eux au moins est combinaison linéaire des autres).

PROPOSITION XIII.1.3

Soit E un K -ev et $n \in \mathbb{N}^*$. On a

(I) $\Lambda_n^*(E) \subset A_n(E)$.

Si le corps K est de caractéristique différente de 2, on a :

(II) $\Lambda_n^*(E) = A_n(E)$.

Démonstration :

Pour $n = 1$, on a $\Lambda_n^*(E) = A_n(E) = \text{ML}_n(E) = E^*$. Supposons donc $n \geq 2$ et soit $f \in \Lambda_n^*(E)$. Considérons une suite $(x_1, \dots, x_n) \in E^n$ et fixons i et j dans $[[1, n]]$, $i < j$. Soit $(y_1, \dots, y_n) \in E^n$ telle que $y_k = x_k$ si $k \notin \{i, j\}$ et $y_i = y_j = x_i + x_j$. Alors $f(y_1, \dots, y_n) = 0$ puisqu'un même vecteur figure deux fois dans la suite.

Mais en développant f par multilinéarité en les i -ième et j -ième facteurs, on trouve :

$$f(y_1, \dots, y_n) = f(x_1, \dots, x_n) + f(x_{\tau(1)}, \dots, x_{\tau(n)}) + f(u_1, \dots, u_n) + f(v_1, \dots, v_n)$$

en notant τ la transposition de i et j dans \mathfrak{S}_n et avec $u_i = u_j = x_i$, $v_i = v_j = x_j$, $u_k = v_k = x_k$ si $k \notin \{i, j\}$. Puisque $f \in \Lambda_n^*(E)$, on a $f(u_1, \dots, u_n) = f(v_1, \dots, v_n) = 0$, donc il reste : $(f + \tau * f)(x_1, \dots, x_n) = 0$, et c'est vrai pour tout $(x_1, \dots, x_n) \in E^n$, d'où $\tau * f = -f$, i.e. $\tau * f = \varepsilon(\tau) f$ puisqu'on sait qu'une transposition est une permutation impaire. C'est vrai avec toutes les transpositions de \mathfrak{S}_n , et puisqu'elles engendrent le groupe \mathfrak{S}_n , on en déduit, à l'aide de la proposition XIII.1.1, que $\sigma * f = \varepsilon(\sigma) f$ pour toute $\sigma \in \mathfrak{S}_n$, i.e. que $f \in A_n$, d'où (I).

Supposons maintenant la caractéristique de K différente de 2, et soit $f \in A_n(E)$. Considérons une suite $(x_1, \dots, x_n) \in E^n$ telle que $x_i = x_j$ (i, j donnés, $1 \leq i < j \leq n$). Si τ est la transposition de i et j dans \mathfrak{S}_n , on a, puisque $\varepsilon(\tau) = -1_K$:

$$(\tau * f)(x_1, \dots, x_n) = f(x_{\tau(1)}, \dots, x_{\tau(n)}) = -f(x_1, \dots,$$

(car $f \in A_n(E)$) et $f(x_{\tau(1)}, \dots, x_{\tau(1)}) = f(x_1, \dots, x_n)$ (car $x_i = x_j$), d'où $2 f(x_1, \dots, x_n) = 0$ et en simplifiant par 2 dans K : $f(x_1, \dots, x_n) = 0$. ■

Exercice 1 : Soit E_1, \dots, E_n des K -ev de dimension finie non nuls, et $f : E_1 \times \dots \times E_n \rightarrow K$, $(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$ une fonction non nulle. Pour que $f \in \text{ML}(E_1, \dots, E_n; K)$, il faut et il suffit que f soit *polynomiale*, et *homogène de degré 1* en chaque x_i .

Exercice 2 : On suppose que $K = \mathbb{Z}/2\mathbb{Z}$. Montrer que la forme trilinéaire définie sur K^2 par $((x_1, y_1), (x_2, y_2), (x_3, y_3)) \mapsto x_1 y_2 y_3 + y_1 x_2 y_3 + y_1 y_2 x_3$ est antisymétrique, mais qu'elle n'est pas alternée.

Exercice 3 : On donne un K -ev E et $n \geq 1$. On suppose K de caractéristique $\neq 2$. Soit $\Gamma_{\mathfrak{U}_n}(E)$ l'ensemble des $f \in \text{ML}_n(E)$ telles que $\sigma * f = f$ pour tout $\sigma \in \mathfrak{U}_n$. Démontrer que $\Gamma_{\mathfrak{U}_n}(E)$ est un sous- K -ev de $\text{ML}_n(E)$, et qu'on a : $\Gamma_{\mathfrak{U}_n}(E) = S_n(E) \oplus \Lambda_n^*(E)$.

Exercice 4 : Soit E un K -ev et $n \in \mathbb{N}^*$.

a) Montrer que si $f \in \text{ML}_n(E)$, l'application $g = a(f)$ définie par $g = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) (\sigma * f)$

est élément de $A_n(E)$.

b) On suppose K de caractéristique nulle. Montrer que l'endomorphisme α de $\text{ML}_n(E)$ tel que $\alpha(f) = \frac{1}{n!} a(f)$ pour tout f , est un *projecteur*, d'image $\Lambda_n^*(E)$ et dont, si $n \geq 2$, le noyau contient $S_n(E)$ ($\alpha(f)$ s'appelle l'*antisymétrisée* de f).

c) Montrer, sous l'hypothèse du b), que l'application $\beta : \text{ML}_n(E) \rightarrow \text{ML}_n(E)$, $f \mapsto \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sigma * f$ est un *projecteur*, d'image $S_n(E)$, et dont le noyau contient $\Lambda_n^*(E)$

($\beta(f)$ s'appelle la *symétrisée* de f).

Exercice 5 : On donne un K -ev et $n \in \mathbb{N}^*$; K est supposé de caractéristique 0. Pour $u \in E$, si f est une application de E dans K , on note $T_u(f)$ l'application $E \rightarrow K$, $x \mapsto f(x + u)$ et $\Delta_u(f)$ l'application $x \mapsto f(x + u) - f(x) = T_u(f) - f$.

a) Soit $s \in S_n(E)$. On note \tilde{s} l'application $E \rightarrow K$, $x \mapsto s(\overbrace{x, x, \dots, x}^{n \text{ fois}})$. Si $(u_1, \dots, u_n) \in E^n$, montrer que l'application $(\Delta_{u_1} \circ \Delta_{u_2} \circ \dots \circ \Delta_{u_n}) \tilde{s} : E \rightarrow K$ est *constante*, sa valeur étant $n! s(u_1, u_2, \dots, u_n)$.

Indication : raisonner par récurrence sur n .

b) En déduire la *formule de restitution* : si $s \in \mathfrak{S}_n(E)$, pour

$$(u_1, \dots, u_n) \in E^n, \quad s(u_1, \dots, u_n) = \frac{1}{n!} [(\Delta_{u_1}) \circ \dots \circ (\Delta_{u_n}) \tilde{s}](0) = \\ = \frac{1}{n!} \sum_{k=0}^n (-1)^{n-k} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \tilde{s}(u_{i_1} + u_{i_2} + \dots + u_{i_k}).$$

(Observer que les Δ_{u_i} (resp. les T_{u_i}) sont deux à deux permutables, et que

$$\Delta_{u_1} \circ \dots \circ \Delta_{u_n} = (T_{u_1} - \text{Id}) \circ \dots \circ (T_{u_n} - \text{Id}),$$

et en déduire que $s \mapsto \tilde{s}$ est *injective*.

c) On suppose maintenant E de dimension finie $N \geq 1$. Montrer que si $s \in S_n(E)$, \tilde{s} est *polynomiale homogène* de degré n sur E . On note $\mathcal{H}_n(E)$ le K -ev des fonctions polynomiales homogènes de degré n sur E . Si $P \in \mathcal{H}_n(E)$, montrer que l'application

$$\hat{P} : E^n \rightarrow K, \quad (u_1, \dots, u_n) \mapsto \frac{1}{n!} \sum_{k=0}^n (-1)^k \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(u_{i_1} + \dots + u_{i_k})$$

est dans $S_n(E)$. En déduire que $s \mapsto \tilde{s}$ est un isomorphisme du K -ev $S_n(E)$ sur $\mathcal{H}_n(E)$, et en particulier, que $\dim S_n(E) = \binom{N+n-1}{N-1}$.

N.B. Si $P \in \mathcal{H}_n(E)$, \hat{P} s'appelle la polarisée de P .

Exercice 6 (suite de l'exercice 5) : E est supposé de dimension finie $N \geq 1$ et on le munit d'une base $\mathcal{B} = (e_1, \dots, e_N)$ et l'on note (x_1, \dots, x_N) sa base duale.

a) Soit $P = \sum_{\substack{\alpha \in \mathbb{N}^N \\ \alpha = (\alpha_1, \dots, \alpha_N)}} c_\alpha x_1^{\alpha_1} \dots x_N^{\alpha_N} \in \mathcal{H}_n(E)$. Montrer que la polarisée \hat{P} de P peut

s'exprimer de la manière suivante : pour chaque $\alpha \in \mathbb{N}^N$ tel que $\|\alpha\| = n$, $\alpha = (\alpha_1, \dots, \alpha_N)$, soit \mathcal{E}_α l'ensemble des applications $\varphi : \llbracket 1, n \rrbracket \rightarrow \llbracket 1, N \rrbracket$ telles que $\text{card}(\varphi^{-1}(i)) = \alpha_i$ pour tout i . Alors, si $(u_1, \dots, u_n) \in E^n$, $u_i = \sum_{j=1}^N u_{j,i} e_j$, on a :

$$\hat{P}(u_1, \dots, u_n) = \frac{1}{n!} \sum_{\alpha \in \mathbb{N}^N, \|\alpha\| = n} \alpha_1! \dots \alpha_N! c_\alpha \left(\sum_{\varphi \in \mathcal{E}_\alpha} u_{\varphi(1),1} \dots u_{\varphi(n),n} \right).$$

Expliciter cette formule pour $n = 2$ et $n = 3$.

b) On donne $p \in \llbracket 1, n-1 \rrbracket$ et on pose $q = n - p$. Soit $x = \sum_{i=1}^N x_i e_i$ et $y = \sum_{i=1}^N y_i e_i$ éléments de E . Démontrer : si $u_i = x$ pour $1 \leq i \leq q$ et $u_i = y$ pour $q+1 \leq i \leq n$, alors

$$n(n-1) \dots (n-p+1) \hat{P}(u_1, \dots, u_n) = \sum_{\alpha \in \mathbb{N}^N, \|\alpha\| = p} \frac{p!}{\alpha_1! \dots \alpha_N!} \frac{\partial^p P(x)}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}} y_1^{\alpha_1} \dots y_N^{\alpha_N}.$$

§ XIII.2 FORMES n -LINÉAIRES ALTERNÉES SUR E DE DIMENSION n

THÉORÈME XIII.2.1

Soit E un K -ev de dimension finie $n \geq 1$ muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$.

(I) Pour $k > n$, on a $\Lambda_k^*(E) = \{0\}$.

(II) Le K -ev $\Lambda_n^*(E)$ est de dimension 1. De plus, notons $\Delta_{\mathcal{B}}$ la fonction $E^n \rightarrow K$ ainsi définie : pour tous $(x_1, \dots, x_n) \in E^n$, si $(x_{ji})_{1 \leq j \leq n}$ est la suite des coordonnées de x_i pour chaque $i \in \llbracket 1, n \rrbracket$

$$\Delta_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1),1} \dots x_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n x_{\sigma(i),i}.$$

Alors $(\Delta_{\mathcal{B}})$ est une base du K -ev $\Lambda_n^*(E)$.

Démonstration :

L'assertion (I) résulte du fait que pour $k > n$ tout système de k vecteurs de E est lié (proposition IX.4). Donc $\wedge^k E = \{0\}$.

(proposition XIII.1.2). Pour démontrer (II) procédons à quelques remarques préliminaires. D'abord, l'ensemble $\llbracket 1, n \rrbracket^n$ s'identifie à l'ensemble I des applications de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, n \rrbracket$, et une telle application est injective ssi elle est bijective, c'est-à-dire ssi c'est un élément de \mathfrak{S}_n .

Ensuite, pour chaque $\varphi \in I$, définissons

$$g_\varphi : E^n \rightarrow E, \quad (x_1, \dots, x_n) \mapsto \prod_{i=1}^n x_{\varphi(i), i}$$

(les notations étant celles de l'énoncé). La démonstration du théorème XIII.1.1 (avec $E_1 = E_2 = \dots = E_n = E$, et $\mathcal{B}_1 = \mathcal{B}_2 = \dots = \mathcal{B}_n = \mathcal{B}$) nous montre que la famille $(g_\varphi)_{\varphi \in I}$ est une base du K -ev $\text{ML}_n(E)$, et que les coordonnées de $f \in \text{ML}_n(E)$ dans cette base sont les

$$(f(e_{\varphi(1)}, \dots, e_{\varphi(n)}))_{\varphi \in I}.$$

Cela dit, cherchons $\Lambda_n^*(E)$ par analyse et synthèse :

Analyse. Soit $f \in \Lambda_n^*(E)$. D'après ce qui précède, on a :

$$(1) \quad f = \sum_{\varphi \in I} f(e_{\varphi(1)}, \dots, e_{\varphi(n)}) g_\varphi$$

soit $\varphi \in I$. Si φ est non injective, c'est-à-dire si $\varphi \notin \mathfrak{S}_n$, on a : $f(e_{\varphi(1)}, \dots, e_{\varphi(n)}) = 0$, car f est alternée. Et si $\varphi \in \mathfrak{S}_n$, on a :

$$f(e_{\varphi(1)}, \dots, e_{\varphi(n)}) = (\varphi * f)(e_1, \dots, e_n) = \varepsilon(\varphi) f(e_1, \dots, e_n),$$

car f étant alternée est antisymétrique (cf. proposition XIII.1.3). Donc, de (1) on déduit, en appelant σ l'élément générique de \mathfrak{S}_n :

$$f = f(e_1, \dots, e_n) \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) g_\sigma = f(e_1, \dots, e_n) \Delta_{\mathcal{B}}.$$

Synthèse. On vérifie immédiatement que $\Delta_{\mathcal{B}} \in \text{ML}_n(E)$. Montrons que $\Delta_{\mathcal{B}}$ est *alternée*. Soit $(x_1, \dots, x_n) \in E^n$ et i, j tels que $1 \leq i < j \leq n$, et $x_i = x_j$ (d'où $x_{k,i} = x_{k,j}$ pour $k \in \llbracket 1, n \rrbracket$). Notons τ la transposition de i et j dans \mathfrak{S}_n . On a $\tau \in \mathfrak{S}_n \setminus \mathcal{U}_n$, donc l'application $\mathcal{U}_n \rightarrow \mathfrak{S}_n \setminus \mathcal{U}_n$, $\sigma \mapsto \tau\sigma$ est bijective, d'où :

$$\begin{aligned} \Delta_{\mathcal{B}}(x_1, \dots, x_n) &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) g_\sigma(x_1, \dots, x_n) = \\ &= \sum_{\sigma \in \mathcal{U}_n} \varepsilon(\sigma) g_\sigma(x_1, \dots, x_n) + \sum_{\sigma \in \mathcal{U}_n} \varepsilon(\sigma\tau) g_{\sigma\tau}(x_1, \dots, x_n). \end{aligned}$$

Mais, si $\sigma \in \mathfrak{U}_n$, on a $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -\varepsilon(\sigma)$ et

$$g_{\sigma\tau}(x_1, \dots, x_n) = \prod_{k=1}^n x_{\sigma\tau(k), k}.$$

Comme $x_{\sigma(i), j} = x_{\sigma(i), i}$ et $x_{\sigma(j), i} = x_{\sigma(j), j}$ et $\tau(k) = k$ pour $k \notin \{i, j\}$, on voit ⁽¹⁾ que $\prod_{k=1}^n x_{\sigma\tau(k), k} = \prod_{k=1}^n x_{\sigma(k), k} = g_{\sigma}(x_1, \dots, x_n)$. Par suite :

$$\Delta_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{U}_n} \varepsilon(\sigma) g_{\sigma}(x_1, \dots, x_n) - \sum_{\sigma \in \mathfrak{U}_n} \varepsilon(\sigma) g_{\sigma}(x_1, \dots, x_n) = 0.$$

Enfin, on vérifie que $\Delta_{\mathcal{B}} \neq 0$, car $\Delta_{\mathcal{B}}(e_1, \dots, e_n) = g_{\text{Id}}(e_1, \dots, e_n) = 1_K$. Donc $\Delta_{\mathcal{B}} \in \Lambda_n^*(E) \setminus \{0\}$, et comme on a vu, dans l'Analyse, que $\Lambda_n^*(E) \subset K\Delta_{\mathcal{B}}$, il s'ensuit bien que $\Lambda_n^*(E) = K\Delta_{\mathcal{B}}$ et que

$$\dim(\Lambda_n^*(E)) = 1. \quad \blacksquare$$

Notons que pour la partie (II) de la démonstration, on n'utilise que la structure d'anneau commutatif de K , et que nulle part l'intégrité de cet anneau n'intervient.

THÉORÈME XIII.2.2

|| Soit E un K -ev de dimension $n \geq 1$. Pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de E et tout $\lambda \in K$, il existe un, et un seul, élément $\varphi \in \Lambda_n^*(E)$ tel que $\varphi(e_1, \dots, e_n) = \lambda$.

Démonstration :

En reprenant les notations du théorème XIII.2.1, on constate que la forme $\varphi = \lambda \Delta_{\mathcal{B}}$ répond à la question. Soit $\psi \in \Lambda_n^*(E)$ telle que $\psi(e_1, \dots, e_n) = \lambda$, puisque $(\Delta_{\mathcal{B}})$ est une base de $\Lambda_n^*(E)$, il existe $\alpha \in K$ tel que $\psi = \alpha \Delta_{\mathcal{B}}$, d'où $\lambda = \psi(e_1, \dots, e_n) = \alpha \Delta_{\mathcal{B}}(e_1, \dots, e_n) = \alpha$. Donc $\psi = \lambda \Delta_{\mathcal{B}}$. \blacksquare

Pour le lecteur qui trouverait le théorème XIII.2.1 incomplet, signalons que la détermination de la dimension des espaces vectoriels $\Lambda_k^*(E)$ pour $k < n$ est proposée en exercice ⁽²⁾. Il aura déjà remarqué que $\Lambda_1^*(E)$ est évidemment E^* (donc est de dimension n), ce qui explique en partie la notation astérisquée.

⁽¹⁾ C'est à cet endroit précis qu'intervient la commutativité du corps K .

⁽²⁾ Cf. exercice 2 du § XIII.3.

Exercice 1 : Soit E un K -ev et deux sous- K -ev F et G de E supplémentaires : $E = F \oplus G$. On note φ le projecteur associé à (F, G) d'image F . Si $f \in \Lambda_k^*(F)$, vérifier que l'application $E^k \rightarrow K$, $(x_1, \dots, x_k) \mapsto f(\varphi(x_1), \dots, \varphi(x_k))$ appartient à $\Lambda_k^*(E)$. Lorsque E est de dimension finie, en déduire que, pour tout $k \in \llbracket 1, \dim(E) \rrbracket$, l'espace $\Lambda_k^*(E)$ est non nul.

Exercice 2 : On reprend les notations utilisées dans la démonstration du théorème XIII.2.1.

a) Pour $\sigma \in \mathfrak{S}_n$ et $\tau \in \mathfrak{S}_n$, montrer : $\tau * g_\sigma = g_{\sigma\tau^{-1}}$.

b) Soit $U_{\mathcal{B}} : E^n \rightarrow K$, $(x_1, \dots, x_n) \mapsto \sum_{\sigma \in \mathfrak{U}_n} g_\sigma(x_1, \dots, x_n) \left(U_{\mathcal{B}} = \sum_{\sigma \in \mathfrak{U}_n} g_\sigma \right)$. Quelle est l'orbite de $U_{\mathcal{B}}$ sous l'action de \mathfrak{S}_n ? (Attention au cas où K est de caractéristique 2.)

c) Si K est de caractéristique 0, comparer $\Delta_{\mathcal{B}}$ avec l'antisymétrisée (cf. exercice 4 du § XIII.1) de θ_{id} .

Exercice 3 : Soit E un K -ev de dimension $n = 6$. Dans le développement de $\Delta_{\mathcal{B}}$ figure le produit $x_{61} x_{23} x_{45} x_{36} x_{12} x_{54}$. De quel signe est-il précédé ?

§ XIII.3 DÉTERMINANT DE n VECTEURS DANS UNE BASE ; DÉTERMINANT D'UN ENDOMORPHISME

Soit $n \in \mathbb{N}^*$ et E un K -ev de dimension n .

DÉFINITION XIII.3.1

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base du K -ev E . On appelle **déterminant de n vecteurs** x_1, \dots, x_n de E dans la base \mathcal{B} , et on note $\det_{\mathcal{B}}(x_1, \dots, x_n)$ le scalaire

$$\Delta_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{\sigma(1),1} \dots x_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n x_{\sigma(i),i},$$

où, pour chaque $i \in \llbracket 1, n \rrbracket$, $(x_{ji})_{1 \leq j \leq n}$ est la suite des coordonnées de x_i dans \mathcal{B} .

L'étude faite au § XIII.3 prouve que la fonction $\det_{\mathcal{B}} : E^n \rightarrow K$ est **n -linéaire** et **alternée** sur E (donc antisymétrique), et comme on a vu que $\Delta_{\mathcal{B}}(e_1, \dots, e_n) = 1_K$, le théorème XIII.2.2 montre que $\det_{\mathcal{B}}$ est l'**unique** forme **n -linéaire alternée** sur E prenant la valeur 1 sur la suite (e_1, \dots, e_n) .

THÉORÈME XIII.3.1 (formule de Chasles)

Soit $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_n)$ deux bases d'un K -ev E ($n \geq 1$). Pour tout $(x_1, \dots, x_n) \in E^n$ on a :

$$(I) \quad \det_{\mathcal{B}}(x_1, \dots, x_n) = \det_{\mathcal{B}}(f_1, \dots, f_n) \det_{\mathcal{C}}(x_1, \dots, x_n)$$

Démonstration :

C'est une conséquence immédiate du théorème XIII.2.2 car, considérés comme fonctions de $(x_1, \dots, x_n) \in E^n$, les deux membres de (I) sont n -linéaires, alternés, et prennent la valeur $\det_{\mathcal{B}}(f_1, \dots, f_n)$ sur la base (f_1, \dots, f_n) . ■

COROLLAIRE

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base d'un K -ev E ($n \geq 1$), et soit $(x_1, \dots, x_n) \in E^n$. Pour que les vecteurs (x_i) soient **linéairement indépendants** (c.-à-d. **constituent une base de E**), il faut et il suffit que $\det_{\mathcal{B}}(x_1, \dots, x_n) \neq 0$.

Démonstration :

Comme $\dim(E) = n$, les vecteurs (x_i) sont indépendants ssi ils forment une base de E . Si les (x_i) sont liés, $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0$ (proposition XIII.1.2). Si les (x_i) sont indépendants, ils forment une base \mathcal{C} de E , et alors, pour tous y_1, \dots, y_n dans E :

$$\det_{\mathcal{C}}(y_1, \dots, y_n) = \det_{\mathcal{C}}(e_1, \dots, e_n) \det_{\mathcal{B}}(y_1, \dots, y_n).$$

En prenant $(y_1, \dots, y_n) = (x_1, \dots, x_n)$ on trouve :

$$\det_{\mathcal{C}}(e_1, \dots, e_n) \det_{\mathcal{B}}(x_1, \dots, x_n) = 1_K \quad \text{d'où} \quad \det_{\mathcal{B}}(x_1, \dots, x_n) \neq 0. \quad \blacksquare$$

Déterminant d'un endomorphisme

Soit E un K -ev de dimension n ($n \in \mathbb{N}^*$). Donnons-nous f et g dans $\Lambda_n^*(E) \setminus \{0\}$, et $u \in \text{Hom}_K(E)$. Chacune des applications

$$\varphi : E^n \rightarrow K, \quad (x_1, \dots, x_n) \mapsto f(u(x_1), \dots, u(x_n))$$

et
$$\psi : E^n \rightarrow K, \quad (x_1, \dots, x_n) \mapsto g(u(x_1), \dots, u(x_n))$$

est n -linéaire (propriété (ML5) du § XIII.1), et visiblement alternée.

D'où, par le théorème XIII.2.1, $\varphi = \alpha f$ et $\psi = \beta g$ avec $\alpha \in K$, $\beta \in K$. Mais, toujours par le théorème XIII.2.1 on sait que $g = \lambda f$, avec $\lambda \in K^*$, d'où :

$$\psi = \lambda \varphi = \lambda \alpha f = \alpha (\lambda f) = \alpha g, \quad \text{et par suite} \quad \beta = \alpha \quad \text{car} \quad g \neq 0.$$

Donc le scalaire α tel que $\varphi = \alpha f$ ne dépend pas du choix de $f \in \Lambda_n^*(E)$.

DÉFINITION XIII.3.2

Soit E un K -ev de dimension $n \in \mathbb{N}^*$, et $u \in \text{Hom}_K(E)$. On appelle **déterminant de u** , et on note $\det(u)$, le scalaire $\alpha \in K$ tel que :

$\forall f \in \Lambda_n^*(E) \setminus \{0\}, \forall (x_1, \dots, x_n) \in E^n$, on ait

$$(1) \quad f(u(x_1), \dots, u(x_n)) = \alpha f(x_1, \dots, x_n).$$

Les propriétés fondamentales de ce déterminant sont rassemblées dans le :

THÉOREME XIII.3.2

- Soit E un K -ev de dimension $n \in \mathbb{N}^*$, et $u \in \text{Hom}_K(E)$, $v \in \text{Hom}_K(E)$.
- (I) On a : $\det(\text{Id}_E) = 1$.
- (II) $\det(v \circ u) = \det(v) \det(u)$.
- (III) $\det(u) \neq 0 \Leftrightarrow u \in \text{GL}_K(E)$.
- (IV) Pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de E et tout $(x_1, \dots, x_n) \in E$, on a : $\det_{\mathcal{B}}(x_1, \dots, x_n) = \det(w)$, où w est l'endomorphisme tel que $w(e_i) = x_i$ pour $1 \leq i \leq n$.

Démonstration :

(I) se déduit immédiatement de (1).

Pour prouver (II), fixons $f \in \Lambda_n^*(E) \setminus \{0\}$. La définition XIII.3.2 appliquée à $\det(u)$, $\det(v)$ et $\det(v \circ u)$ montre que, pour tout

$$(x_1, \dots, x_n) \in E^n, \quad f[v \circ u(x_1), \dots, v \circ u(x_n)] = \det(v \circ u) f(x_1, \dots, x_n) = \\ = \det(v) f(u(x_1), \dots, u(x_n)) = \det(v) \det(u) f(x_1, \dots, x_n),$$

d'où en choisissant pour (x_1, \dots, x_n) une base de E (ce qui entraîne $f(x_1, \dots, x_n) \neq 0$), $\det(v \circ u) = \det(v) \det(u)$. De (I) et (II) on tire : $\det(u) \det(u^{-1}) = \det(\text{Id}_E) = 1_K$, d'où $\det(u) \neq 0$ si $u \in \text{GL}_K(E)$. En revanche si $u \notin \text{GL}_K(E)$, pour tous $(x_1, \dots, x_n) \in E^n$, $u(x_1), \dots, u(x_n)$ sont liés, d'où $f(u(x_1), \dots, u(x_n)) = 0$ pour toute $f \in \Lambda_n^*(E)$, et, à nouveau avec (1), $\det(u) = 0$.

Enfin, pour prouver (IV), prenons $f = \det_{\mathcal{B}}$, d'où $f(e_1, \dots, e_n) = 1_K$. Alors (1) entraîne :

$$\det_{\mathcal{B}}(x_1, \dots, x_n) = \det_{\mathcal{B}}(w(e_1), \dots, w(e_n)) = \\ = \det(w) \det_{\mathcal{B}}(e_1, \dots, e_n) = \det(w). \quad \blacksquare$$

Exemple 1 : Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $\lambda_i \in K$ ($1 \leq i \leq n$). Notons u l'endomorphisme de E tel que $u(e_i) = \lambda_i e_i$ ($1 \leq i \leq n$). Fixons $f \in \Lambda_n^*(E) \setminus \{0\}$ (par exemple $f = \det_{\mathcal{B}}$). La n -linéarité de f donne : $f[u(x_1), \dots, u(x_n)] = \lambda_1 \dots \lambda_n f(e_1, \dots, e_n)$. Par suite $\det(u) = \lambda_1 \dots \lambda_n$.

En particulier

$$\det(\lambda \text{Id}_E) = \lambda^n.$$

Si $\lambda_1 = \dots = \lambda_{n-1} = 1_K$ et $\lambda_n = \lambda \in K$, on a $\det(u) = \lambda$, ce qui prouve la *surjectivité* de l'application $\text{Hom}_K(E) \rightarrow K, u \mapsto \det(u)$.

Exemple 2 : Avec les notations de l'exemple 1, donnons-nous $\sigma \in \mathfrak{S}_n$, et définissons $u_\sigma \in \text{Hom}_K(E)$ par $u_\sigma(e_i) = e_{\sigma(i)}$ pour $1 \leq i \leq n$. Puisque f est antisymétrique, on a :

$$f(u_\sigma(e_1), \dots, u_\sigma(e_n)) = f(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma) f(e_1, \dots, e_n) = \varepsilon(\sigma) \times 1_K,$$

d'où $\det(u_\sigma) = \varepsilon(\sigma) \cdot 1_K$.

Exemple 3 : Toujours avec les mêmes notations, soit $u \in \text{Hom}_K(E)$ dont la matrice $M = [a_{ij}]$ dans \mathcal{B} soit *trigonale supérieure*. On a donc :

$$u(e_j) = \sum_{i=1}^j a_{ij} e_i \text{ pour tout } j, \text{ et}$$

$$\det(u) = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = a_{11} \det_{\mathcal{B}}(e_1, u(e_2), \dots, u(e_n)).$$

Supposons prouvé que

$$\det(u) = a_{11} \dots a_{kk} \det_{\mathcal{B}}(e_1, \dots, e_k, u(e_{k+1}), \dots, u(e_n)) \text{ avec } k \leq n-1.$$

$$\text{Alors } u(e_{k+1}) = a_{k+1, k+1} e_{k+1} + \sum_{i=1}^k a_{i, k+1} e_i, \text{ d'où (proposition XIII.1.2)}$$

$$\begin{aligned} \det(u) &= a_{11} \dots a_{kk} \det_{\mathcal{B}}(e_1, \dots, e_k, a_{k+1, k+1} e_{k+1}, u(e_{k+2}), \dots, u(e_n)) = \\ &= a_{11} \dots a_{k+1, k+1} \det_{\mathcal{B}}(e_1, \dots, e_{k+1}, u(e_{k+2}), \dots, u(e_n)). \end{aligned}$$

Par récurrence on voit donc que $\det(u) = a_{11} a_{22} \dots a_{nn}$ (ce qui généralise le résultat de l'exemple 1).

On démontrerait de même que si $[a_{ij}]$ est *trigonale inférieure*, on a encore : $\det(u) = a_{11} \dots a_{nn}$.

Groupe spécial linéaire

Conservons les notations du théorème XIII.3.2. En conséquence de (II) et de (III) de cet énoncé, l'application $\text{GL}_K(E) \rightarrow K^*, u \mapsto \det(u)$ est bien définie, et c'est un *homomorphisme de groupes*. Donc, son noyau est un sous-groupe de $\text{GL}_K(E)$.

DÉFINITION XIII.3.3

⎧ Soit E un K -ev de dimension $n \geq 1$. On appelle **groupe spécial**
 ⎧ **linéaire** de E , et on note $\text{SL}_K(E)$, le sous-groupe de $\text{GL}_K(E)$ formé
 ⎧ des $u \in \text{GL}_K(E)$ tels que $\det(u) = 1_K$.

L'exemple 1 ci-dessus a montré que l'homomorphisme $\text{GL}_K(E) \rightarrow K^*, u \mapsto \det(u)$ est surjectif. Etant un noyau d'homomorphisme

groupe $SL_K(E)$ est *distingué* dans $GL_K(E)$, d'où, par passage au quotient, un isomorphisme de groupes $GL_K(E)/SL_K(E) \cong K^*$.

Exercice 1 : Le corps K est supposé de caractéristique nulle. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base d'un K -ev E ($n \geq 1$). Pour $i \in \llbracket 1, n \rrbracket$ on pose $\varepsilon_i = e_i - \sum_{j \neq i} e_j$. Démontrer que

$$\det_{\mathcal{B}}(\varepsilon_1, \dots, \varepsilon_n) = -2^{n-1}(n-2).$$

Indication : Pour $k \in \mathbb{N}^*$, soit A_k la somme des $\varepsilon(\sigma)$ lorsque σ parcourt l'ensemble des *dérangements* dans \mathfrak{S}_K . On pose $A_0 = 1$ et on remarque que $A_1 = 0$. Etablir d'abord que $A_k = (-1)^{k-1} (k-1)$ pour $k \geq 1$, en calculant la série formelle $\sum_{n \geq 0} \frac{A_n}{n!} X^n$. Enfin, si $B_n = \det_{\mathcal{B}}(\varepsilon_1, \dots, \varepsilon_n)$, calculer de même $\sum_{n \geq 0} \frac{B_n}{n!} X^n$. Plus simplement on pourra écrire $\varepsilon_i = 2e_i - s$, avec $s = \sum_{i=1}^n e_i$ et développer $\det(\varepsilon_1, \dots, \varepsilon_n)$.

Exercice 2 : Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base d'un K -ev de dimension $n \geq 3$. On donne $p \in \llbracket 2, n-1 \rrbracket$ et on note \mathcal{E}_p l'ensemble des p -parties de $\llbracket 1, n \rrbracket$, qu'on identifiera à l'ensemble des suites strictement croissantes $\{i_1, \dots, i_p\} \in \llbracket 1, n \rrbracket^p$ ($1 \leq i_1 < \dots < i_p \leq n$). Pour $I \in \mathcal{E}_p$, on désigne par E_I le sous- K -ev engendré par les $(e_i)_{i \in I}$, par F_I celui engendré par les $(e_j)_{j \in \llbracket 1, n \rrbracket \setminus I}$ (d'où $F_I = E \setminus I$, et $E_I \oplus F_I = E$), et par φ_I le projecteur de E sur E_I parallèlement à F_I . Si \mathcal{B}_I est la base $(e_i)_{i \in I}$ de E_I , on considère sur E_I^p la fonction $D_I = \det_{\mathcal{B}_I} \in \Lambda_p^*(E_I)$, et on pose :

$$\Delta_I(x_1, \dots, x_p) = D_I(\varphi_I(x_1), \dots, \varphi_I(x_p)) \quad \text{pour tous } (x_1, \dots, x_p) \in E^p.$$

a) Soit $I = \{i_1, \dots, i_p\} \in \mathcal{E}_p$ et $J = \{j_1, \dots, j_p\} \in \mathcal{E}_p$. Montrer :

$$\Delta_I(e_{j_1}, \dots, e_{j_p}) = \delta_{i_1, j_1} \dots \delta_{i_p, j_p} \times 1_K \quad (\delta = \text{symbole de Kronecker}).$$

b) Démontrer que $(\Delta_I)_{I \in \mathcal{E}_p}$ est une *base* du K -ev $\Lambda_p^*(E)$, et donc que $\dim \Lambda_p^*(E) = \binom{n}{p}$ (si $\varphi \in \Lambda_p^*(E)$, et $(x_1, \dots, x_p) \in E^p$ on développera $\varphi(x_1, \dots, x_p)$ à partir des coordonnées $(x_{ji})_{1 \leq j \leq n}$ des x_i dans \mathcal{B} par un calcul analogue à celui du théorème XIII.2.1).

Exercice 3 : Le corps commutatif K est supposé fini, de cardinal q . Pour $n \in \mathbb{N}^*$, si E est un K -ev de dimension n , calculer $\text{card}(SL_K(E))$.

Exercice 4 : Soit E un K -ev de dimension $n \geq 2$ et E_1, \dots, E_p des sous- K -ev de dimensions d_1, \dots, d_p $\left(d_i \geq 1, \sum_{i=1}^p d_i = n \right)$ tels que $E = \bigoplus_{i=1}^p E_i$.

Pour chaque $i \in \llbracket 1, p \rrbracket$, soit $u_i \in \text{Hom}_K(E_i)$. Enfin on pose : $u = \bigoplus_{i=1}^p u_i$. Démontrer que $\det(u) = \prod_{i=1}^p \det(u_i)$.

Indication : Se ramener au cas où $p = 2$.

Exercice 5 : Soit u un endomorphisme *nilpotent* d'un K -ev de dimension finie $n \in \mathbb{N}^*$. Montrer que $\det(u) = 0_K$ (cf. exercice 4 du § XI.6).

§ XIII.4 DÉTERMINANT D'UNE MATRICE CARRÉE

Considérons une matrice $M = [a_{ij}] \in \mathfrak{M}_n(K)$ ($n \geq 1$). On peut la considérer comme la matrice d'un endomorphisme $u_M \in \text{Hom}_K(K^n)$, K^n étant rapporté à sa base canonique $\mathcal{B} = (e_1, \dots, e_n)$. L'étude du § XIII.3 nous a prouvé que les scalaires $\det(u_M)$ et

$$\det_{\mathcal{B}}(u_M(e_1), \dots, u_M(e_n)) = \det_{\mathcal{B}}(\mathcal{C}_1(M), \dots, \mathcal{C}_n(M))$$

sont égaux, leur valeur commune étant

$$\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i),i}.$$

DÉFINITION XIII.4.1

On appelle **déterminant de la matrice carrée** $M = [a_{ij}] \in \mathfrak{M}_n(K)$, et on note $\det(M)$, le scalaire $\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}$. On écrit aussi $\det(M)$ sous la forme $|a_{ij}|_{(i,j) \in \llbracket 1,n \rrbracket^2}$, ou encore

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Pour $n = 1$, $\det(M) =$ unique terme de M .

Pour $n = 2$, $M = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, et $\det(M) = a_{11} a_{22} - a_{21} a_{12}$.

La formule (1) $\det M = \det_{\mathcal{B}}(\mathcal{C}_1(M), \dots, \mathcal{C}_n(M))$ montre que $\det(M)$ est une **fonction n -linéaire et alternée des colonnes** de M . En conséquence, si l'on rajoute à l'une des colonnes de M une combinaison linéaire des autres colonnes, le déterminant ne change pas ; et si l'on permute les colonnes de M , ce qui change la matrice M en la matrice M_{σ} de colonnes $\mathcal{C}_{\sigma(1)}, \dots, \mathcal{C}_{\sigma(n)}$ (où $\sigma \in \mathfrak{S}_n$), on a : $\det(M_{\sigma}) = \varepsilon(\sigma) \det(M)$. La formule (2) $\det M = \det(u_M)$ permet de déduire des propriétés du déterminant d'un endomorphisme les suivantes :

THÉORÈME XIII.4.1

On a :

(I) $\det(I_n) = 1_K$

(II) pour $M \in \mathfrak{M}_n(K)$ et $N \in \mathfrak{M}_n(K)$:

$$\det(MN) = \det(M) \det(N)$$

(III) $M \in \mathfrak{M}_n(K)$ est inversible ssi $\det(M) \neq 0$.

Démonstration :

(I) provient du fait que la matrice unité I_n représente Id_{K^n} dans la base canonique, et que $\det(\text{Id}_{K^n}) = 1$.

Pour (II) on sait d'après l'interprétation « géométrique » du produit matriciel (§ XI.3) que $u_{MN} = u_M \circ u_N$, et d'après le théorème XIII.3.2, que

$$\det(u_M \circ u_N) = \det(u_M) \det(u_N),$$

d'où

$$\det(MN) = \det(u_{MN}) = \det(u_M) \det(u_N) = \det(M) \det(N).$$

Quant à (III), on sait que M est inversible ssi $u_M \in \text{GL}_K(K^n)$, et toujours par le théorème XIII.3.2,

$$u_M \in \text{GL}_K(K^n) \text{ ssi } \det(u_M) \neq 0. \quad \blacksquare$$

Rappelons que l'inversibilité de u_M équivaut à l'indépendance linéaire des colonnes $\mathcal{C}_1(M), \dots, \mathcal{C}_n(M)$: donc cette indépendance linéaire équivaut à son tour à : $\det(M) \neq 0$.

Vu son importance nous allons donner de la propriété (II) de multiplicativité une démonstration par un calcul direct, indépendant de toute interprétation, et l'on va voir que ce calcul n'utilise que la structure d'anneau commutatif de K , le fait que cet anneau est intègre n'intervenant pas.

Soit donc $M = [a_{ij}]$, $N = [b_{ij}]$ deux matrices carrées d'ordre n , et posons $MN = [c_{ij}]$, d'où $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. Alors

$$\det(MN) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n c_{\sigma(i), i},$$

mais, en notant $\mathcal{F} = \mathcal{F}(\llbracket 1, n \rrbracket, \llbracket 1, n \rrbracket)$

$$\prod_{i=1}^n c_{\sigma(i), i} = \sum_{\varphi \in \mathcal{F}} a_{\sigma(1), \varphi(1)} b_{\varphi(1), 1} \cdots a_{\sigma(n), \varphi(n)} b_{\varphi(n), n},$$

$$\text{d'où } \det(MN) = \sum_{\varphi \in \mathcal{F}} b_{\varphi(1), 1} \cdots b_{\varphi(n), n} \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1), \varphi(1)} \cdots a_{\sigma(n), \varphi(n)} \right).$$

Pour $\varphi \in \mathcal{F}$, notons $D_\varphi = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1), \varphi(1)} \cdots a_{\sigma(n), \varphi(n)}$. Il est clair que $D_\varphi =$

$\det(M_\varphi)$, où M_φ est la matrice des colonnes $\mathcal{C}_{\varphi(1)}(M), \dots, \mathcal{C}_{\varphi(n)}(M)$. Si φ est *non injective* (i.e. $\varphi \notin \mathfrak{S}_n$), il s'ensuit que $D_\varphi = 0$, et si $\varphi \in \mathfrak{S}_n$: $D_\varphi = \varepsilon(\varphi) \det(M)$. Donc

$$\begin{aligned} \det(MN) &= \sum_{\varphi \in \mathfrak{S}_n} b_{\varphi(1), 1} \cdots b_{\varphi(n), n} \varepsilon(\varphi) \det M = \det M \sum_{\varphi \in \mathfrak{S}_n} \varepsilon(\varphi) b_{\varphi(1), 1} \cdots b_{\varphi(n), n} \\ &= \det(M) \det(N). \end{aligned}$$

Exemple 1 : Les formules (1) et (2) permettent parfois le calcul direct du déterminant d'une matrice. Par exemple, tenant compte des exemples 1, 2 et 3 du § XIII.3, on voit que si $M = [a_{ij}]$ est *trigonale*, alors $\det(M) = a_{11} a_{22} \dots a_{nn}$; et si M est une *matrice de permutation* $M = [\delta_{\sigma^{-1}(i), j}]_{(i, j) \in \llbracket 1, n \rrbracket^2}$, avec $\sigma \in \mathfrak{S}_n$, alors $\det(M) = \varepsilon(\sigma) \cdot 1_K$.

Exemple 2 : Reprenons les notations du théorème XI.5.1 et écrivons $M \in \text{GL}(n, K)$ sous la forme

$$M = DT_1 \dots T_r = U_1 \dots U_s \Delta,$$

où les T_i et les U_j sont des matrices de transvection, et où

$$D = D_n(\lambda) \quad \text{et} \quad \Delta = D_n(\mu) \quad (\lambda \in K^*, \mu \in K^*)$$

sont des matrices de dilatation. Alors $\det(T_i) = \det(U_j) = 1_K$ pour tous i et j d'après l'exemple 1 ci-dessus. Donc $\det(M) = \det(D) = \det(\Delta) = \lambda = \mu$, d'où $D = \Delta$ et ainsi se trouve justifiée la remarque 2 du § XI.5.

THÉORÈME XIII.4.2

|| Soit E un K -ev de dimension $n \geq 1$, et $u \in \text{Hom}_K(E)$. Pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de E , on a : $\det(u) = \det(\text{Mat}_{\mathcal{B}}(u))$.

Démonstration :

On sait que $\det(u) = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$.

Si $\text{Mat}_{\mathcal{B}}(u) = [a_{ij}]$, les définitions mêmes de ces scalaires montrent que

$$\det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \det([a_{ij}]). \quad \blacksquare$$

THÉORÈME XIII.4.3

|| Pour $M = [a_{ij}] \in \mathfrak{M}_n(K)$ ($n \geq 1$), on a : $\det({}^t M) = \det(M)$.

Démonstration :

On a $\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i), i}$ et ${}^t M = [b_{ij}]$ avec

$b_{ij} = a_{ji}$ pour tous i et j .

Pour $\sigma \in \mathfrak{S}_n$, on a $\prod_{i=1}^n a_{\sigma(i), i} = \prod_{i=1}^n a_{i, \sigma^{-1}(i)}$. Comme de plus l'application $\sigma \mapsto \sigma^{-1}$ est bijective, il s'ensuit :

$$\det(M) = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau^{-1}) \prod_{i=1}^n a_{i, \tau(i)} = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau^{-1}) \prod_{i=1}^n b_{\tau(i), i}.$$

Mais évidemment $\varepsilon(\tau^{-1}) = \varepsilon(\tau)$ pour tout $\tau \in \mathfrak{S}_n$, d'où

$$\det({}^t M) = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau) \prod_{i=1}^n b_{\tau(i), i} = \det(M). \quad \blacksquare$$

En conséquence de ce théorème, *toute propriété des déterminants d'une matrice vraie pour les colonnes est encore vraie avec les lignes de M .*

Ainsi, $\det(M)$ est une *forme n -linéaire et alternée* (donc antisymétrique) en les *lignes* de M , et $\det(M) \neq 0$ équivaut à l'indépendance linéaire des lignes de M .

Dérivation d'un déterminant

Soit K un corps commutatif de caractéristique 0, et M une matrice dont les coefficients sont dans $K(X)$. Dans ces conditions $\det(M)$ devient une fonction rationnelle de l'indéterminée X , et la question se pose naturellement de calculer la dérivée de cette fonction. Soit donc $I \subset K$ ($I \neq \emptyset$) où est définie $\psi : I \rightarrow K$, $t \mapsto \det(M(t)) = \det(\mathcal{C}_1(t), \dots, \mathcal{C}_n(t))$. Il suffit de se reporter au théorème XIII.1.2 pour conclure que la fonction ψ a une dérivée rationnelle donnée par la formule

$$\psi'(t) = \sum_{k=1}^n \det(\mathcal{C}_1(t), \dots, \mathcal{C}_{k-1}(t), \mathcal{C}'_k(t), \mathcal{C}_{k+1}(t), \dots, \mathcal{C}_n(t)),$$

où $\mathcal{C}'_k(t)$ désigne bien entendu le vecteur colonne qui a pour coordonnées les dérivées des coordonnées du vecteur colonne $\mathcal{C}_k(t)$.

Exemple 3 : Considérons le déterminant

$$\Delta_n = \begin{vmatrix} 1+x & 1 & \dots & 1 \\ 1 & 1+x & & 1 \\ 1 & 1 & & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & & 1+x \end{vmatrix}$$

que l'on se propose de calculer. On obtient immédiatement

$$\Delta'_n(x) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & 1+x & & 1 \\ 0 & 1 & & \cdot \\ \vdots & 1 & & \vdots \\ 0 & 1 & \dots & 1+x \end{vmatrix} + \dots + \begin{vmatrix} 1+x & \dots & 0 \\ 1 & & 0 \\ 1 & & 0 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{vmatrix} = n\Delta_{n-1}(x).$$

En partant de $\Delta_1 = x + 1$, $\Delta_2 = x^2 + 2x$, et en tenant compte de $\Delta_n(0) = 0$, on en déduit immédiatement $\Delta_n = x^n + nx^{n-1}$.

Mineurs d'une matrice, développements d'un déterminant

Les déterminants des *sous-matrices carrées* d'une matrice $M \in \mathfrak{M}_{p,n}(K)$ s'appellent les **mineurs** de M . Dans ce qui suit, nous emploierons

suivante : si $I \subset \llbracket 1, p \rrbracket$ et $J \subset \llbracket 1, n \rrbracket$ avec

$$\text{card}(I) = \text{card}(J) = r \geq 1, \quad r \leq \text{Min}(p, n),$$

$\Delta_{I,J}(M)$ désignera le mineur $\det(\mathcal{M}_{I,J}(M))$ de M . L'entier r s'appellera l'ordre de ce mineur. On a bien sûr, à cause du théorème XIII.4.3 $\Delta_{I,J}(M) = \Delta_{J,I}({}^t M)$. Nous conviendrons que $\Delta_{\emptyset \emptyset}(M) = 1_K$. Si M est elle-même carrée, les mineurs $\Delta_{I,I}(M)$, avec $I \subset \llbracket 1, n \rrbracket$, seront appelés les *mineurs centrés* (sous-entendu sur la diagonale), et les mineurs $\Delta_{\llbracket 1, r \rrbracket, \llbracket 1, r \rrbracket}(M)$ ($1 \leq r \leq n$) sont appelés les **mineurs principaux** de M . Toujours en supposant M carrée, les mineurs d'ordre $n-1$ seront ainsi désignés à l'aide des complémentaires de I et J dans $\llbracket 1, n \rrbracket$: si $I = \llbracket 1, n \rrbracket \setminus \{i\}$ et $J = \llbracket 1, n \rrbracket \setminus \{j\}$, on pose ainsi : $D_{i,j}(M) = \Delta_{I,J}(M)$. On a besoin de cette notation spéciale car les mineurs d'ordre $n-1$ jouent un rôle particulièrement important : la matrice $[D_{i,j}(M)]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ se note $\Lambda^{n-1}(M)$ et s'appelle **puissance extérieure $(n-1)$ -ième** de M .

La matrice $\check{M} = [(-1)^{i+j} D_{i,j}(M)]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ s'appelle la **comatrice** de M (ou mieux **matrice des cofacteurs de M** , le scalaire $(-1)^{i+j} D_{i,j}(M)$ étant, par définition, le **cofacteur d'indice (i, j)** de M (ou de a_{ij})). Et la matrice ${}^t(\check{M}) = ({}^t\check{M})$, qu'on notera simplement ${}^t\check{M}$, ou \tilde{M} , s'appelle la **matrice complémentaire** de M .

THÉORÈME XIII.4.4 (Développement d'un déterminant par rapport à une ligne ou à une colonne).

On suppose n entier ≥ 2 .

Soit $M = [a_{ij}] \in \mathfrak{M}_n(K)$. Pour tous i et j dans $\llbracket 1, n \rrbracket$, on a :

$$(3) \quad \det(M) = \sum_{k=1}^n (-1)^{k+j} a_{k,j} D_{k,j}(M).$$

*(C'est le développement de $\det(M)$ par rapport à la j -ième colonne)
et*

$$(4) \quad \det(M) = \sum_{k=1}^n (-1)^{i+k} a_{i,k} D_{i,k}(M)$$

(c'est le développement de $\det(M)$ par rapport à la i -ième ligne).

Démonstration :

En raison du théorème XIII.4.3, il suffit de prouver (3). Si $j < n$, soit M' la matrice dont les colonnes sont :

$$\mathcal{C}_1(M), \mathcal{C}_2(M), \dots, \mathcal{C}_{j-1}(M), \mathcal{C}_{j+1}(M), \dots, \mathcal{C}_n(M), \mathcal{C}_j(M).$$

Alors $D_{k,n}(M') = D_{k,j}(M)$ pour tout k , et $\det(M') = (-1)$

Par suite, il suffit de prouver (3) pour $j = n$, ce que nous supposons. Or, on a :

$$(5) \quad \det(M) = \sum_{k=1}^n B_k \quad \text{avec, pour } k \in \llbracket 1, n \rrbracket :$$

$$B_k = \sum_{\sigma \in \mathcal{E}_k} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = a_{k,n} \sum_{\sigma \in \mathcal{E}_k} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n-1),n-1},$$

en posant $\mathcal{E}_k = \{\sigma \in \mathfrak{S}_n \mid \sigma(n) = k\}$. En effet, les $(\mathcal{E}_k)_{1 \leq k \leq n}$ forment une partition de \mathfrak{S}_n . Pour chaque $k \in \llbracket 1, n \rrbracket$, soit s_k l'élément de \mathfrak{S}_n tel que $s_k(k) = n$ et que $s_k|_{\llbracket 1, n-1 \rrbracket}$ soit la bijection croissante de $\llbracket 1, n \rrbracket \setminus \{k\}$ sur $\llbracket 1, n-1 \rrbracket$, et posons $\tau_k = s_k^{\langle -1 \rangle}$. L'application $\mathcal{E}_k \rightarrow \mathcal{E}_n$, $\sigma \mapsto s_k \circ \sigma$ est bijective, d'où :

$$(6) \quad \begin{aligned} B_k &= a_{k,n} \sum_{\tau \in \mathcal{E}_n} \varepsilon(\tau_k \circ \tau) a_{\tau_k \circ \tau(1),1} \cdots a_{\tau_k \circ \tau(n-1),n-1} \\ &= \varepsilon(\tau_k) a_{k,n} \sum_{\tau \in \mathcal{E}_n} \varepsilon(\tau) a_{\tau_k \circ \tau(1),1} \cdots a_{\tau_k \circ \tau(n-1),n-1}. \end{aligned}$$

La bijection $\tau \mapsto \tau|_{\llbracket 1, n-1 \rrbracket}$, $\mathcal{E}_n \rightarrow \mathfrak{S}_{n-1}$ préserve la signature (compter les inversions !) et par définition

$$D_{k,n}(M) = \det (a_{\tau_k(\lambda), \mu})_{(\lambda, \mu) \in \llbracket 1, n-1 \rrbracket^2}.$$

On déduit donc de (6) :

$$B_k = \varepsilon(\tau_k) a_{k,n} \sum_{s \in \mathfrak{S}_{n-1}} \varepsilon(s) a_{\tau_k \circ s(1),1} \cdots a_{\tau_k \circ s(n-1),n-1} = \varepsilon(\tau_k) a_{k,n} D_{k,n}(M)$$

et on voit facilement que $\varepsilon(\tau_k) = (-1)^{n-k}$, d'où, avec (5) :

$$\det(M) = \sum_{k=1}^n (-1)^{k+n} a_{k,n} D_{k,n}(M). \quad \blacksquare$$

Le théorème XIII.4.4 constitue bien sûr un outil important pour calculer des déterminants, mais il a aussi des conséquences théoriques :

THÉORÈME XIII.4.5

Soit $M = [a_{ij}] \in \mathfrak{M}_n(K)$.

(I) On a : $M\tilde{M} = \tilde{M}M = \det(M) I_n$.

(II) Supposons que tous les $a_{i,j}$ appartiennent à un même sous-anneau A de K ; alors $M \in \text{GL}(n, A)$ ssi $\det(A)$ est inversible dans A , et s'il en est ainsi, l'inverse de M dans $\mathfrak{M}_n(A)$ est $(\det(A))^{-1} \tilde{M}$, et c'est l'inverse de M dans $\mathfrak{M}_n(K)$.

$$\left\| \begin{array}{l} \text{En particulier, } M \in \text{GL}(n, K) \text{ ssi } \det(M) \neq 0, \text{ et lorsque} \\ \det(M) \neq 0, \text{ on a } M^{-1} = (\det(M))^{-1} \tilde{M}, \text{ et } \det(M^{-1}) = \\ (\det(M))^{-1}. \end{array} \right.$$

Démonstration :

Posons $\tilde{M} = [\tilde{a}_{i,j}]$, i.e.

$$\tilde{a}_{i,j} = (-1)^{i+j} D_{j,i}(M) \quad ((i,j) \in \llbracket 1, n \rrbracket^2).$$

$$\text{Alors } \tilde{M}M = [c_{ij}], \text{ où } c_{ij} = \sum_{k=1}^n \tilde{a}_{i,k} a_{k,j} = \sum_{k=1}^n (-1)^{k+i} D_{k,i}(M) a_{k,j}.$$

Si $i = j$, le théorème XIII.4.4 montre directement que $c_{i,i} = \det(M)$. Si $i \neq j$, il montre que $c_{i,j}$ est le déterminant de la matrice des colonnes $\Gamma_1, \dots, \Gamma_n$, avec $\Gamma_r = \mathcal{C}_r(M)$ pour $r \neq i$ et $\Gamma_i = \mathcal{C}_j(M)$. En particulier, dans ce cas, $\Gamma_i = \Gamma_j$, donc $c_{i,j} = 0$. Finalement $\tilde{M}M = \det(M) I_n$. En raisonnant sur les lignes, on verrait de même que $MM\tilde{M} = \det(M) I_n$, d'où (I).

Prouvons maintenant l'assertion (II) : si $M \in \text{GL}(n, A)$, on a, en notant M^{-1} son inverse dans $\mathfrak{M}_n(A)$ (qui est aussi son inverse dans $\mathfrak{M}_n(K)$) :

$$\det(MM^{-1}) = \det(M) \det(M^{-1}) = \det(I_n) = 1_K = 1_A,$$

et la formule de définition du déterminant montre que $\det(M) \in A$, $\det(M^{-1}) \in A$, donc $\det(M)$ est inversible dans A , et on a prouvé de plus que $(\det(M))^{-1} = \det(M^{-1})$. Réciproquement, si $\det(M)$ est inversible dans A , on a, d'après (I) :

$$M \times [(\det(M))^{-1} \tilde{M}] = [(\det(M))^{-1} \tilde{M}] M = I_n,$$

et comme il est clair que $\tilde{M} \in \mathfrak{M}_n(A)$, on en déduit que $M \in \text{GL}(n, A)$, son inverse dans $\mathfrak{M}_n(A)$ étant $(\det(M))^{-1} \tilde{M}$. ■

Matrices trigonales par blocs

THÉOREME XIII.4.6

$$\left\| \begin{array}{l} \text{Soit } M = [a_{ij}] \in \mathfrak{M}_n(K) \text{ avec } n \geq 2, \text{ soit } p \in \llbracket 1, n-1 \rrbracket \text{ et} \\ q = n - p. \text{ Supposons } a_{\lambda, \mu} = 0 \text{ pour } \lambda > p \text{ et } \mu \leq p. \text{ Alors} \\ \det(M) = \Delta_{\llbracket 1, p \rrbracket \times \llbracket 1, p \rrbracket}(M) \times \Delta_{\llbracket p+1, n \rrbracket \times \llbracket p+1, n \rrbracket}(M). \end{array} \right.$$

Démonstration :

Notons Γ le sous-groupe de \mathfrak{S}_n formé des $\sigma \in \mathfrak{S}_n$ tels que $\sigma(\llbracket 1, p \rrbracket) = \llbracket 1, p \rrbracket$, G (resp. \tilde{G}) le sous-groupe des $\sigma \in \mathfrak{S}_n$ tels que $\sigma(i) = i$ pour

$$i \in \llbracket p+1, n \rrbracket \text{ (resp. } i \in \llbracket 1, p \rrbracket).$$

L'application $G \times \tilde{G} \rightarrow \Gamma$, $(s, t) \mapsto st$ est bijective, et si on munit $G \times \tilde{G}$ de sa structure de *groupe produit*, c'est un isomorphisme de groupes. Enfin, les applications

$$G \rightarrow \mathfrak{S}_p, \sigma \mapsto \sigma \parallel \llbracket 1, p \rrbracket \quad \text{et} \quad \tilde{G} \rightarrow \mathfrak{S}_q, \sigma \mapsto \bar{\sigma}$$

(où $\bar{\sigma}(j) = \sigma(p+j) - p$ pour $p+1 \leq j \leq n$) sont des isomorphismes de groupes qui conservent la signature.

Cela dit, on a : $\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}$,

et l'hypothèse faite montre que $a_{\sigma(1),1} \dots a_{\sigma(n),n} = 0$ pour $\sigma \notin \Gamma$. Donc

$$\begin{aligned} \det(M) &= \sum_{(s,t) \in G \times \tilde{G}} \varepsilon(st) a_{s(1),1} \dots a_{s(p),p} a_{t(p+1),p+1} \dots a_{t(n),n} \\ &= \sum_{(s,t) \in G \times \tilde{G}} (\varepsilon(s) a_{s(1),1} \dots a_{s(p),p}) (\varepsilon(t) a_{t(p+1),p+1} \dots a_{t(n),n}) = ST, \end{aligned}$$

avec

$$S = \sum_{s \in G} \varepsilon(s) a_{s(1),1} \dots a_{s(p),p} = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(p),p} = \Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(M)$$

et

$$\begin{aligned} T &= \sum_{t \in \tilde{G}} \varepsilon(t) a_{t(p+1),p+1} \dots a_{t(n),n} \\ &= \sum_{\sigma \in \mathfrak{S}_q} \varepsilon(\sigma) a_{p+\sigma(1),p+1} \dots a_{p+\sigma(q),p+q} = \Delta_{\llbracket p+1, n \rrbracket, \llbracket p+1, n \rrbracket}(M). \quad \blacksquare \end{aligned}$$

A l'aide du théorème XIII.4.3, on voit que la conclusion du théorème XIII.4.6 subsiste si $a_{\lambda,\mu} = 0$ pour $\lambda \leq p$ et $\mu > p$, c'est-à-dire si le bloc de zéros est en haut à droite au lieu d'être en bas à gauche.

Par récurrence on étend le théorème à un nombre de blocs supérieur à 4 :

COROLLAIRE

$$\left\| \begin{array}{l} \text{Supposons } M = [a_{ij}] \in \mathfrak{M}_n(K) \text{ **trigonale par blocs** relativement à} \\ \text{une suite d'entiers } d_1, d_2, \dots, d_r \text{ (} d_1 + \dots + d_r = n, d_i \geq 1 \text{ pour tout} \\ \text{ } i, r \geq 2 \text{), de blocs diagonaux} \\ \\ M_{1,1}, M_{2,2}, \dots, M_{r,r} \text{ (} M_{i,i} \in \mathfrak{M}_{d_i}(K) \text{)}. \\ \\ \text{Alors } \det(M) = \det(M_{11}) \times \dots \times \det(M_{r,r}). \end{array} \right.$$

Naturellement, si M est *trigonale*, on retrouve le fait, déjà vu dans l'exemple 1, que $\det(M) = a_{11} a_{22} \dots a_{nn}$.

Formule de Laplace

Les théorèmes XIII.4.4 et XIII.4.6 sont des cas particuliers d'une relation qui exprime $\det(M)$ en fonction des mineurs d'ordre donné de M .

commodément cette propriété, introduisons quelques notations : soit $n \in \mathbb{N}$ ($n \geq 2$) et $p \in \llbracket 1, n-1 \rrbracket$. Nous noterons $\mathcal{F}_p(n)$ l'ensemble des p -parties de $\llbracket 1, n \rrbracket$. Si $I \in \mathcal{F}_p(n)$, σ_I désignera l'élément de \mathfrak{S}_n tel que $\sigma_I|_{\llbracket 1, p \rrbracket}$ (resp. $\sigma_I|_{\llbracket p+1, n \rrbracket}$) définisse une bijection croissante de $\llbracket 1, p \rrbracket$ sur I (resp. de $\llbracket p+1, n \rrbracket$ sur $\llbracket 1, n \rrbracket \setminus I$). Le complémentaire $\llbracket 1, n \rrbracket \setminus I$ de $I \subset \llbracket 1, n \rrbracket$ sera noté en abrégé \tilde{I} . On vérifie (en calculant le nombre d'inversions), que $\varepsilon(\sigma_I) = (-1)^{\nu(I, \tilde{I})}$, où $\nu(I, \tilde{I}) =$ nombre des $(i, j) \in I \times \tilde{I}$ tels que $i > j$.

THÉORÈME XIII.4.7 (formule de Laplace)

$$\left\| \begin{array}{l} \text{Soit } M = [a_{ij}] \in \mathfrak{M}_n(K) \text{ } (n \geq 2) \text{ et } p \in \llbracket 1, n-1 \rrbracket. \text{ Pour toute partie} \\ J \in \mathcal{F}_p(n), \text{ on a :} \end{array} \right\| \quad (7) \quad \boxed{\det(M) = \sum_{I \in \mathcal{F}_p(n)} \varepsilon(\sigma_I) \varepsilon(\sigma_J) \Delta_{I,J}(M) \Delta_{\tilde{I}, \tilde{J}}(M)}.$$

Démonstration :

a) Ramenons-nous d'abord au cas où $J = \llbracket 1, p \rrbracket$. Pour cela, soit N la matrice dont les colonnes sont $\mathcal{C}_{\sigma_J(1)}(M), \dots, \mathcal{C}_{\sigma_J(n)}(M)$. Alors $\det(N) = \varepsilon(\sigma_J) \det(M)$. Supposant (7) prouvée avec $J = \llbracket 1, p \rrbracket$ on a :

$$\det(N) = \sum_{I \in \mathcal{F}_p(n)} \varepsilon(\sigma_I) \Delta_{I, \llbracket 1, p \rrbracket}(N) \Delta_{\tilde{I}, \llbracket p+1, n \rrbracket}(N).$$

Mais, pour $I \in \mathcal{F}_p(n)$, $\Delta_{I, \llbracket 1, p \rrbracket}(N) = \Delta_{I,J}(M)$ et $\Delta_{\tilde{I}, \llbracket p+1, n \rrbracket}(N) = \Delta_{\tilde{I}, \tilde{J}}(M)$, d'où (7) avec M .

b) On peut donc supposer maintenant $J = \llbracket 1, p \rrbracket$ et il s'agit de prouver (7). Pour cela posons $q = n - p$, et notons, comme dans le théorème XIII.4.6, G (resp. \tilde{G}) le groupe des $\sigma \in \mathfrak{S}_n$ tels que $\sigma(j) = j$ si $j > p$ (resp. $\sigma(j) = j$ si $j \leq p$), et Γ le groupe des $\sigma \in \mathfrak{S}_n$ tels que $\sigma(\llbracket 1, p \rrbracket) = \llbracket 1, p \rrbracket$. L'application $G \times \tilde{G} \rightarrow \Gamma$, $(s, t) \mapsto st$ est un isomorphisme de groupes ($G \times \tilde{G}$ étant muni de sa structure de groupe produit). Les applications $G \rightarrow \mathfrak{S}_p$, $\sigma \mapsto \sigma|_{\llbracket 1, p \rrbracket}$ et $\tilde{G} \rightarrow \mathfrak{S}_q$, $\sigma \mapsto \bar{\sigma}$ (où $\bar{\sigma}(j) = \sigma(p+j) - p$ pour $j \in \llbracket 1, q \rrbracket$) sont des isomorphismes qui conservent la signature.

Notons, pour $I \in \mathcal{F}_p(n)$, $\mathcal{E}_I = \{\sigma \in \mathfrak{S}_n \mid \sigma(\llbracket 1, p \rrbracket) = I\}$. L'application $\Gamma \rightarrow \mathcal{E}_I$, $\sigma \mapsto \sigma_I \sigma$ est bijective, et les $(\mathcal{E}_I)_{I \in \mathcal{F}_p(n)}$ forment une partition de \mathfrak{S}_n . On a donc

$$(8) \quad \det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} = \sum_{I \in \mathcal{F}_p(n)} S_I,$$

avec, pour $I \in \mathcal{F}_p(n)$:

$$S_I = \sum_{\sigma \in \mathcal{E}_I} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

Mais, d'après ce qui précède, pour I fixé, $I \in \mathcal{F}_p(n)$, on a :

$$\begin{aligned} S_I &= \sum_{\sigma \in \Gamma} \varepsilon(\sigma_I \sigma) a_{\sigma_I \sigma(1), 1} \cdots a_{\sigma_I \sigma(n), n} = \varepsilon(\sigma_I) \sum_{\sigma \in \Gamma} \varepsilon(\sigma) a_{\sigma_I \sigma(1), 1} \cdots a_{\sigma_I \sigma(n), n} \\ &= \varepsilon(\sigma_I) \sum_{(s, t) \in \tilde{G} \times \tilde{G}} \varepsilon(s) \varepsilon(t) a_{\sigma_I s(1), 1} \cdots a_{\sigma_I s(p), p} a_{\sigma_I t(p+1), p+1} \cdots a_{\sigma_I t(n), n} \end{aligned}$$

(car $st(i) = s(i)$ pour $i \leq p$ et $st(i) = t(i)$ pour $i \geq p+1$), d'où :

$$(9) \quad S_I = UV,$$

avec

$$(10) \quad U = \sum_{s \in \tilde{G}} \varepsilon(s) a_{\sigma_I s(1), 1} \cdots a_{\sigma_I s(p), p} = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) a_{\sigma_I \sigma(1), 1} \cdots a_{\sigma_I \sigma(p), p} = \Delta_{I, \llbracket p+1, p \rrbracket} (M),$$

et :

$$(11) \quad V = \sum_{t \in \tilde{G}} \varepsilon(t) a_{\sigma_I t(p+1), p+1} \cdots a_{\sigma_I t(n), n} = \sum_{\sigma \in \mathfrak{S}_q} \varepsilon(\sigma) a_{\sigma_I(p+\sigma(1)), p+1} \cdots a_{\sigma_I(p+\sigma(q)), p+q} = \Delta_{\tilde{I}, \llbracket p+1; n \rrbracket} (M).$$

On voit alors que (7) découle de (8), (9), (10) et (11). ■

Le mineur $\Delta_{\tilde{I}, \tilde{J}}(M)$ est souvent appelé *mineur complémentaire* de $\Delta_{I, J}(M)$.

Remarque 1 : Du fait que $\det(M) = \det({}^t M)$, on a aussi, pour tout $I = \mathcal{F}_p(n)$,

$$(7') \quad \det(M) = \sum_{J \in \mathcal{F}_p(n)} \varepsilon(\sigma_I) \varepsilon(\sigma_J) \Delta_{I, J}(M) \Delta_{\tilde{I}, \tilde{J}}(M)$$

qui est une autre façon de développer un déterminant par la règle de Laplace.

Exemple 4 : Ecrivons la formule (7) pour $n = 4$, $p = 2$ et $J = \{1, 2\}$, avec $M = [a_{ij}]$. On a le tableau suivant :

I	$\{1, 2\}$	$\{1, 3\}$	$\{1, 4\}$	$\{2, 3\}$	$\{2, 4\}$	$\{3, 4\}$
$v(I, \tilde{I})$	0	1	2	2	3	4
$\varepsilon(\sigma_I)$	+ 1	- 1	+ 1	+ 1	- 1	+ 1

d'où le développement :

$$\begin{aligned} \det(M) &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \begin{vmatrix} a_{23} & a_{24} \\ a_{43} & a_{44} \end{vmatrix} + \\ &+ \begin{vmatrix} a_{11} & a_{12} \\ a_{41} & a_{42} \end{vmatrix} \begin{vmatrix} a_{23} & a_{24} \\ a_{33} & a_{34} \end{vmatrix} + \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \begin{vmatrix} a_{13} & a_{14} \\ a_{43} & a_{44} \end{vmatrix} \\ &- \begin{vmatrix} a_{21} & a_{22} \\ a_{41} & a_{42} \end{vmatrix} \begin{vmatrix} a_{13} & a_{14} \\ a_{33} & a_{34} \end{vmatrix} + \begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix} \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix} \end{aligned}$$

A titre d'application, remplaçons $\mathcal{C}_3(M)$ et $\mathcal{C}_4(M)$ par $\mathcal{C}_1(M)$ et $\mathcal{C}_2(M)$ respectivement, et appliquons la formule ci-dessus à la matrice obtenue, dont le déterminant est évidemment nul. On obtient, en posant

$$\alpha = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \beta = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}, \gamma = \begin{vmatrix} a_{11} & a_{12} \\ a_{41} & a_{42} \end{vmatrix},$$

$$u = \begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix}, v = \begin{vmatrix} a_{21} & a_{22} \\ a_{41} & a_{42} \end{vmatrix} \quad \text{et} \quad w = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} : 2(\alpha u - \beta v + \gamma w) = 0.$$

Supposant la caractéristique de K différente de 2, on a donc la relation $\alpha u - \beta v + \gamma w = 0$ entre les six mineurs d'ordre 2 de la matrice générale de $\mathfrak{M}_{4,2}(K)$. Ainsi, dans l'algèbre $K[(a_{ij})_{(i,j) \in \llbracket 1,4 \rrbracket \times \llbracket 1,2 \rrbracket}]$, les 6 éléments $\alpha, \beta, \gamma, u, v, w$ ne sont pas algébriquement libres sur K (on conçoit qu'une méthode analogue permettrait d'établir de telles relations polynomiales vérifiées par les mineurs d'une matrice rectangulaire quelconque).

Remarque 2 : Le lecteur se sera certainement aperçu que les démonstrations données dans ce § pour les théorèmes XIII.4.3, XIII.4.4, XIII.4.5, XIII.4.6 et XIII.4.7 ne font intervenir que la structure d'anneau commutatif de K (sans utiliser le fait que cet anneau est intègre).

Exercice 1 : Soit $D = |a_{ij}|_{(i,j) \in \llbracket 1,n \rrbracket^2}$, n entier ≥ 1 et $x \in K$. On considère la matrice M de terme général $a_{ij} + x$.

a) Calculer $\det(M)$.

b) On suppose en particulier que $a_{ii} \neq 0$ pour tout $i \in \llbracket 1,n \rrbracket$, $x \neq 0$ et $a_{ij} \neq 0$ pour $i \neq j$. Montrer que :

$$\begin{vmatrix} a_{11} + x & x & \dots & x \\ x & a_{22} + x & & x \\ \vdots & \vdots & \ddots & \vdots \\ x & x & & a_{nn} + x \end{vmatrix} = a_{11} a_{22} \dots a_{nn} x \left(\frac{1}{x} + \frac{1}{a_{11}} + \dots + \frac{1}{a_{nn}} \right).$$

Exercice 2 : Le corps de base est \mathbb{C} . Etablir la formule

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = (af - be + cd)^2.$$

Exercice 3 : Montrer (sans le calculer) que le déterminant $\begin{vmatrix} 2 & 9 & 9 \\ 4 & 6 & 8 \\ 7 & 4 & 1 \end{vmatrix}$ est un entier divisible par 13.

Exercice 4 : Le corps de base est \mathbb{C} .

a) Calculer

$$D = \det \begin{bmatrix} 0 & c & b & d \\ c & 0 & a & e \\ b & a & 0 & f \\ d & e & f & 0 \end{bmatrix}$$

(par exemple à l'aide de la formule de Laplace).

b) Dans $\mathbb{C}[X, Y, Z]$, factoriser en facteurs de degré 1 le polynôme

$$X^4 + Y^4 + Z^4 - 2(Y^2 Z^2 + Z^2 X^2 + X^2 Y^2).$$

c) En déduire une factorisation en facteurs de degré 2 du déterminant :

$$\Delta = \det \begin{bmatrix} 0 & \gamma^2 & \beta^2 & \lambda^2 \\ \gamma^2 & 0 & \alpha^2 & \mu^2 \\ \beta^2 & \alpha^2 & 0 & \nu^2 \\ \lambda^2 & \mu^2 & \nu^2 & 0 \end{bmatrix}.$$

Exercice 5 : En calculant le produit

$$\begin{vmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{vmatrix} \cdot \begin{vmatrix} x & -y & -z & -t \\ y & x & -t & z \\ z & t & x & -y \\ t & -z & y & x \end{vmatrix}$$

établir l'identité d'Euler :

$$(a^2 + b^2 + c^2 + d^2)(x^2 + y^2 + z^2 + t^2) = (ax + by + cz + dt)^2 + (ay - bx + ct - dz)^2 + (az + bt - cx + dy)^2 + (at + bz - cy - dx)^2.$$

Combien obtient-on d'égalités distinctes quand on permute les éléments de $\{a, b, c, d\}$ et ceux de $\{x, y, z, t\}$, compte tenu du fait qu'un carré a 2 racines (Rép. 96).

Exercice 6 : Combien d'additions et de multiplications doit-on en principe effectuer pour calculer un déterminant d'ordre 10 ?

Exercice 7 : Calculer, à l'aide de la formule de Laplace, les déterminants suivants :

$$\begin{array}{ll} a) \begin{vmatrix} 1 & 1 & 3 & 4 \\ 2 & 0 & 0 & 8 \\ 3 & 0 & 0 & 2 \\ 4 & 4 & 7 & 5 \end{vmatrix} & b) \begin{vmatrix} 2 & 1 & 4 & 3 & 5 \\ 3 & 4 & 0 & 5 & 0 \\ 3 & 4 & 5 & 2 & 1 \\ 1 & 5 & 2 & 4 & 3 \\ 4 & 6 & 0 & 7 & 0 \end{vmatrix} \\ c) \begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 3 \\ 6 & 5 & 7 & 8 & 4 & 2 \\ 9 & 8 & 6 & 7 & 0 & 0 \\ 3 & 2 & 4 & 5 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 5 & 6 & 0 & 0 & 0 & 0 \end{vmatrix} & d) \begin{vmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & a & b & c & d \\ 0 & a^2 & b^2 & c^2 & d^2 \end{vmatrix} \end{array}$$

§ XIII.5 EXEMPLES DE DÉTERMINANTS

Quand on désire calculer le déterminant d'une matrice carrée de $\mathfrak{M}_n(K)$, il suffit de se rappeler quelques principes simples, d'abord que $\det(M) = \det_{\mathcal{B}}(\mathcal{C}_1(M), \dots, \mathcal{C}_n(M))$, \mathcal{B} désignant la base canonique de K^n , et que c'est donc une fonction *n-linéaire et alternée des vecteurs colonnes* $\mathcal{C}_i(M)$. Si donc une colonne (resp. une ligne) est multipliée par un scalaire $\lambda \in K$, $\det(M)$ est multiplié par λ . Si on transpose deux colonnes (resp. deux lignes), le déterminant est multiplié par (-1) . Si à une colonne donnée (resp. une ligne) on ajoute une combinaison linéaire arbitraire des autres colonnes (resp. des autres lignes), le déterminant est inchangé. On se souviendra également que si les vecteurs colonnes (resp. lignes) sont liés, le déterminant $\det(M)$ est nul. On n'hésitera donc pas à utiliser au mieux les *opérations élémentaires* sur la matrice M mais il ne faudra pas trop s'y attarder (sauf à vouloir en faire un usage systématique) et le pl

finira par développer le déterminant par rapport à une ligne ou à une colonne en utilisant le théorème XIII.4.4.

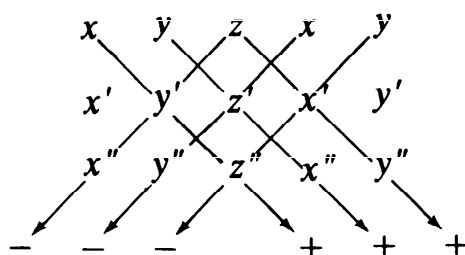
Exemple 1 : Le déterminant d'ordre 1 : $\det [a]$, vaut a .

Le déterminant d'ordre 2 : $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ est égal à $ad - bc$.

Le déterminant d'ordre 3 : $\Delta = \begin{vmatrix} x & y & z \\ x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix}$ s'obtient par développement par rapport à une ligne ou une colonne quelconque, ce qui donne

$$\Delta = xy'z'' + x'y''z + x''yz' - xz'y'' - x'yz'' - x''y'z.$$

On retiendra facilement le résultat en remarquant que Δ est la somme des trois produits des termes situés sur les parallèles à la diagonale principale affectés du signe +, et des trois produits affectés du signe – des termes situés sur les parallèles à la diagonale secondaire dans la matrice (3, 5) obtenue en répétant les deux premières colonnes de Δ :



C'est la « règle de Sarrus », très pratique, dont le seul défaut est de ne s'appliquer qu'aux déterminants d'ordre 2 et 3. Dès que l'ordre de la matrice M est ≥ 4 il n'est plus question de calculer « de tête » $\det (M)$.

Exemple 2 : Le corps de base est \mathbb{C} . Mettre la valeur du déterminant $\det (M)$ ci-après sous forme d'un produit de facteurs de degré 1 :

$$M = \begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix}.$$

Solution : On remplace d'abord, dans M , $\mathcal{C}_1(M)$ par $\sum_{i=1}^4 \mathcal{C}_i(M)$, ce qui ne

change pas la valeur du déterminant :

$$\det (M) = \det (M_1), \quad \text{avec} \quad M_1 = \begin{bmatrix} a+b+c+d & b & c & d \\ a+b+c+d & a & d & c \\ a+b+c+d & d & a & b \\ a+b+c+d & c & & \end{bmatrix}.$$

On peut déjà mettre en facteur $P_1 = a + b + c + d$ dans $\mathcal{C}_1(M)$, d'où

$$\det(M_1) = P_1 \det(M_2), \quad \text{avec} \quad M_2 = \begin{bmatrix} 1 & b & c & d \\ 1 & a & d & c \\ 1 & d & a & b \\ 1 & c & b & a \end{bmatrix}.$$

Il est facile maintenant de faire apparaître des zéros en remplaçant, dans M_2 , $\mathcal{L}_i(M_2)$ par $\mathcal{L}_i(M_2) - \mathcal{L}_1(M_2)$ pour $i \in \{2, 3, 4\}$, ce qui donne :

$$\det(M_2) = \det(M_3) = \det(M_4),$$

avec

$$M_3 = \begin{bmatrix} 1 & b & c & d \\ 0 & & & \\ 0 & & M_4 & \\ 0 & & & \end{bmatrix} \quad \text{et} \quad M_4 = \begin{bmatrix} a-b & d-c & c-d \\ d-b & a-c & b-d \\ c-b & b-c & a-d \end{bmatrix}$$

(la dernière égalité a été obtenue en développant $\det(M_3)$ suivant la première colonne). Dans M_4 , on remplace $\mathcal{C}_1(M_4)$ par $\mathcal{C}_1(M_4) + \mathcal{C}_2(M_4)$, et dans la matrice obtenue, la première colonne contient $P_2 = a - b + d - c$ en facteur, d'où

$$\det(M_4) = P_2 \det(M_5) \quad \text{avec} \quad M_5 = \begin{bmatrix} 1 & d-c & c-d \\ 1 & a-c & b-d \\ 0 & b-c & a-d \end{bmatrix}.$$

On fait apparaître le facteur $P_3 = a - d + b - c$ en remplaçant, dans M_5 , $\mathcal{L}_2(M_5)$ par $\mathcal{L}_2(M_5) - \mathcal{L}_1(M_5) + \mathcal{L}_3(M_5)$, ce qui donne :

$$\det(M_5) = P_3 \det(M_6), \quad \text{avec} \quad M_6 = \begin{bmatrix} 1 & d-c & c-d \\ 0 & 1 & 1 \\ 0 & b-c & a-d \end{bmatrix}.$$

On achève en développant $\det(M_6)$ suivant la première colonne, d'où en récapitulant :

$$\det(M) = (a + b + c + d) (a + b - c - d) (a - b + c - d) (a - b - c + d).$$

Exemple 3 : Déterminant de Vandermonde

Le corps de base K étant quelconque, soit $n \in \mathbb{N}$ ($n \geq 2$), et x_1, x_2, \dots, x_n dans K . Posons

$$V_n = [(x_i)^{j-1}]_{(i,j) \in \llbracket 1, n \rrbracket^2} \quad \text{et} \quad D_n = \det(V_n).$$

Au lieu de faire apparaître des facteurs comme dans l'exemple 2, nous allons calculer D_n en établissant une *relation de récurrence*.

D_{n-1} . On a d'abord $D_2 = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$. Puis, si $n \geq 3$, remplaçons, dans V_n , la ligne $\mathcal{L}_i(V_n)$ par $\mathcal{L}_i(V_n) - \mathcal{L}_n(V_n)$ pour $1 \leq i \leq n-1$. Mettons en facteur $x_i - x_n$ dans la ligne i ($i \leq n-1$) et développons le déterminant obtenu suivant sa première colonne. On obtient :

$$D_n = \left(\prod_{i=1}^{n-1} (x_i - x_n) \right) (-1)^{n+1} \det(M), \quad \text{avec } M \in \mathfrak{M}_{n-1}(K);$$

$$M = \begin{bmatrix} 1 & P_1(x_1) & \dots & P_{n-2}(x_1) \\ 1 & P_1(x_2) & \dots & P_{n-2}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & P_1(x_n) & \dots & P_{n-2}(x_n) \end{bmatrix} = [P_{j-1}(x_i)]_{(i,j) \in \llbracket 0, n-1 \rrbracket^2},$$

où

$$P_k(X) = X^k + X^{k-1}x_n + \dots + Xx_n^{k-1} + x_n^k \quad \text{si } k \geq 2, \quad \text{et } P_0(X) = 1.$$

Or, on passe aisément de M à V_{n-1} par opérations élémentaires ; en effet posons $M_{n-1} = V_{n-1}$, et supposons prouvé que la matrice

$$M_k = \begin{bmatrix} 1 & x_1 & \dots & x_1^{k-1} & P_k(x_1) & \dots & P_{n-2}(x_1) \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 1 & x_{n-1} & \dots & x_{n-1}^{k-1} & P_k(x_{n-1}) & \dots & P_{n-2}(x_{n-1}) \end{bmatrix}$$

soit déduite de M par opérations élémentaires laissant invariant le déterminant (où $k \leq n-2$). En remplaçant, dans M_k , $\mathcal{C}_{k+1}(M_k)$ par

$$\mathcal{C}_{k+1}(M_k) - (x_n \mathcal{C}_k(M_k) + x_n^2 \mathcal{C}_{k-1}(M_k) + \dots + x_n^k \mathcal{C}_1(M_k)),$$

on obtient exactement M_{k+1} .

Comme l'opération $\mathcal{C}_2(M) \rightsquigarrow \mathcal{C}_2(M) - x_n \mathcal{C}_1(M)$ fait passer de M à M_1 , on voit par récurrence que $V_{n-1} = M_{n-1}$ se déduit de M par opérations élémentaires ne changeant pas le déterminant, d'où finalement :

$$(1) \quad D_n = \left[\prod_{i=1}^{n-1} (x_n - x_i) \right] \times D_{n-1},$$

ce qui, à partir de D_2 , donne :

$$(2) \quad \boxed{D_n = \prod_{1 \leq i < j \leq n} (x_j - x_i)}.$$

Remarque 1 : Cette démonstration de (2) a l'avantage de n'utiliser que la structure d'anneau commutatif de K . Indiquons cependant une autre méthode qui, elle, a l'avantage d'être plus rapide. On voit tout d'abord que si $i \mapsto x_i$ n'est pas injective, alors $D_n = 0$ puisque la matrice a au moins deux lignes égales. Supposons donc $i \mapsto x_i$ injective, c'est-à-dire les x_i donnés distincts deux à deux. Soit alors

$$P(X) = \det(V_n(x_1, \dots, x_{n-1}, X)) = D_n(x_1, \dots, x_{n-1}, X),$$

où X est une indéterminée sur K . On raisonne en remplaçant le corps K par son extension $L = K(X)$. Alors $P(X)$ est élément de $K[X]$, et en le développant suivant la dernière ligne de $V_n(x_1, \dots, x_{n-1}, X)$, on voit que $P(X) \in K_{n-1}[X]$, le terme dominant étant $V_{n-1} X^{n-1}$. Mais $P(x_i) = 0$ pour $1 \leq i \leq n-1$, d'où (cf. corollaire 1 du théorème VII.4.2)

$$P(X) = V_{n-1} \prod_{i=1}^{n-1} (X - x_i),$$

et en prenant la valeur de P en x_n , on retrouve (1), et donc (2).

Exemple 4 : Déterminant de matrices circulantes :

Le corps de base est \mathbb{C} . Soit $n \in \mathbb{N}$ ($n \geq 2$) et a_0, \dots, a_{n-1} des éléments de \mathbb{C} . Appelons *matrice circulante* $\Gamma(a_0, \dots, a_{n-1})$ la matrice $[a_{j-i}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$, où, pour $k \in \mathbb{Z}$, \bar{k} désigne le *reste mod* (n) de k ($\bar{k} \in \llbracket 0, n-1 \rrbracket$), ce qui donne :

$$\Gamma(a_0, \dots, a_{n-1}) = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}$$

(chaque ligne de $\Gamma(a_0, \dots, a_{n-1})$ se déduit de la précédente par la même *permutation circulaire*).

Notons ζ une racine primitive n -ième de 1, et utilisons la matrice ⁽¹⁾

$$\Omega(\zeta) = [\zeta^{(i-1)(j-1)}]_{(i,j) \in \llbracket 1, n \rrbracket^2},$$

qui n'est autre que la matrice de Vandermonde $V_n(1, \zeta, \zeta^2, \dots, \zeta^{n-1})$. D'après l'exemple 3, on a : $\det(\Omega(\zeta)) \neq 0$ puisque $1, \zeta, \dots, \zeta^{n-1}$ sont distincts.

⁽¹⁾ La matrice $\Omega(e^{2i\pi/n})$ est parfois appelée *matrice alternante* d'ordre n

Le produit $\Gamma(a_0, \dots, a_{n-1}) \Omega(\zeta) = [c_{ij}]$ se calcule facilement :

$$(\forall i, \forall j) \quad c_{ij} = \sum_{k=1}^n a_{\overline{k-i}} \zeta^{(k-1)(j-1)}.$$

En remarquant que l'application $\llbracket 1, n \rrbracket \rightarrow \llbracket 0, n-1 \rrbracket, k \mapsto \overline{k-i}$ est bijective, cela s'écrit :

$$c_{ij} = \sum_{q=0}^{n-1} a_q \zeta^{(q+i-1)(j-1)} = \zeta^{(i-1)(j-1)} L_j,$$

où $L_j = \sum_{q=0}^{n-1} a_q \zeta^{(j-i)q}$. On a ainsi prouvé que :

$$\Gamma(a_0, \dots, a_{n-1}) \Omega(\zeta) = \Omega(\zeta) \times \text{Diag}(L_1, L_2, \dots, L_n),$$

d'où

$$\det(\Gamma(a_0, \dots, a_{n-1})) \det(\Omega(\zeta)) = \det(\Omega(\zeta)) \times L_1 L_2 \dots L_n,$$

et puisque $\det(\Omega(\zeta)) \neq 0$,

$$\det(\Gamma(a_0, \dots, a_{n-1})) = L_1 L_2 \dots L_n = \prod_{\lambda=0}^{n-1} \left(\sum_{q=0}^{n-1} \zeta^{q\lambda} a_q \right).$$

Exemple 5 : Formules de Jacobi

Soit $M = [a_{ij}] \in \mathfrak{M}_n(K)$ avec $n \geq 2$. Nous allons calculer les mineurs de la comatrice $\check{M} = [\check{a}_{ij}]$. Avec les notations du § XIII.4, montrons, pour $p \in \llbracket 1, n-1 \rrbracket$, la relation

$$(3) \quad \forall I \in \mathcal{F}_p(n), \forall J \in \mathcal{F}_p(n), \quad \Delta_{I,J}(\check{M}) = \varepsilon(\sigma_I) \varepsilon(\sigma_J) (\det(M))^{p-1} \Delta_{\tilde{I}, \tilde{J}}(M).$$

Pour cela observons d'abord que (3) est évidente si $p = 1$.

Si $p \geq 2$, les deux membres sont des fonctions polynomiales de la famille (a_{ij}) . Quitte, s'il le faut, à remplacer le corps K par son extension infinie $K(X)$, on peut supposer K infini.

La fonction polynomiale $M \mapsto \det(M)$ sur $\mathfrak{M}_n(K)$ n'est pas nulle, puisque $\det(I_n) = 1_K \neq 0$. D'après le théorème X.1.4, il suffit donc d'établir (3) dans l'hypothèse supplémentaire où $\det(M) \neq 0$, que nous postulons donc.

Remarquons encore que le cas où $p = n$ est conséquence du théorème XIII.4.5

$$(I), \text{ puisqu'alors} \quad \det(\check{M}) = \det(\tilde{M})$$

$$\text{et} \quad \det(M\tilde{M}) = \det(M) \det(\tilde{M}) = \det[\det(M) I_n] = (\det(M))^n,$$

d'où, après division par $\det(M)$: $\det(\tilde{M}) = [\det(M)]^{n-1}$, ce qui est bien (3) lorsque $p = n$. Il reste donc à prouver (3) sous les hypothèses suivantes : K est infini, $n \geq 2$, $2 \leq p \leq n-1$, et $\det(M) \neq 0$.

Premier cas : $I = J = \llbracket 1, p \rrbracket$. Posons alors $\tilde{M} = [\tilde{a}_{ij}]$ ($\tilde{a}_{ij} = \check{a}_{ji}$). On a : $\Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(\check{\tilde{M}}) = \Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(\tilde{M})$. Introduisons la matrice

$$N = \left[\begin{array}{c|c} U & V \\ \hline 0 & I_{n-p} \end{array} \right],$$

où $[U \mid V]$ est la matrice à p lignes et n colonnes formée avec les p premières lignes de \tilde{M} . Utilisant le théorème XIII.4.5 (I) on constate que le produit NM est égal à

$$NM = \left[\begin{array}{c|c} \det(M) I_p & 0 \\ \hline R & S \end{array} \right]$$

où $[R \mid S]$ est la matrice à $(n-p)$ lignes et n colonnes formée avec les $(n-p)$ dernières lignes de M . On en déduit :

$$\det(N) \det(M) = \det(NM) = [\det(M)]^p \det(S) = \det(N) \det(M) = \det(U) \det(M).$$

Mais $\det(U) = \Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(\tilde{M})$ et $\det(S) = \Delta_{\llbracket p+1, n \rrbracket, \llbracket p+1, n \rrbracket}(M)$. Donc, après division par $\det(M) \neq 0$, il s'ensuit bien :

$$\Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(\tilde{M}) = (\det(M))^{p-1} \Delta_{\llbracket p+1, n \rrbracket, \llbracket p+1, n \rrbracket}(M).$$

Cas général : Soit P la matrice

$$[a_{\sigma_I(i), \sigma_J(j)}]_{(i,j) \in \llbracket 1, n \rrbracket^2} = [b_{ij}],$$

d'où : $\det(P) = \varepsilon(\sigma_I) \varepsilon(\sigma_J) \det(M)$. Posons $\check{P} = [\check{b}_{ij}]$. Lorsque i est fixé quelconque, on a $\det(P) = \sum_{j=1}^n b_{ij} \check{b}_{ij}$, $\det(M) = \sum_{k=1}^n a_{\sigma_I(i), k} \check{a}_{\sigma_I(i), k} = \sum_{j=1}^n b_{ij} \check{a}_{\sigma_I(i), \sigma_J(j)}$.

Par identification de ces expressions considérées comme formes linéaires en $(b_{i1}, b_{i2}, \dots, b_{in})$ il s'ensuit :

$$\forall j \in \llbracket 1, n \rrbracket \quad \check{b}_{ij} = \varepsilon(\sigma_I) \varepsilon(\sigma_J) \check{a}_{\sigma_I(i), \sigma_J(j)}$$

et c'est vrai pour tous i et j .

$$\text{Or,} \quad \Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(\check{\tilde{M}}) = [\varepsilon(\sigma_I) \varepsilon(\sigma_J)]^p \Delta_{\llbracket 1, p \rrbracket, \llbracket 1, p \rrbracket}(Q),$$

(où $Q = [\check{a}_{\sigma_I(i), \sigma_J(j)}]) = [\varepsilon(\sigma_I) \varepsilon(\sigma_J)]^p \Delta_{I,J}(\check{\tilde{M}}) =$ (à cause du premier cas)
 $= (\det(P))^{p-1} \Delta_{\llbracket p+1, n \rrbracket, \llbracket p+1, n \rrbracket}(P) = [\varepsilon(\sigma_I) \varepsilon(\sigma_J) \det(M)]^{p-1} \Delta_{\tilde{I}, \tilde{J}}(M)$,
 d'où en rapprochant les deux expressions trouvées :

$$\varepsilon(\sigma_I) \varepsilon(\sigma_J) \Delta_{I,J}(\check{\tilde{M}}) = (\det(M))^{p-1} \Delta_{\tilde{I}, \tilde{J}}(M),$$

ce qui prouve (3).

Expression de (3) à l'aide de la matrice $\Lambda^{n-1}(M)$.

Montrons (4) :

$$(4) \quad \forall I \in \mathcal{F}_p(n), \quad \forall J \in \mathcal{F}_p(n) \quad \Delta_{I,J}(\Lambda^{n-1}(M)) = (\det(M))^{p-1}$$

Pour cela, posons $\Lambda^{n-1}(M) = [c_{ij}]$, d'où $\check{M} = [(-1)^{i+j} c_{ij}]$.

Pour $I \in \mathcal{F}_p(n)$, $J \in \mathcal{F}_p(n)$, il s'ensuit :

$$\Delta_{I,J}(\check{M}) = (-1)^{s+t} \Delta_{I,J}(\Lambda^{n-1}(M)),$$

avec $s = \sum_{i \in I} i$ et $t = \sum_{j \in J} j$.

Mais, si $I = \{i_1, i_2, \dots, i_p\}$, $i_1 < i_2 < \dots < i_p$, il est clair que

$$\nu(I, \tilde{I}) = i_1 + i_2 + \dots + i_p - \frac{p(p+1)}{2} = s - \frac{p(p+1)}{2},$$

et de même $\nu(J, \tilde{J}) = t - \frac{p(p+1)}{2}$, d'où

$$(-1)^{s+t} = (-1)^{\nu(I, \tilde{I}) + \nu(J, \tilde{J})} = \varepsilon(\sigma_I) \varepsilon(\sigma_J).$$

En utilisant (3), on en déduit bien (4).

Exemple 6 : Le corps de base est $K = \mathbb{C}$. On donne $n \in \mathbb{N}$ ($n \geq 2$), a et b distincts dans \mathbb{C} et des éléments x_1, \dots, x_n de \mathbb{C} . On se propose de calculer $D = \det [a_{ij}] = \det (M)$, avec $a_{ij} = a$ pour $i > j$, $a_{ij} = b$ pour $i < j$ et $a_{ii} = x_i$ pour $i \in \llbracket 1, n \rrbracket$.

Solution : Désignons par S la matrice carrée d'ordre n dont tous les termes sont égaux à 1. Notant t une variable dans \mathbb{C} , et $f(t)$ le polynôme $\prod_{i=1}^n (x_i - t)$, soit $\Delta(t)$ la fonction $\det (M + tS)$. On voit que $\Delta(t)$ est

polynomiale, de degré ≤ 1 (par exemple, en retranchant la première colonne à toutes les autres, et en développant ensuite suivant la première colonne). Donc, $\Delta(t) = \alpha t + \beta$, avec $\beta = \Delta(0) = D$.

Mais il est clair que $\Delta(-a) = f(a)$ et $\Delta(-b) = f(b)$, d'où $\alpha a - \beta = -f(a)$, $\alpha b - \beta = -f(b)$, ce qui fournit α et β puisque $a \neq b$:
 $\alpha = -\frac{f(b) - f(a)}{b - a}$, $\beta = \frac{bf(a) - af(b)}{b - a}$. En particulier, $D = \beta$.

Exercice 1 : Le corps de base est \mathbb{C} . On donne $(s_1, s_2, \dots, s_n) \in \mathbb{C}^n$ ($n \geq 2$). Calculer :

$$\Delta_n = \det \begin{bmatrix} s_1 & \dots & \dots & s_1 \\ \vdots & s_2 & \dots & s_2 \\ \vdots & \vdots & \ddots & \vdots \\ s_1 & s_2 & \dots & s_n \end{bmatrix}.$$

Expliciter la valeur de Δ_n pour $s_i = \sum_{j=0}^i j$ ($1 \leq i \leq n$).

Exercice 2 : Calculer $\det [|i - j|]_{(i,j) \in \llbracket 1, n \rrbracket^2}$.

Exercice 3 : Calculer $\det \left[\binom{i}{j-1} \right]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ (on rappelle que $\binom{i}{k} = 0$ dès que $k > i$). Réponse : $D = 1$.

Exercice 4 : Calculer $\det \left[\binom{i+j-2}{i-1} \right]_{(i,j) \in \llbracket 1, n \rrbracket^2}$. Réponse : $D = 1$.

Exercice 5 : Calculer $\det [(i+j-1)^2]_{(i,j) \in \llbracket 1, n \rrbracket^2}$.

Exercice 6 : Le corps de base est \mathbb{C}

a) pour $(a_1, a_2, \dots, a_n) \in \mathbb{C}^n$, calculer $\det (M_n)$ où $M_n = [b_{ij}]$, avec $b_{ij} = 1 + \delta_{ij} a_i$ (δ = symbole de Kronecker).

b) Etudier le cas particulier où $a_1 = a_2 = \dots = a_n = a$.

c) On développe le déterminant D_n suivant : $D_n = \det [c_{ij}]$ où $c_{ii} = -x_{ii}$ pour $i \in \llbracket 1, n \rrbracket$ et $c_{ij} = x_{ij}$ pour $i \neq j$, les (x_{ij}) étant des indéterminées sur \mathbb{C} . Le développement est du type $\sum \varepsilon \mu(x)$, où les $\mu(x)$ sont des monômes en les x_{ij} et $\varepsilon \in \{-1, +1\}$. Calculer le nombre de termes de ce développement pour lesquels $\varepsilon = +1$.

Exercice 7 : On suppose acquise la formule de l'exemple 4 donnant la factorisation du déterminant d'une matrice circulante sur \mathbb{C} .

a) En déduire, pour $(x, y, z) \in \mathbb{C}^3$:

$$x^3 + y^3 + z^3 - 3xyz = (x + y + z)(x + jy + j^2z)(x + j^2y + jz).$$

b) Pour $n \geq 3$ et $x \in \mathbb{C}$, calculer le déterminant circulant :

$$D_n = \det \begin{bmatrix} 0 & 1 & 2x & 3x^2 & \dots & (n-1)x^{n-2} \\ (n-1)x^{n-2} & 0 & 1 & 2x & \dots & (n-2)x^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2x & 3x^2 & \dots & \dots & 0 \end{bmatrix}$$

c) Calculer sous forme de nombre rationnel le déterminant circulant $\Gamma(1, 2, \dots, n)$, avec les notations de l'exemple 4.

d) En utilisant un déterminant circulant, démontrer, dans $\mathbb{C}[X, Y, Z]$,

$$\prod_{\zeta \in \mu_5} (X + \zeta Y + \bar{\zeta} Z) = X^5 + Y^5 + Z^5 - 5XYZ(X^2 - YZ).$$

Exercice 8 : Soit

$$A = \begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix} \quad \text{et} \quad E = \begin{bmatrix} e & f & g & h \\ f & e & h & g \\ g & h & e & f \\ h & g & f & e \end{bmatrix}.$$

Décomposer $\det(A)$ en produit de 4 facteurs linéaires.

On considère la matrice $M = \begin{bmatrix} A & E \\ E & A \end{bmatrix}$. Décomposer $\det(M)$ en produit de 8 facteurs linéaires du type $a \pm b \pm \dots \pm h$.

Exercice 9 : On suppose connu le déterminant de Vandermonde (cf. exemple 3)

$$\det [(x_i)^{j-1}]_{(i,j) \in \llbracket 1, n \rrbracket^2} = \prod_{i < j} (x_j - x_i),$$

où les (x_i) sont dans \mathbb{C} .

a) Soit $P_k(X) \in \mathbb{C}_k[X]$ un polynôme *normalisé* de degré k pour $k \in \llbracket 0, n-1 \rrbracket$, avec $n \geq 2$. Donc

$$P_0 = 1, \dots, P_k = X^k + a_{k1}X^{k-1} + \dots + a_{kk}.$$

Démontrer que $\det [P_{j-1}(x_i)]_{(i,j) \in \llbracket 1, n \rrbracket^2} = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

b) Calculer ensuite $\det \left[\binom{x_i}{j-1} \right]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ où $\binom{x_i}{k}$ désigne le coefficient binomial généralisé $\frac{x_i(x_i-1)\dots(x_i-k+1)}{k!}$.

c) En déduire que si $(a_1, a_2, \dots, a_n) \in \mathbb{Z}^n$, avec $a_1 < a_2 < \dots < a_n$, alors le rationnel $\prod_{1 \leq i < j \leq n} \frac{a_j - a_i}{j - i}$ est un entier.

Exercice 10 : Soit $(x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ ($n \geq 1$). On considère la matrice

$$M = [1 + (x_i)^j]_{(i,j) \in \llbracket 1, n \rrbracket^2}.$$

Montrer que $\det(M) = \left[2x_1 \dots x_n - \prod_{k=1}^n (x_k - 1) \right] \prod_{i < j} (x_j - x_i)$.

Exercice 11 : Soit $n \in \mathbb{N}$ ($n \geq 3$), $p \in \llbracket 1, n-1 \rrbracket$ et $(x_1, \dots, x_n) \in \mathbb{C}^n$. Calculer $\det(V_{n,p})$,

avec
$$V_{n,p} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{p-1} & x_1^{p+1} & \dots & x_1^n \\ \vdots & & & & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^{p-1} & x_n^{p+1} & \dots & x_n^n \end{bmatrix}.$$

Exercice 12 : On donne $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$, $n \geq 2$ et $(x_1, \dots, x_n) \in \mathbb{C}^n$. Pour $i \in \llbracket 1, n \rrbracket$,

soit
$$P_i(X) = \sum_{k=1}^n a_{ik} X^{k-1} \in \mathbb{C}[X],$$

et soit enfin W la matrice $[(x_j)^{i-1}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$. Calculer MW . En déduire, pour $(a_1, \dots, a_n) \in \mathbb{C}^n$, la valeur du déterminant de $[(a_i + x_j)^{n-1}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$.

Exercice 13 : Soit $(x, y) \in \mathbb{C}^2$ et M la matrice

$$M = \begin{bmatrix} 1 & x & x^2 & x^3 & x^4 \\ 1 & 2x & 3x^2 & 4x^3 & 5x^4 \\ 1 & 4x & 9x^2 & 16x^3 & 25x^4 \\ 1 & y & y^2 & y^3 & y^4 \\ 1 & 2y & 3y^2 & 4y^3 & 5y^4 \end{bmatrix}.$$

Montrer : $\det(M) = 2x^3y(x-y)^6$.

Exercice 14 : a) Soit $\theta \in \mathbb{R}$. On pose $D_0 = \Delta_0 = 1$, $D_1 = 2 \cos \theta$, $\Delta_1 = \cos \theta$ et, pour $n \geq 2$, $D_n = \det[a_{ij}]$, $\Delta_n = \det[b_{ij}]$, avec $a_{ii} = 2 \cos \theta$ pour $1 \leq i \leq n$, $a_{ij} = 1$ pour $|i-j| = 1$, $a_{ij} = 0$ pour les autres couples ; $b_{11} = \cos \theta$, $b_{ii} = 2 \cos \theta$ pour $2 \leq i \leq n$, $b_{ij} = 1$ pour $|i-j| = 1$, $b_{ij} = 0$ pour tous les autres couples de $\llbracket 1, n \rrbracket^2$. Démontrer que $D_n = \frac{\sin n\theta}{\sin \theta}$ et $\Delta_n = \cos n\theta$.

Indication : Les suites (D_n) et (Δ_n) satisfont à une relation de récurrence $(\forall n \geq 2) \quad u_n = 2u_{n-1} \cos \theta - u_{n-2}$ que l'on résoudra soit en introduisant la série formelle $S(t) = \sum_{n \geq 0} u_n t^n \in \mathbb{C}[[t]]$, soit en cherchant une base de l'espace vectoriel (sur \mathbb{C}) des solutions.

b) Soit $(a, b, c) \in \mathbb{C}^3$ avec $bc \neq 0$. On pose $D_0 = \Delta_0 = 1$, $D_1 = 2a$, $\Delta_1 = a$ et pour $n \geq 2$, $D_n = \det[a_{ij}]$, $\Delta_n = \det[b_{ij}]$, avec : $a_{i,i+1} = b_{i,i+1} = b$ pour $1 \leq i \leq n-1$, $a_{j+1,j} = b_{j+1,j} = c$ pour $1 \leq j \leq n-1$, $a_{11} = 2a$, $b_{11} = a$, $a_{ii} = b_{ii} = 2a$ pour $2 \leq i \leq n$, tous les autres coefficients étant nuls. Montrer que les suites (Δ_n) et (D_n) satisfont à la relation de récurrence :

$$(\forall n \geq 2) \quad u_n = 2au_{n-1} - bcu_{n-2}$$

et en déduire l'expression de Δ_n et de D_n comme polynômes en α, β , à coefficients dans \mathbb{Z} ou sous la forme $\rho^n \cos n\theta$ et $\rho^n \frac{\sin(n+1)\theta}{\sin \theta}$, α, β, ρ et θ étant convenable

Applications numériques : $2a = 3, b = 2, c = 1$ ou $2a = b = c = 1$.

Exercice 15 : Le terme général d'une matrice $M \in \mathfrak{M}_n(K)$ vérifie la relation $a_{ij} = a_{i-1,j} + a_{i,j-1}$ permettant de calculer tous les éléments connaissant ceux de la première ligne et de la première colonne. On donne par exemple $a_{i1} = a_{1i} = i + 1$. Calculer $\det(M)$.

Exercice 16 : Mettre sous forme de produits les déterminants :

$$a) \begin{vmatrix} 1 & \cos x & \cos 2x \\ 1 & \cos y & \cos 2y \\ 1 & \cos z & \cos 2z \end{vmatrix} \quad b) \begin{vmatrix} 1 & \cos x & \cos 3x \\ 1 & \cos y & \cos 3y \\ 1 & \cos z & \cos 3z \end{vmatrix} \quad c) \begin{vmatrix} 1 & x^2(y-z)^2 & x^n \\ 1 & y^2(z-x)^2 & y^n \\ 1 & z^2(x-y)^2 & z^n \end{vmatrix}.$$

Exercice 17 : Soit $n \in \mathbb{N}$ ($n \geq 2$). On désigne par φ l'indicateur d'Euler sur \mathbb{N}^* (cf. § IV.5).

a) Démontrer que $\det ([\text{pgcd}(i, j)]_{(i, j) \in \llbracket 1, n \rrbracket^2}) = \varphi(1) \varphi(2) \dots \varphi(n)$.

b) Généraliser au calcul de $\det ([(\text{pgcd}(i, j))^\lambda]_{(i, j) \in \llbracket 1, n \rrbracket^2})$ pour $\lambda \in \mathbb{C}$.

Exercice 18 : Pour $n \in \mathbb{N}^*$, on donne $(c_1, c_2, \dots, c_{n+1}) \in \mathbb{C}^{n+1}$ et des nombres complexes p, q, b, c . Calculer $\det(M)$ où $M = [a_{ij}] \in \mathfrak{M}_{n+1}(\mathbb{C})$, $a_{ii} = c_i$ pour $1 \leq i \leq n+1$, $a_{ij} = p$ pour $j < i \leq n$, $a_{ij} = q$ pour $i < j \leq n$, $a_{n+1,j} = b$ et $a_{j,n+1} = c$ pour $j \in \llbracket 1, n \rrbracket$.

Exercice 19 (matrices circulantes par blocs).

a) Soit $n \in \mathbb{N}^*$ et M et N dans $\mathfrak{M}_n(\mathbb{C})$. On considère la matrice $P \in \mathfrak{M}_{2n}(\mathbb{C})$ qui s'écrit, par blocs de $\mathfrak{M}_n(\mathbb{C})$: $P = \begin{bmatrix} M & N \\ N & M \end{bmatrix}$. Démontrer : $\det(P) = \det(M+N) \det(M-N)$ (cf. exercice 8).

b) Plus généralement, soit n et p deux entiers, $n \geq 2, p \geq 2$. On donne $A_0, A_1, \dots, A_{p-1} \in \mathfrak{M}_n(\mathbb{C})$ et on considère la matrice $M \in \mathfrak{M}_{np}(\mathbb{C})$ qui s'écrit, par blocs carrés d'ordre n :

$$M = \begin{bmatrix} A_0 & A_1 & \dots & A_{p-1} \\ A_{p-1} & A_0 & \dots & A_{p-2} \\ \dots & \dots & \dots & \dots \\ A_1 & A_2 & \dots & A_0 \end{bmatrix}.$$

Pour $\zeta \in \mu_p$, soit $M_\zeta = \sum_{q=0}^{p-1} \zeta^q A_q$. Démontrer : $\det(M) = \prod_{\zeta \in \mu_p} \det(M_\zeta)$.

Indication : On adaptera « par blocs » la méthode de l'exemple 4.

Exercice 20 : Soit m, n dans \mathbb{N}^* . On donne des matrices $A_{ij} \in \mathfrak{M}_m(K)$ ($(i, j) \in \llbracket 1, n \rrbracket^2$) deux à deux permutables. Soit $M \in \mathfrak{M}_{mn}(K)$ la matrice qui s'écrit, par blocs carrés d'ordre m : $M = [A_{ij}]_{(i, j) \in \llbracket 1, n \rrbracket^2}$. Démontrer :

$$\det(M) = \det \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) A_{\sigma(1),1} A_{\sigma(2),2} \dots A_{\sigma(n),n} \right).$$

Exercice 21 : On donne $m \in \mathbb{N}^*$.

a) Soit $M = [a_{ij}] \in \mathfrak{M}_{2m}(\mathbb{C})$ telle que $\forall (i, j) \quad a_{ij} = a_{2m+1-i, 2m+1-j}$.

Mettre $\det(M)$ sous forme du produit de deux déterminants d'ordre m .

b) Soit $M = [a_{ij}] \in \mathfrak{M}_{2m+1}(\mathbb{C})$ avec $\forall (i, j) \quad a_{ij} = a_{2m+2-i, 2m+2-j}$. Mettre $\det(M)$ sous forme du produit de deux déterminants, l'un d'ordre m , l'autre d'ordre $m+1$.

Exercice 22 (déterminants circulants gauches).

Soit $(a_0, a_1, \dots, a_{n-1}) \in \mathbb{C}^n$ ($n \geq 2$). On considère $D_n = \det(\Gamma'_n)$, où

$$\Gamma'_n = \begin{bmatrix} a_0 & a_1 & \dots & \dots & a_{n-1} \\ -a_{n-1} & a_0 & \dots & \dots & a_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & -a_2 & \dots & -a_{n-1} & a_0 \end{bmatrix}.$$

- a) Si n est impair, ramener D_n à un déterminant circulant ordinaire.
 b) Si $n = 2m$, $m \in \mathbb{N}^*$, utiliser l'exercice 21a pour prouver que le déterminant circulant $\det(\Gamma(b_0, \dots, b_{2m-1}))$, où $b_i = a_i$ si $0 \leq i \leq m-1$, et $b_i = 0$ si $i \geq m$ s'exprime simplement à l'aide de $\det(\Gamma(a_0, a_1, \dots, a_{n-1}))$ et de $\det(\Gamma'_n(a_0, a_1, \dots, a_{n-1}))$. En déduire une méthode de calcul de ce dernier déterminant.
 c) Application numérique : $a_i = i + 1$ pour $0 \leq i \leq n-1$.

Exercice 23 : Soit m et n dans \mathbb{N}^* , $m < n$. On donne $M \in \mathfrak{M}_{m,n}(K)$ et $N \in \mathfrak{M}_{n,m}(K)$ ($M = [a_{ij}]$, $N = [b_{ij}]$). Soit P la matrice produit.

a) Démontrer :

$$\det(P) = \sum_{1 \leq r_1 < r_2 < \dots < r_m \leq n} \Delta_{[1, m], \{r_1, \dots, r_m\}}(M) \Delta_{\{r_1, \dots, r_m\}, [1, m]}(N)$$

(on raisonnera comme dans le calcul du déterminant de la matrice produit).

b) Pour $M \in \mathfrak{M}_{m,n}(K)$, en déduire $\det(M^t M)$.

c) Soit $n \in \mathbb{N}$, $n \geq 2$, et $(p, q) \in (\mathbb{N}^*)^2$ tels que $p + q = n$. On suppose $K = \mathbb{R}$. On donne $M = [A : B] \in \mathfrak{M}_n(\mathbb{R})$, avec $A \in \mathfrak{M}_{n,p}(\mathbb{R})$ et $B \in \mathfrak{M}_{n,q}(\mathbb{R})$. Démontrer :

$$(\det(M))^2 \leq \det(A^t A) \times \det(B^t B).$$

d) Soit $M \in \mathfrak{M}_n(\mathbb{R})$, de colonnes C_1, C_2, \dots, C_n . Démontrer que

$$(\det(M))^2 \leq \prod_{i=1}^n (C_i^t C_i) \quad (\text{inégalité de Hadamard}).$$

Exercice 24 : Le corps K est supposé de caractéristique $p \geq 3$. Soit a_0, a_1, \dots, a_{p-1} éléments de K . Montrer que le déterminant circulant $\Delta_p = \det(\Gamma(a_0, a_1, \dots, a_{p-1}))$ est égal à $\sum_{i=0}^{p-1} (a_i)^p$.

N.B. Cet exercice a reçu de nombreuses solutions. En voici une : on identifie $[0, p-1]$ à $\mathbb{Z}/p\mathbb{Z}$ et $\mathfrak{S}_{\mathbb{Z}/p\mathbb{Z}}$ à $\mathfrak{S}_{[0, p-1]}$. On note

$$E = \{(\alpha_0, \dots, \alpha_{p-1}) \in \mathbb{N}^p \mid \alpha_0 + \alpha_1 + \dots + \alpha_{p-1} = p\}$$

et, pour $\alpha = (\alpha_0, \dots, \alpha_{p-1}) \in E$, $J_\alpha = \left\{ \sigma \in \mathfrak{S}_{\mathbb{Z}/p\mathbb{Z}} \mid a_0^{\alpha_0} \dots a_{p-1}^{\alpha_{p-1}} = \prod_{i=0}^{p-1} a_{\bar{i}-\sigma(\bar{i})} \right\}$.

a) Etablir $\Delta_p = \sum_{\alpha \in E} c_\alpha a_0^{\alpha_0} \dots a_{p-1}^{\alpha_{p-1}}$ avec $c_\alpha = \sum_{\sigma \in J_\alpha} \varepsilon(\sigma)$.

b) Faire opérer le groupe $(\mathbb{Z}/p\mathbb{Z}, +)$ à droite sur $\mathfrak{S}_{\mathbb{Z}/p\mathbb{Z}}$ par la loi $(\sigma, \bar{i}) \mapsto \sigma * \bar{i}$ avec $\sigma * \bar{i}(\bar{k}) = \bar{i} + \sigma(\bar{k} - \bar{i})$ ($k \in \mathbb{Z}/p\mathbb{Z}$) et montrer que chaque J_α est union d'orbites.

c) Soit G le groupe engendré dans $\mathfrak{S}_{\mathbb{Z}/p\mathbb{Z}}$ par la translation $\bar{i} \mapsto \bar{i} + 1$; montrer : si $\sigma \in G$, $\text{Orb}(\sigma) = \{\sigma\}$; et si $\sigma \in \mathfrak{S}_{\mathbb{Z}/p\mathbb{Z}} \setminus G$, alors $\text{card}(\text{Orb}(\sigma)) = p$. Vérifier en outre que $\varepsilon(\sigma)$ est constant sur une même orbite, et en déduire que $\sum_{\sigma \in C} \varepsilon(\sigma) = 0$ si C est une orbite de cardinal p .

d) A l'aide de a) en déduire que $\Delta_p = \sum_{i=0}^{p-1} (a_i)^p$. (Voir aussi exercice 8 du § XV.2.)

Exercice 25 : Soit $(n, p) \in \mathbb{N}^2$, $1 \leq p < n$. On donne $M \in \mathfrak{M}_{n,p}(K)$ et $N \in \mathfrak{M}_{p,n}(K)$. Montrer : $(\forall x \in K) \det(MN + xI_n) = x^{n-p} \det(NM + xI_p)$.

Exercice 26 : a) Soit $(a_1, \dots, a_n) \in K^n$ et $(b_1, \dots, b_n) \in K^n$ ($n \geq 2$), avec $\forall (i, j), a_i + b_j \neq 0$. Démontrer :

$$\det \left(\left[\frac{1}{a_i + b_j} \right]_{(i,j) \in [1, n]^2} \right) = \left[\prod_{1 \leq i < j \leq n} (a_j - a_i)(b_j - b_i) \right] / \prod_{(i,j) \in [1, n]^2} (a_i + b_j)$$

(déterminant de Cauchy). Acheter le calcul pour $a_i = b_i = i$ ($1 \leq i \leq n$).

b) Soit $(\rho_1, \rho_2, \dots, \rho_{n+1}) \in (\mathbb{C}^*)^{n+1}$. On pose, pour $1 \leq i \leq n$,

$$x_{ii} = \frac{(\rho_i + \rho_{n+1})^2}{\rho_i^2 \rho_{n+1}^2}, \quad \text{et pour } i \neq j, \quad x_{i,j} = \frac{(\rho_i + \rho_{n+1})(\rho_j + \rho_{n+1}) - 2\rho_{n+1}^2}{\rho_i \rho_j \rho_{n+1}^2}.$$

Démontrer : $\det ([x_{ij}]_{(i,j) \in \llbracket 1, n \rrbracket^2}) = \frac{2^{n-1}}{\rho_1^2 \rho_2^2 \dots \rho_{n+1}^2} \left[\left(\sum_{i=1}^{n+1} \rho_i \right)^2 - (n-1) \sum_{i=1}^{n+1} \rho_i^2 \right].$

c) *Application géométrique* : Soit E un espace euclidien de dimension $n-1 \geq 2$. On donne des sphères $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{n+1}$ dans E telles que chacune d'elles est tangente à toutes les autres en des points distincts. On note $\Omega_1, \dots, \Omega_{n+1}$ leurs centres, R_1, \dots, R_{n+1} leurs rayons et $\rho_1, \dots, \rho_{n+1}$ leurs courbures $\left(\rho_i = \frac{1}{R_i} \right)$. Prouver d'abord que : ou bien toutes les sphères sont tangentes extérieurement, ou bien l'une contient toutes les autres. Si on est dans le premier cas, les R_i seront pris > 0 , et dans le second cas, si on suppose que c'est la sphère \mathcal{S}_{n+1} qui contient toutes les autres, son rayon sera pris < 0 . Poser ensuite $\vec{V}_i = \frac{\vec{\Omega}_{n+1}\vec{\Omega}_i}{\Omega_{n+1}^2}$ ($1 \leq i \leq n$). On adopte les notations du b). Prouver que $(\vec{V}_i | \vec{V}_j) = (x_{ij})$ ($(i, j) \in \llbracket 1, n \rrbracket^2$).

En s'aidant du résultat du b), montrer la relation : $\left(\sum_{i=1}^{n+1} \rho_i \right)^2 = (n-1) \sum_{i=1}^{n+1} \rho_i^2$. (Référence : W. S. Brown, in Math Monthly, Juin 1969.)

Exercice 27 : On donne $n \in \mathbb{N}^*$.

a) En reprenant la preuve du théorème XI.5.1, montrer que le groupe $SL(n, K)$ est engendré par les matrices de transvection.

b) Soit $f : GL(n, K) \rightarrow K^*$ une fonction polynomiale qui soit en même temps un homomorphisme de groupes, le corps K étant supposé infini. On reprend les notations du § XI.5.

• Si $\lambda \in K^*$, on pose $P(\lambda) = f(D_n(\lambda))$. Montrer que P est polynomiale et vérifie $P(\lambda\mu) = P(\lambda)P(\mu)$ pour tous λ et μ . En déduire la forme de P .

• Si $(i, j) \in \llbracket 1, n \rrbracket^2$, avec $i \neq j$ et $\lambda \in K$, soit $Q_{ij}(\lambda) = f(T_{ij}(\lambda))$. Montrer que Q_{ij} est polynomiale, vérifie $Q_{ij}(\lambda + \mu) = Q_{ij}(\lambda)Q_{ij}(\mu)$ pour tous λ et μ , et en déduire $Q_{ij}(\lambda) = 1$ pour tout λ .

• En déduire que, pour un certain $d \in \mathbb{N}$, on a : $f(M) = (\det(M))^d$ pour toute $M \in GL(n, K)$ en utilisant le théorème XI.5 et la remarque qui le suit.

c) Lorsque K est fini, soit $f : GL(n, K) \rightarrow K^*$ un homomorphisme de groupes quelconque. Utiliser le fait que le groupe K^* est cyclique (cf. théorème VII.5.2 et son corollaire) pour prouver que le résultat du b) demeure.

Exercice 28 : On suppose $K = \mathbb{C}$. On donne $n \in \mathbb{N}$ ($n \geq 2$).

a) Soit $\sigma_k(x_1, \dots, x_n) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} x_{i_1} x_{i_2} \dots x_{i_k}$ ($(x_1, \dots, x_n) \in \mathbb{C}^n$, $k \in \llbracket 1, n \rrbracket$).

Démontrer : $\det \left(\left[\frac{\partial \sigma_j}{\partial x_i} \right]_{(i,j) \in \llbracket 1, n \rrbracket^2} \right) = (-1)^{\frac{n(n-1)}{2}} \prod_{1 \leq i < j \leq n} (x_j - x_i).$

Indication : Utiliser les relations (1) du § X.4 et l'exercice 9a ci-dessus.

b) On pose $\sigma_k = A_k + iB_k$, avec A_k et B_k à valeurs réelles, et

$$x_k = u_k + iv_k \quad (u_k, v_k) \in \mathbb{R}^2,$$

de sorte que A_k et B_k sont, pour tout k , polynomiales en $(u_1, \dots, u_n, v_1, \dots, v_n) \in \mathbb{R}^{2n}$. Soit $(t_1, \dots, t_{2n}) \in \mathbb{R}^{2n}$ et (C_1, \dots, C_{2n}) tels que $t_k = u_k$ si $k \leq n$, $t_k = v_{k-n}$ si $k > n$, $C_k = A_k$ si $k \leq n$ et $C_k = B_{k-n}$ si $k > n$. Démontrer :

$$\det \left(\left[\frac{\partial C_l}{\partial t_k} \right]_{(k,l) \in \llbracket 1, 2n \rrbracket^2} \right) = \left| \prod_{1 \leq k < l \leq n} (x_l - x_k) \right|^2.$$

Indication : penser à l'exercice 19 ci-dessus.

Exercice 29 (le Pfaffien, d'après Kowalewski : *Déterminante theorie*, 1909) : Dans tout le problème, p désigne un entier ≥ 1 , K un corps commutatif de caractéristique différente de 2. $(X_{ij})_{1 \leq i < j < 2p}$ une famille d'indéterminées sur K . En posant $Y_{ii} = 0$

$Y_{ij} = -X_{ji}$ si $i > j$ et $Y_{ij} = X_{ij}$ si $i < j$, on obtient une matrice notée $M_p = [Y_{ij}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$, avec $n = 2p$.

On désigne par \mathcal{D} l'ensemble des *dérangements* de \mathfrak{S}_n ; \mathcal{P} l'ensemble des $\sigma \in \mathcal{D}$ tels que tout cycle de σ soit de longueur paire; $\mathcal{J} = \mathcal{D} \setminus \mathcal{P}$; \mathcal{E} l'ensemble des ensembles $\{(\alpha_1, \beta_1), \dots, (\alpha_p, \beta_p)\}$ tels que $(\forall i) \alpha_i < \beta_i$ et $\{\alpha_1, \beta_1, \dots, \alpha_p, \beta_p\} = \llbracket 1, 2p \rrbracket$; \mathcal{F} l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ telles que :

$$\sigma(1) < \sigma(3) < \dots < \sigma(2p-1) \quad \text{et} \quad \forall i \in \llbracket 1, p \rrbracket \quad \sigma(2i-1) < \sigma(2p).$$

Enfin, si $(u, v) \in \llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket$ avec $u \neq v$, on pose $(\widehat{u, v}) = \begin{cases} (u, v) & \text{si } u < v \\ (v, u) & \text{si } v < u \end{cases}$.

1° a) Calculer $\text{card}(\mathcal{E})$.

b) Etablir une bijection naturelle ψ entre \mathcal{E} et \mathcal{F} .

Dans toute la suite du problème, on notera Δ_p le déterminant $\det(M_p)$, c'est-à-dire $\Delta_p = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) Y_{\sigma(1),1} \dots Y_{\sigma(n),n}$.

2° a) Montrer que $\Delta_p = \sum_{\sigma \in \mathcal{D}} \varepsilon(\sigma) Y_{\sigma(1),1} \dots Y_{\sigma(n),n}$.

b) Pour $\sigma \in \mathcal{J}$ on note : $q(\sigma) =$ le plus petit entier de $\llbracket 1, 2p \rrbracket$ égal au minimum du support de l'un des cycles de longueur impaire de σ ; $c(\sigma)$ le cycle de σ dont le support contient $q(\sigma)$, $\tau(\sigma)$ le produit des autres cycles de σ . Soit $\varphi : \mathcal{J} \rightarrow \mathcal{J}$, $\sigma \mapsto (c(\sigma))^{-1} \tau(\sigma)$. En utilisant \mathcal{J} , prouver successivement :

$$\sum_{\sigma \in \mathcal{J}} \varepsilon(\sigma) Y_{\sigma(1),1} \dots Y_{\sigma(n),n} = 0, \quad \text{et} \quad \Delta_p = \sum_{\sigma \in \mathcal{F}} \varepsilon(\sigma) Y_{\sigma(1),1} \dots Y_{\sigma(n),n}.$$

3° On fixe $\sigma \in \mathcal{P}$. On note $G_\sigma = \varepsilon(\sigma) Y_{\sigma(1),1} \dots Y_{\sigma(n),n}$. On désigne par c_1, \dots, c_q les cycles de σ , par S_i le support de c_i , par e_i le minimum de S_i , la numérotation étant choisie telle que $e_1 < e_2 < \dots < e_q$. Enfin on note $2r_i$ la longueur de c_i ($r_i \geq 1$); $a_0 = 0$, $a_k = 2r_1 + 2r_2 + \dots + 2r_k$ pour $k \geq 1$.

On définit les éléments f_σ de \mathfrak{S}_n , g_σ de \mathfrak{S}_n , ainsi :

$$\forall k \in \llbracket 1, q \rrbracket, \quad \text{pour } 1 \leq i \leq 2r_k, \quad f_\sigma(a_{k-1} + i) = (c_k)^{i-1}(e_k),$$

$$\text{et} \quad \forall k \in \llbracket 1, q \rrbracket, \quad \text{pour } 1 \leq i \leq 2r_k, \quad g_\sigma(a_{k-1} + i) = (c_k)^i(e_k).$$

On pose aussi :

$$H_{\sigma,i} = \prod_{\lambda=0}^{r_i-1} Y_{c_i^{2\lambda+1}(e_i), c_i^{2\lambda}(e_i)}; \quad L_{\sigma,i} = \prod_{\lambda=0}^{r_i-1} Y_{c_i^{2\lambda+2}(e_i), c_i^{2\lambda+1}(e_i)} \quad (1 \leq i \leq q);$$

$$H_\sigma = \prod_{i=1}^q H_{\sigma,i}; \quad L_\sigma = \prod_{i=1}^q L_{\sigma,i}.$$

a) Démontrer que $\varepsilon(\sigma) = \varepsilon(f_\sigma) \varepsilon(g_\sigma)$, puis que $G_\sigma = \varepsilon(f_\sigma) \varepsilon(g_\sigma) H_\sigma L_\sigma$.

b) Pour chaque $s \in \mathfrak{S}_n$, soit \tilde{s} l'inverse de s . Démontrer : $\varepsilon(g_\sigma) L_\sigma = \varepsilon(f_{\tilde{\sigma}}) H_{\tilde{\sigma}}$ (on pourra étudier la permutation $(f_{\tilde{\sigma}})^{-1} \circ g_\sigma$).

4° Pour $Z \in \mathcal{E}$, $Z = \{(\alpha_1, \beta_1), \dots, (\alpha_p, \beta_p)\}$, on pose : $\Phi(Z) = \prod_{i=1}^p X_{\alpha_i, \beta_i}$; $\varepsilon(Z) =$ signature de la permutation $\psi(Z)$. Soit $\sigma \in \mathcal{P}$. On pose

$$Z_\sigma = \{(\overline{c_i^{2\lambda+1}(e_i)}), (\overline{c_i^{2\lambda}(e_i)})_{1 \leq i \leq q, 0 \leq \lambda \leq r_i-1}\}.$$

Donc $Z_\sigma \in \mathcal{E}$. $N_\sigma =$ nombre des $k \in \llbracket 1, p \rrbracket$ tels que $f_\sigma(2k) > f_\sigma(2k-1)$.

a) Démontrer : $G_\sigma = (-1)^{N_\sigma} \varepsilon(f_\sigma) \Phi(Z_\sigma) (-1)^{N_{\tilde{\sigma}}} \varepsilon(f_{\tilde{\sigma}}) \Phi(Z_{\tilde{\sigma}})$.

b) Démontrer : $(-1)^{N_\sigma} \varepsilon(f_\sigma) = \varepsilon(Z_\sigma) (-1)^p$.

5° Soit $(Z, T) \in \mathcal{E} \times \mathcal{E}$, $Z = \{(\alpha_1, \beta_1), \dots, (\alpha_p, \beta_p)\}$, $T = \{(\gamma_1, \delta_1), \dots, (\gamma_p, \delta_p)\}$. On note $A = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$, et $e_1 = \text{Min}(A)$. On définit $c_1 \in \mathfrak{S}_n$ ainsi :

$c_1(e_1) =$ autre terme que e_1 dans le couple de Z où est e_1 .

$c_1^2(e_1) =$ terme autre que $c_1(e_1)$ dans le couple de T où est $c_1(e_1)$, etc.

On s'arrête à $r_1 \geq 1$ dès que $c_1^{2r_1}(e_1) = e_1$, d'où : c_1 a été défini sur $\{e_1, \dots, c_1^{2r_1-1}(e_1), \dots\}$ prolonge c_1 à \mathfrak{S}_n par l'identité.

a) Vérifier que tout cela a un sens, et que c_1 ainsi défini est un cycle de longueur $2r_1$ (bien expliquer pourquoi le premier $k \in \mathbb{N}^*$ tel que $c_1^k(e_1) = e_1$ est *pair*).

b) Soit alors $e_2 = \text{Min}(A \setminus \text{supp}(c_1))$. On pose $c_2(e_2) =$ terme autre que e_2 dans le couple de Z où est e_2 , etc.

On continue ainsi jusqu'à épuisement des éléments de $\llbracket 1, 2p \rrbracket$, ce qui arrive quand on a construit c_1, c_2, \dots, c_q pour un certain entier q .

Démontrer que les cycles c_1, \dots, c_q sont disjoints, de longueur paire, et que si $\sigma = c_1 c_2 \dots c_q$, on a : $Z = Z_\sigma, T = Z_{\bar{\sigma}}$.

c) En déduire que l'application $\theta : \mathcal{P} \rightarrow \mathcal{E} \times \mathcal{E}, \sigma \mapsto (Z_\sigma, Z_{\bar{\sigma}})$ est bijective.

6° On pose : $Pf((X_{ij})_{1 \leq i < j \leq n}) = \sum_{Z \in \mathcal{E}} \varepsilon(Z) \Phi(Z)$ (pfaffien des X_{ij}).

a) Vérifier que : $Pf((X_{ij})) = \sum_{\sigma \in \mathcal{F}} \varepsilon(\sigma) X_{\sigma(1), \sigma(2)} X_{\sigma(3), \sigma(4)} \dots X_{\sigma(2p-1), \sigma(2p)}$.

b) Démontrer : $\Delta_p = [Pf(X_{ij})]^2$.

c) Ecrire le pfaffien des X_{ij} pour $n = 4$ (cf. exercice 2 du § XIII.4).

Exercice 30 : Soit \mathcal{R} l'ensemble des matrices $M \in \mathfrak{M}_2(\mathbb{C})$ du type $\begin{bmatrix} a & -\bar{b} \\ b & \bar{a} \end{bmatrix}$ ($(a, b) \in \mathbb{C}^2$).

a) Montrer que \mathcal{R} est une sous- \mathbb{R} -algèbre de $\mathfrak{M}_2(\mathbb{C})$, et que cette sous- \mathbb{R} -algèbre est une algèbre à division, i.e. un corps non commutatif, et vérifier que $\dim_{\mathbb{R}}(\mathcal{R}) = 4$.

b) Vérifier que le groupe quaternionique, engendré dans $\text{SL}(2, \mathbb{C})$ par $A = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$ et $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (cf. exercice 12 du § V.7) est un sous-groupe du groupe multiplicatif $\mathcal{R} \setminus \{0\}$, et que (I_2, A, B, C) est une base du \mathbb{R} -ev \mathcal{R} , si $C = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$.

(L'algèbre \mathcal{R} s'appelle **algèbre des quaternions ordinaires**).

Exercice 31 : (Autre preuve du déterminant de Vandermonde, due à Mr Escurier, Professeur de Mathématiques Spéciales.)

On reprend le déterminant de (V_n) de l'exemple 3 du § XIII.5.

a) Soit $(\lambda_1, \dots, \lambda_{n-1}) \in K^{n-1}$ et $P = X^{n-1} + \lambda_1 X^{n-2} + \dots + \lambda_{n-1} \in K[X]$. Vérifier que $\det(V_n)$ est le déterminant de la matrice obtenue en remplaçant la dernière colonne de V_n par $(P(x_1), \dots, P(x_n))$.

b) Prendre $\lambda_k = (-1)^k \sigma_k$ pour $1 \leq k \leq n-1$, où :

$$\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n-1} x_{i_1} \dots x_{i_k}.$$

Conclure à l'aide d'une récurrence.

Chapitre XIV

ÉQUATIONS LINÉAIRES SUR UN CORPS

Dans tout le chapitre, K désigne un corps commutatif fixé une fois pour toutes.

§ XIV.1 LANGAGE DE LA GÉOMÉTRIE AFFINE DANS UN ESPACE VECTORIEL

Considérons un K -ev E . Le groupe additif de E opère à gauche sur E par la loi d'addition $E \times E \rightarrow E$, $(t, x) \mapsto t + x = f_t(x)$. Pour chaque $t \in E$, la bijection $f_t : E \rightarrow E$ est appelée **translation de vecteur t** . L'application $t \mapsto f_t$, $E \rightarrow \mathfrak{S}_E$ est un homomorphisme de groupes, et il est *injectif*, car si $f_t = \text{Id}_E$, alors $f_t(0_E) = 0_E = t + 0_E = t$, d'où $t = 0_E$ (cela signifie que l'opération de E sur E que nous considérons est *fidèle* selon la terminologie du § V.6).

Le sous-groupe $\mathcal{T}(E)$ formé des $\{f_t\}_{t \in E}$ est donc isomorphe au groupe additif de E : on l'appelle **groupe des translations de E** .

Si A est une partie de E , et si $t \in E$, l'ensemble $f_t(A)$ sera appelé *translaté de A par t* , et noté $t + A$ (ou $A + t$).

DÉFINITION XIV.1.1

Un sous-ensemble non vide \mathcal{V} d'un K -ev E est appelé une **sous-variété linéaire affine de E** (en abrégé **s.v.l.a. de E**) ssi il existe un sous- K -ev V de E et $t \in E$ tels que $\mathcal{V} = t + V$.
On convient aussi que \emptyset est une sous-variété linéaire affine particulière, à laquelle on attribue la dimension -1 .

Exemple 1 : Tout sous- K -ev de E est une s.v.l.a. (prendre $t = 0_E$). Tout singleton de E est une s.v.l.a. (prendre $V = \{0_E\}$). Soit \mathcal{V} une s.v.l.a. de E , $\mathcal{V} = t + V$ (où $t \in E$ et où V est un sous- K -ev de E). Alors $V = -t + \mathcal{V}$, et $t = t + 0_E \in \mathcal{V}$. L'application $\psi_t : x \mapsto x - t$, $\mathcal{V} \rightarrow V$ est bijective. Réciproquement, si $t_1 \in \mathcal{V}$, on a $t_1 = t + x_1$ avec $x_1 \in V$ et l'application $x \mapsto x - t_1 = x - t - x_1$ définit une autre bijection φ de \mathcal{V} sur V . D'ailleurs $\mathcal{V} = t - x_1 + V$ et $-x_1 + V = V$, donc $\mathcal{V} =$

pouvait donc être considéré comme translaté de V par t_1 , et φ n'est autre que ψ_{t_1} .

Enfin, si $t_1 \in E$ et $\mathcal{V} = t_1 + V_1$ avec un sous- K -ev V_1 de E_1 , on a $t + V = t_1 + V_1$ avec $t \in \mathcal{V}$, $t_1 \in \mathcal{V}$, donc $t_1 = t + x_1$ ($x_1 \in V$), donc $V = x_1 + V_1$, d'où $V = V_1$. En résumé on a prouvé :

PROPOSITION XIV.1.1

Soit \mathcal{V} une s.v.l.a. non vide d'un K -ev E ; il existe un et un seul sous- K -ev V de E tel que \mathcal{V} soit un translaté de V . Pour chaque $t \in \mathcal{V}$, l'application $\psi_t : \mathcal{V} \rightarrow V$, $x \mapsto x - t$ est bijective, et \mathcal{V} est égal à l'ensemble des $t \in E$ tels que $\mathcal{V} = t + V$.

DÉFINITION XIV.1.2

Etant donnée une s.v.l.a. non vide \mathcal{V} d'un K -ev E , l'unique sous- K -ev V de E tel que $\mathcal{V} = t + V$ pour tout $t \in \mathcal{V}$ s'appelle l'espace directeur, ou la direction, de \mathcal{V} ; \mathcal{V} est dite de dimension (resp. codimension) finie ou infinie selon que V l'est. Si V est de dimension (resp. codimension) finie, l'entier $\dim(V)$ (resp. $\text{Codim}_E(V)$) s'appelle dimension de \mathcal{V} (resp. codimension de \mathcal{V}) et se note $\dim(\mathcal{V})$ (resp. $\text{Codim}_E(\mathcal{V})$).

Les éléments de l'espace directeur V de \mathcal{V} s'appellent **vecteurs directeurs** de \mathcal{V} . Les s.v.l.a. de dimension 1 (resp. 2) s'appellent les **droites affines** de E (resp. les **plans affines** de E). Les s.v.l.a. de codimension 1 s'appellent les **hyperplans affines** de E . Les **singletons** de E sont les s.v.l.a. d'espace directeur $\{0_E\}$.

Pour chaque couple $(x, y) \in \mathcal{V} \times \mathcal{V}$, il existe, d'après la proposition XIV.1.1 un élément $t \in V$ unique tel que $y = t + x$. Cet élément se note souvent \overrightarrow{xy} (« vecteur de x à y »). On a : $\overrightarrow{yx} = -\overrightarrow{xy}$ et $\overrightarrow{xy} = 0_E$ ssi $x = y$. Ces notations ont déjà été utilisées au chapitre VI, lorsqu'a été introduit le plan d'Argaud-Cauchy.

Supposons la s.v.l.a. de dimension finie $n \geq 1$. Pour tout $t \in \mathcal{V}$, et pour toute base $\mathcal{B} = (e_1, \dots, e_n)$ de son espace directeur V , l'application

$$(1) \quad K^n \rightarrow \mathcal{V}, \quad (\lambda_1, \dots, \lambda_n) \mapsto t + \sum_{i=1}^n \lambda_i e_i$$

est une bijection. On dit que $(t, \mathcal{B}) = (t; e_1, e_2, \dots, e_n)$ est un **repère cartésien** de \mathcal{V} , d'**origine** t , de **vecteurs de base** $(e_i)_{1 \leq i \leq n}$, et que la bijection (1) est la **représentation paramétrique** de \mathcal{V} associée à ce repère.

Parallélisme

DÉFINITION XIV.1.3

Deux s.v.l.a. non vides \mathcal{V} et \mathcal{W} d'un K -ev E , d'espaces directeurs V et W , sont dites **parallèles** ssi on a $V \subset W$ ou $W \subset V$.

Supposons \mathcal{V} et \mathcal{W} parallèles, par exemple avec $V \subset W$. Si $a \in \mathcal{V} \cap \mathcal{W}$, les applications $x \mapsto a + x$ de V dans \mathcal{V} et de W dans \mathcal{W} sont bijectives, et puisque $V \subset W$, il s'ensuit $\mathcal{V} \subset \mathcal{W}$. Par conséquent, ou bien $\mathcal{V} \cap \mathcal{W} = \emptyset$, ou bien $\mathcal{V} \subset \mathcal{W}$. On dit que \mathcal{V} et \mathcal{W} sont **strictement parallèles** ssi : elles sont parallèles et $\mathcal{V} \cap \mathcal{W} = \emptyset$.

La relation « les deux s.v.l.a. \mathcal{V} et \mathcal{W} ont même espace directeur » est d'**équivalence** entre les s.v.l.a. non vides de E . Les classes d'équivalence de cette relation sont appelées les *directions* des s.v.l.a. de E .

Elles sont en bijection canonique avec les sous- K -ev de E . Lorsque \mathcal{V} et \mathcal{W} sont de dimension finie, si $\mathcal{V} \subset \mathcal{W}$, on a $\mathcal{V} = \mathcal{W}$ ssi $\dim(\mathcal{V}) = \dim(\mathcal{W})$.

Si V est un sous- K -ev de E , par $a \in E$ il passe une et une seule s.v.l.a. d'espace directeur V , et c'est $\mathcal{V} = a + V$.

Intersection de sous-variétés linéaires affines

PROPOSITION XIV.1.2

|| Dans un K -ev E , l'intersection \mathcal{V} de toute famille $(\mathcal{V}_i)_{i \in I}$ de s.v.l.a. de E est une s.v.l.a. de E . Si $\mathcal{V} \neq \emptyset$, l'espace directeur V de \mathcal{V} est l'intersection des espaces directeurs V_i ($i \in I$) des \mathcal{V}_i .

Démonstration :

Si $\mathcal{V} = \emptyset$, il n'y a rien à prouver. Soit donc $a \in \mathcal{V}$. Alors chaque \mathcal{V}_i est non vide et admet donc un espace directeur V_i . L'application $x \mapsto a + x$ définit, pour chaque i , une bijection de V_i sur \mathcal{V}_i . Sa restriction à $V = \bigcap_{i \in I} V_i$ est donc une bijection de V sur \mathcal{V} , et par suite $\mathcal{V} = a + V$. ■

Par exemple, soit A une partie non vide de E . Il existe des s.v.l.a. (par exemple E) contenant A . L'intersection de toutes ces s.v.l.a. est donc une s.v.l.a. non vide de E , dite **engendrée** par A et que nous noterons $\text{vla}(A)$. Il résulte aussitôt des définitions que, pour tout $a \in A$, on a : $\text{vla}(A) = a + V_a$, où $V_a = \text{Vect}((x - a)_{x \in A})$. En particulier l'espace directeur V de $\text{vla}(A)$ est égal à V_a pour tout $a \in A$.

Applications affines

DÉFINITION XIV.1.4

|| Soit E et F deux K -ev, et \mathcal{V} et \mathcal{W} des s.v.l.a. non vides de E et F , d'espaces directeurs respectifs V et W . Une application $f : \mathcal{V} \rightarrow \mathcal{W}$ est dite **affine** ssi il existe $a \in \mathcal{V}$ et $\varphi \in \text{Hom}_K(V, W)$ tels que

$$(2) \quad (\forall x \in V) \quad f(a + x) = f(a) + \varphi(x)$$

|| i.e. $(\forall X \in \mathcal{V}) \quad f(X) = f(a) + \varphi(X - a)$.

Exemple 2 : $\text{Id}_{\mathcal{V}} : \mathcal{V} \rightarrow \mathcal{V}$ est une application affine (prendre $\varphi = \text{Id}_V$).
Toute application K -linéaire : $E \rightarrow F$ est affine.

Toute translation $E \rightarrow E$ est affine (on prend $\varphi = \text{Id}_E$).

Si $\varphi \in \text{Hom}_K(E, F)$, pour tout $a \in E$ et tout $b \in F$, l'application $E \rightarrow F$, $x \mapsto b + \varphi(x - a)$ est affine.

PROPOSITION XIV.1.3

En gardant les notations de la définition XIV.1.4, soit $f : \mathcal{V} \rightarrow \mathcal{W}$ une application affine. Il existe une et une seule $\varphi \in \text{Hom}_K(E, F)$ vérifiant (2) avec **au moins un** $a \in \mathcal{V}$. Avec cette application φ , on a alors (2) avec **tout** $a \in \mathcal{V}$.

Démonstration :

Soit en effet a et a_1 dans \mathcal{V} , et φ et φ_1 dans $\text{Hom}_K(V, W)$ tels que

$$(\forall x \in V) \quad f(a + x) = f(a) + \varphi(x)$$

et $(\forall y \in V) \quad f(a_1 + y) = f(a_1) + \varphi_1(y).$

On a : $a - a_1 \in V$. Posant $y = x + a - a_1$ il s'ensuit :

$$\begin{aligned} (\forall x \in V) \quad f(a) + \varphi(x) &= f(a_1) + \varphi_1(x + a - a_1) = \\ &= f(a_1) + \varphi_1(a - a_1) + \varphi_1(x). \end{aligned}$$

Prenant $x = 0_V$, on en tire d'abord $f(a) = f(a_1) + \varphi_1(a - a_1)$, puis

$$(\forall x \in V) \quad \varphi(x) = \varphi_1(x); \quad \text{d'où} \quad \varphi = \varphi_1.$$

Soit alors $b \in \mathcal{V}$ et $x \in V$. On a : $b = a + z$ ($z \in V$) ; d'où

$$f(b + x) = f(a + x + z) = f(a) + \varphi(x + z) = f(a) + \varphi(z) + \varphi(x).$$

Avec $x = 0_V$, cela donne $f(a) + \varphi(z) = f(b)$, donc

$$(\forall x \in V) \quad f(b + x) = f(b) + \varphi(x). \quad \blacksquare$$

DÉFINITION XIV.1.5

Avec les notations de la proposition XIV.1.3, si $f : \mathcal{V} \rightarrow \mathcal{W}$ est une application affine, on appelle **partie linéaire** de f l'unique application linéaire $\varphi : V \rightarrow W$ telle que

$$(\forall a \in \mathcal{V}) \quad (\forall x \in V) \quad f(a + x) = f(a) + \varphi(x)$$

que nous noterons souvent \vec{f} .

Exemple 3 : L'application $f : E \rightarrow E$ est une translation ssi elle est affine et de partie linéaire $\vec{f} = \text{Id}_E$.

Exemple 4 : Soit $\varphi \in \text{Hom}_K(E, F)$ et $a \in E$, $b \in F$. Chacune des applications $f_a : E \rightarrow F$, $x \mapsto \varphi(x - a)$ et ${}_b f : E \rightarrow F$, $x \mapsto b + \varphi(x)$ est affine, de partie linéaire φ .

Toute application affine $f : E \rightarrow F$ peut s'écrire d'une manière et d'une seule sous la forme $f = f_a$, avec $a \in E$; et d'une manière et d'une seule sous la forme ${}_b f$, avec $b \in F$.

THÉORÈME XIV.1.1

Soit E, F, G trois K -ev et des s.v.l.a. non vides $\mathcal{U}, \mathcal{V}, \mathcal{W}$ de E, F, G d'espaces directeurs respectifs U, V, W . Si $f : \mathcal{U} \rightarrow \mathcal{V}$ et $g : \mathcal{V} \rightarrow \mathcal{W}$ sont des applications affines, $g \circ f$ est affine, et on a :

$$\overrightarrow{g \circ f} = \vec{g} \circ \vec{f}.$$

Démonstration :

Soit $a \in \mathcal{U}$ et $b = f(a)$. Pour $x \in U$, on a :

$$\begin{aligned} g \circ f(a + x) &= g(b + \vec{f}(x)) = g(b) + \vec{g} \circ \vec{f}(x) = \\ &= g(f(a)) + (\vec{g} \circ \vec{f})(x). \quad \blacksquare \end{aligned}$$

En particulier, si $E = F$ et $\mathcal{U} \subset \mathcal{V}$ (d'où $U \subset V$), en prenant pour f l'injection canonique $\mathcal{U} \rightarrow \mathcal{V}$ on voit que la restriction de l'application affine g à \mathcal{U} est encore une application affine.

Bijections affines

THÉORÈME XIV.1.2

Les notations étant celles de la proposition XIV.1.3, pour qu'une application affine $f : \mathcal{V} \rightarrow \mathcal{W}$ soit bijective, il faut et il suffit que sa partie linéaire $\vec{f} : V \rightarrow W$ le soit. S'il en est ainsi, $f^{\langle -1 \rangle} : \mathcal{W} \rightarrow \mathcal{V}$ est affine, de partie linéaire $\vec{f}^{\langle -1 \rangle}$.

Démonstration :

Fixons $a \in \mathcal{V}$ et soit $b = f(a)$. Les applications $\alpha : V \rightarrow \mathcal{V}$, $x \mapsto a + x$ et $\beta : W \rightarrow \mathcal{W}$, $y \mapsto b + y$ sont bijectives, et on a : $f = \beta \circ \vec{f} \circ \alpha^{\langle -1 \rangle}$. Donc f est bijective ssi \vec{f} l'est.

Si \vec{f} est bijective, on a

$$f^{\langle -1 \rangle} = \alpha \circ \vec{f}^{\langle -1 \rangle} \circ \beta^{\langle -1 \rangle},$$

donc $f^{\langle -1 \rangle}$ est affine, de partie linéaire $\vec{f}^{\langle -1 \rangle}$. \blacksquare

En conséquence, fixons la s.v.l.a. non vide \mathcal{V} du K -ev E , d'espace directeur V . Si on se souvient que $\text{Id}_{\mathcal{V}} : \mathcal{V} \rightarrow \mathcal{V}$ est affine, en combinant les théorèmes XIV.1.1 et XIV.1.2, on obtient :

THÉORÈME XIV.1.3

|| Avec les notations ci-dessus, l'ensemble, noté $\text{GA}(\mathcal{V})$, des bijections affines de \mathcal{V} dans \mathcal{V} est un sous-groupe du groupe $\mathfrak{S}_{\mathcal{V}}$ des permutations de \mathcal{V} , et l'application $f \mapsto \vec{f}$, $\text{GA}(\mathcal{V}) \rightarrow \text{GL}_K(V)$ est un homomorphisme de groupes.

Le noyau de cet homomorphisme est formé du groupe des bijections de \mathcal{V} dans \mathcal{V} du type $x \mapsto x + a$, où $a \in V$ est fixe. Autrement dit, c'est le groupe des $g_t = f_t|_{\mathcal{V}}$ pour t parcourant V : il est isomorphe au groupe additif de V . On l'appelle **groupe des translations de \mathcal{V}** .

Puisque c'est un noyau d'homomorphisme, ce groupe est un **sous-groupe distingué** de $\text{GA}(\mathcal{V})$.

Images directes ou réciproques de s.v.l.a. par une application affine**THÉORÈME XIV.1.4**

|| Soit E, F deux K -ev, \mathcal{V} et \mathcal{W} des s.v.l.a. non vides de E et F , d'espaces directeurs respectifs V et W et $f : \mathcal{V} \rightarrow \mathcal{W}$ une application affine.

(I) Pour toute s.v.l.a. non vide \mathcal{U} de E incluse dans \mathcal{V} , d'espace directeur U (donc $U \subset V$), $f(\mathcal{U})$ est une s.v.l.a. de F , d'espace directeur $\vec{f}(U)$.

En particulier, $f(\mathcal{V})$ est une s.v.l.a. de F , d'espace directeur $\vec{f}(V)$.

(II) Pour toute s.v.l.a. non vide \mathcal{U}' de F incluse dans \mathcal{W} , d'espace directeur U' , $f^{-1}(\mathcal{U}')$ est une s.v.l.a. de E , vide ou d'espace directeur $\vec{f}^{-1}(U')$. En particulier, pour $b \in f(\mathcal{W})$, $f^{-1}(b)$ est une s.v.l.a. de E , d'espace directeur $\text{Ker}(\vec{f})$.

Démonstration :

(I) Prenons $a \in \mathcal{U}$, et soit $b = f(a)$. On a alors $f(\mathcal{U}) = b + \vec{f}(U)$ car l'application $x \mapsto a + x$ définit une bijection de U sur \mathcal{U} , et car $f(a + x) = b + \vec{f}(x)$.

(II) Soit $b \in \mathcal{U}'$ et choisissons $a \in \mathcal{V}$ tel que $f(a) = b$. On a : $\mathcal{U}' = b + U'$. Si $x \in V$, $f(a + x) = b + \vec{f}(x)$, donc $f(a + x) \in \mathcal{U}'$ ssi $f(x) \in U'$, c'est-à-dire ssi $x \in \vec{f}^{-1}(U')$; (II) s'ensuit puisque $x \mapsto a + x$ définit une bijection de V sur \mathcal{V} . ■

Equations d'une sous-variété linéaire affine

Dans ce qui suit, le K -ev E sera fixé et supposé *non nul*. Les applications affines de E dans K sont appelées les **fonctions affines** sur E . Une telle fonction f s'écrit de manière unique sous la forme $g_{C, \varphi} : x \mapsto C + \varphi(x)$, où $\varphi = \vec{f} \in E^*$, et où $C \in K$; et réciproquement toute fonction de ce type est affine sur E . On en déduit aisément que les fonctions affines :

un sous- K -ev, que nous noterons $\text{Aff}(E, K)$ du K -ev de toutes les fonctions de E dans K .

L'application $K \times E^* \rightarrow \text{Aff}(E, K)$, $(C, \varphi) \mapsto g_{C, \varphi}$ est un isomorphisme de K -ev.

En particulier, si E est de dimension finie n , on voit que $\text{Aff}(E, K)$ est de dimension finie, sa dimension étant $n + 1$: dans ce cas c'est l'ensemble des *fonctions polynomiales de degré ≤ 1 sur E* .

Une fonction affine $f = g_{C, \varphi}$ ($(C, \varphi) \in K \times E^*$) est *constante* ssi $\varphi = \vec{f} = 0$. Lorsque f n'est pas constante, elle est évidemment surjective, puisque φ l'est.

THÉORÈME XIV.1.5

- (I) Une partie \mathcal{H} de E est un **hyperplan affine** ssi il existe une **fonction affine non constante** f sur E telle que $\mathcal{H} = f^{-1}(0)$.
- (II) Si \mathcal{H} est un hyperplan affine de E , l'ensemble des fonctions affines sur E telles que $\mathcal{H} \subset f^{-1}(0)$ est une **droite vectorielle**, et pour $f \neq 0$, on a : $\mathcal{H} = f^{-1}(0)$.

Démonstration :

(I) Si $f : E \rightarrow K$ est affine non constante, elle est surjective, donc $\mathcal{H} = f^{-1}(0)$ est non vide ; donc (théorème XIV.1.4) \mathcal{H} est une s.v.l.a. de E , d'espace directeur $H = \text{Ker}(\vec{f})$. Puisque $\vec{f} \in E^* \setminus \{0\}$, H est un hyperplan vectoriel (théorème IX.5.2), donc \mathcal{H} est un hyperplan affine.

Réciproquement, si \mathcal{H} est un hyperplan affine, d'espace directeur H , on a $H = \text{Ker}(\varphi)$ avec $\varphi \in E^* \setminus \{0\}$ (théorème IX.5.2). Choisissons $a \in \mathcal{H}$. Il est clair que $\mathcal{H} = f^{-1}(0)$ avec $f : E \rightarrow K$, $x \mapsto \varphi(x) - \varphi(a)$, et f est bien affine non constante sur E .

(II) Fixons $f : E \rightarrow K$ affine non constante, telle que $\mathcal{H} = f^{-1}(0)$, et soit $\varphi = \vec{f}$. Alors $H = \text{Ker}(\varphi)$ est l'espace directeur de \mathcal{H} . Pour tout $\lambda \in K^*$, on a : $\mathcal{H} = (\lambda f)^{-1}(0)$. Si $g : E \rightarrow K$ est affine, et si $\mathcal{H} \subset g^{-1}(0)$, on a : $H \subset \text{Ker}(\vec{g})$, donc (théorème IX.5.2) $\vec{g} = \lambda \varphi$ avec $\lambda \in K$; d'où $g(x) = g(0) + \lambda \varphi(x)$ pour tout $x \in E$, soit : $g(x) = g(0) - \lambda f(0) + \lambda f(x)$. Prenons $a \in \mathcal{H}$; du fait que $f(a) = 0 = g(a)$ on déduit $g(0) = \lambda f(0)$, d'où enfin $g = \lambda f$. ■

Montrons enfin comment, *en dimension finie*, on peut définir une s.v.l.a. par un *système d'équations affines*. Pour cela, nous utiliserons la théorie de la dualité exposée au chapitre XII, § 2.

THÉORÈME XIV.1.6

- Supposons le K -ev E de dimension finie $n \geq 1$. Soit $p \in \llbracket 1, n \rrbracket$, et soit f_1, f_2, \dots, f_p des fonctions affines sur E dont les parties linéaires $\varphi_1, \varphi_2, \dots, \varphi_p$ sont linéairement indépendantes. Alors l'ensemble $\mathcal{V} = \bigcap_{i=1}^p f_i^{-1}(0)$ est une s.v.l.a. de codimension p de E , et l'ensemble des fonctions affines f qui s'annulent sur \mathcal{V} est $\text{Vect}(f_1, f_2, \dots)$.

Démonstration :

Considérons $f : E \rightarrow K^p, x \mapsto (f_1(x), \dots, f_p(x))$. La vérification que f est une fonction affine est immédiate, sa partie linéaire est $\varphi : E \rightarrow K^p, x \mapsto (\varphi_1(x), \dots, \varphi_p(x))$. L'hypothèse entraîne que φ est surjective (cf. théorème XII.2.3), donc f est surjective (théorème XIV.1.4-(I)). On en déduit que $v = f^{-1}(0) = \bigcap_{i=1}^p f_i^{-1}(0)$ est une s.v.l.a. de E d'espace directeur $V = \text{Ker}(\varphi)$ (théorème XIV.1.4-(II)), et la surjectivité de φ , avec la formule du rang, montre que $\dim(V) = n - p$, d'où la première assertion. On notera qu'en particulier $\mathcal{V} \neq \emptyset$ (si $p = n$, \mathcal{V} est un singleton).

Il est clair que toute combinaison linéaire de f_1, \dots, f_p s'annule sur \mathcal{V} . Inversement, soit $f \in \text{Aff}(E, K)$ qui s'annule sur \mathcal{V} . Prenons $a \in \mathcal{V}$. Puisque $f(a+x) = f(a) + \vec{f}(x) = \vec{f}(x)$ pour $x \in E$, on voit que \vec{f} s'annule sur V , d'où

$$\vec{f} = \lambda_1 \varphi_1 + \dots + \lambda_p \varphi_p$$

avec $(\lambda_1, \dots, \lambda_p) \in K^p$ convenable (corollaire 1 du théorème XII.2.5). De plus, $f_i(a+x) = \varphi_i(x) + f_i(a) = \varphi_i(x)$ pour $x \in E$. Donc

$$(\forall x \in E) \quad f(a+x) = \sum_{i=1}^p \lambda_i f_i(a+x),$$

ce qui signifie : $(\forall y \in E) \quad f(y) = \sum_{i=1}^p \lambda_i f_i(y)$. ■

Exercice 1 : Soit E et F deux K -ev. Si \mathcal{V} et \mathcal{W} sont des s.v.l.a. de E et F respectivement d'espaces directeurs V et W , alors $\mathcal{V} \times \mathcal{W}$ est une s.v.l.a. de $E \times F$, d'espace directeur $V \times W$.

Exercice 2 : Soit E un K -ev non nul. Démontrer que les hyperplans affines de E sont les s.v.l.a. \mathcal{V} de E distinctes de E et *maximales* pour l'inclusion, i.e. telles que les seules s.v.l.a. de E contenant \mathcal{V} sont \mathcal{V} et E .

Exercice 3 : On suppose K de caractéristique $\neq 2$. Soit \mathcal{V} une partie non vide d'un K -ev E . Pour que \mathcal{V} soit une s.v.l.a. de E , il faut et il suffit que :

$$(\forall (x, y) \in \mathcal{V} \times \mathcal{V}, \forall \lambda \in K) \quad \lambda x + (1 - \lambda)y \in \mathcal{V}.$$

Exercice 4 : Soit \mathcal{V} et \mathcal{W} deux s.v.l.a. non vides de dimension finie d'un K -ev E . On note \mathcal{U} la s.v.l.a. de E engendrée par $\mathcal{V} \cup \mathcal{W}$. Démontrer :

$$\dim \mathcal{U} = \begin{cases} \dim(\mathcal{V}) + \dim(\mathcal{W}) - \dim(\mathcal{V} \cap \mathcal{W}) & \text{si } \mathcal{V} \cap \mathcal{W} \neq \emptyset \\ \dim(\mathcal{V}) + \dim(\mathcal{W}) + 1 - \dim(V \cap W) & \text{si } \mathcal{V} \cap \mathcal{W} = \emptyset \end{cases}$$

V et W désignant les espaces directeurs de \mathcal{V} et \mathcal{W} .

Exercice 5 : Soit \mathcal{V} et \mathcal{W} deux s.v.l.a. non vides d'un K -ev E , d'espaces directeurs V et W .

a) Si $V \oplus W = E$, alors $\mathcal{V} \cap \mathcal{W}$ est un singleton.

b) Si $V + W = E$, alors $\mathcal{V} \cap \mathcal{W} \neq \emptyset$.

Exercice 6 : Le corps K est supposé fini, de cardinal q , et le K -ev E est supposé de dimension finie $n \geq 2$.

a) Pour $p \in \llbracket 1, n-1 \rrbracket$, calculer le nombre de s.v.l.a. de E de dimension p .

b) On donne une s.v.l.a. \mathcal{V} de E de dimension p ($1 \leq p \leq n-2$, ce qui sous-entend $n \geq 3$). Pour $q \in \llbracket 1, n-1-p \rrbracket$, calculer le nombre de s.v.l.a. \mathcal{W} de E de dimension q qui rencontrent \mathcal{V} .

Exercice 7 : On suppose E de dimension finie $n \geq 1$. Soit \mathcal{V} une s.v.l.a. non vide égale à $\bigcap_{i=1}^p f_i^{-1}(0)$, où les $f_i : E \rightarrow K$ sont des fonctions affines de parties linéaires φ_i indépendantes.

Pour $\lambda = (\lambda_1, \dots, \lambda_p) \in K^p$, montrer que l'ensemble $\mathcal{V}_\lambda = \{x \in E \mid (\forall i) f_i(x) = \lambda_i\}$ est une s.v.l.a. de E ayant même espace directeur V que \mathcal{V} . Montrer qu'on obtient ainsi toutes les s.v.l.a. d'espace directeur V , et étudier l'injectivité de $\lambda \mapsto \mathcal{V}_\lambda$.

Exercice 8 : On suppose E de dimension finie $n \geq 2$. Si \mathcal{V} et \mathcal{W} sont deux s.v.l.a. de E de même dimension, il existe au moins une bijection affine $f : E \rightarrow E$ telle que $f(\mathcal{V}) = \mathcal{W}$.

Exercice 9 : Soit E et F deux K -ev.

a) L'ensemble $\text{Aff}(E, F)$ des applications affines de E dans F est un sous- K -ev du K -ev $\mathcal{F}(E, F)$.

b) Si E et F sont de dimensions finies n et p , montrer que $\text{Aff}(E, F)$ est aussi de dimension finie et donner sa dimension.

Exercice 10 : Dans un K -ev E non nul, on donne des droites affines parallèles $\mathcal{D}_1, \dots, \mathcal{D}_n$ ($n \geq 2$). Soit $(\alpha_1, \dots, \alpha_n) \in K^n$ tel que $\sum_{i=1}^n \alpha_i = 1$. Démontrer que l'application

$$\mathcal{D}_1 \times \mathcal{D}_2 \times \dots \times \mathcal{D}_n \rightarrow E, (x_1, \dots, x_n) \mapsto \sum_{i=1}^n \alpha_i x_i$$

admet pour image une droite affine de E , de même direction que les \mathcal{D}_i .

§ XIV.2 ÉQUATIONS LINÉAIRES SUR UN CORPS ; CAS D'UN SYSTÈME DE CRAMER

Considérons deux K -ev non nuls E et F , et une application linéaire $f : E \rightarrow F$. Pour chaque $b \in F$ il est naturel de se demander quel est l'ensemble des $x \in E$ dont l'image par f est b , autrement dit de chercher à résoudre l'équation

$$(\mathcal{L}) \quad f(x) = b$$

à l'inconnue $x \in E$. On dit que (\mathcal{L}) est une **équation linéaire**, de **second membre** b .

Cette équation est dite **homogène** ssi $b = 0_F$. Dans ce cas elle admet au moins la solution $x = 0_E$. De manière générale, l'équation (\mathcal{L}) est dite **compatible** ssi elle admet au moins une solution, et **incompatible** dans le cas contraire.

Le langage introduit au § XIV.1 permet une discussion géométrique facile de l'équation (\mathcal{L}) . En effet $\text{Im}(f)$ est un sous- K -ev de F . La compatibilité de (\mathcal{L}) signifie que : $b \in \text{Im}(f)$, et au moins en dimension finie, nous verrons que cette condition peut s'explicitier facilement. En supposant la condition $b \in \text{Im}(f)$ remplie, l'ensemble des solutions de (\mathcal{L}) est $f^{-1}(b)$. D'après le théorème XIV.1.4, cet ensemble \mathcal{S}_b est une s.v.l.a. de E , dont l'espace directeur est $\text{Ker}(f) = \mathcal{S}_0$ c'est-à-dire l'ensemble des solutions de l'équation homogène associée (\mathcal{L}_0) . Dire que

s.v.l.a. d'espace directeur \mathcal{S}_0 est rigoureusement équivalent à l'énoncé suivant :

THÉORÈME XIV.2.1

Si (\mathcal{L}) est homogène, l'ensemble \mathcal{S}_0 de ses solutions est un sous-K-ev de E .

Si x est une solution particulière de (\mathcal{L}) supposée compatible, alors l'ensemble des solutions de (\mathcal{L}) est la s.v.l.a. de E :

$$x + \mathcal{S}_0 = \{x + x_0\}_{x_0 \in \mathcal{S}_0}.$$

Bien entendu, la vérification directe de ce théorème est tout à fait immédiate, mais cela n'enlève rien à son importance.

Remarquons enfin que si E et F sont de dimension finie, et si $\dim(E) = \dim(F)$, alors $\text{Im}(f) = F$ ssi f est bijective, cette condition équivalant à : $\text{Ker}(f) = \{0\}$. Donc dans ce cas, ou bien l'équation (\mathcal{L}) est toujours compatible, quel que soit le second membre et alors elle admet, quel que soit b , une et une seule solution ; ou bien f n'est pas bijective, et alors, selon le second membre b , ou bien (\mathcal{L}) n'a pas de solution, ou bien (\mathcal{L}) admet une solution mais il n'y a pas unicité de cette solution.

Systèmes linéaires : vocabulaire de base

Soit p et n deux entiers ≥ 1 , et $M = [a_{ij}] \in \mathfrak{M}_{p,n}(K)$ une matrice donnée, ainsi que $(b_1, \dots, b_p) \in K^p$.

On considère le système de p équations :

$$(\mathcal{S}\mathcal{L}) \quad (\forall i \in \llbracket 1, p \rrbracket)$$

$$\sum_{j=1}^n a_{ij} x_j = b_i$$

où l'inconnue est le n -uplet $(x_1, \dots, x_n) \in K^n$.

Un tel système est appelé un **système d'équations linéaires scalaires en les inconnues (x_i) , sur le corps K** (en abrégé : *système d'équations linéaires sur K* , ou tout simplement : *système linéaire sur K*).

La matrice $M = [a_{ij}]$ est appelée la **matrice du système**. (b_1, \dots, b_p) en est le **second membre**. Les $(x_1, \dots, x_n) \in K^n$ vérifiant $(\mathcal{S}\mathcal{L})$ sont appelés les **solutions** du système.

Pour i fixé ($i \in \llbracket 1, p \rrbracket$), l'équation $(\mathcal{S}\mathcal{L}_i) \sum_{j=1}^n a_{ij} x_j = b_i$ est appelée la

i -ième équation du système.

Le système $(\mathcal{S}\mathcal{L})$ est dit **homogène** ssi son second membre (b_1, \dots, b_p) est nul. Lorsque (b_1, \dots, b_p) est quelconque, le système $(\mathcal{S}\mathcal{L}_0)$ obtenu en remplaçant, dans $(\mathcal{S}\mathcal{L})$, les b_i par 0 est appelé le **système hon**

à (\mathcal{SL}) . Le **rang de la matrice M** est, par définition, le **rang du système (\mathcal{SL})** . Le système (\mathcal{SL}) est dit **compatible** ssi il admet au moins une solution, et sinon, il est dit **incompatible**. Il est clair que si (\mathcal{SL}) est homogène, il est compatible, car il admet alors la solution $(x_1, \dots, x_n) = (0, \dots, 0)$, qui est appelée la **solution nulle**, ou *solution triviale*, de ce système.

Un système linéaire est susceptible de nombreuses interprétations. En voici deux : si nous notons X la matrice-colonne de coordonnées (x_1, \dots, x_n) , et B celle de coordonnées (b_1, \dots, b_p) , le système (\mathcal{SL}) équivaut à l'unique **équation matricielle**

(1)

$$MX = B$$

à l'inconnue $X \in \mathfrak{M}_{n,1}(K)$. Cette équation matricielle (1) est appelée la **forme matricielle** du système (\mathcal{SL}) . Mais nous pouvons également associer à M l'application linéaire $f_M \in \text{Hom}_K(K^n, K^p)$ de matrice M dans les bases canoniques respectives $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (\varepsilon_1, \dots, \varepsilon_p)$ de K^n et K^p . Soit alors b le vecteur $\sum_{j=1}^p b_j \varepsilon_j$. Le système (\mathcal{SL}) équivaut à l'unique équation

(2)

$$f_M(x) = b$$

où l'inconnue x est un vecteur de K^n : l'application $(x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i e_i$

est une bijection de l'ensemble des solutions de (\mathcal{SL}) sur l'ensemble des solutions de (2). Nous dirons que l'équation 2 est une **forme vectorielle** du système (\mathcal{SL}) . En utilisant le théorème XIV.2.1, on obtient le résultat suivant, concernant le système (\mathcal{SL}) , dont bien sûr la démonstration directe serait immédiate :

PROPOSITION XIV.2.1

Supposons le système (\mathcal{SL}) compatible, et soit (x_1, \dots, x_n) une solution particulière. Alors l'ensemble des solutions de (\mathcal{SL}) est l'ensemble des n -uplets $(x_1 + z_1, \dots, x_n + z_n)$, où (z_1, \dots, z_n) parcourt l'ensemble des solutions du système homogène associé (\mathcal{SL}_0) . Ces solutions forment une s.v.l.a. de K^n dont l'espace directeur est le K -ev des solutions de (\mathcal{SL}_0) .

Systèmes linéaires associés à une équation linéaire

Reprenons l'équation linéaire (\mathcal{L}) : $f(x) = b$, où $b \in F$ et $f \in \text{Hom}_K(E, F)$. Supposons E et F non nuls, de dimensions finies respectivement égales à n et p ; munissons-les de bases $\alpha = (\alpha_1, \dots, \alpha_n)$ et $\beta = (\beta_1, \dots, \beta_p)$. Soit (b_i) les coordonnées de b dans

(x_j) les coordonnées de l'inconnue x dans la base α , et soit $M = [a_{ij}]$ la matrice $\text{Mat}_{\alpha, \beta}(f)$. Alors la correspondance $x \mapsto (x_1, \dots, x_n)$ établit une bijection entre l'ensemble des solutions de (\mathcal{L}) et celui des solutions du système linéaire :

$$(\mathcal{S}\mathcal{L})_{\alpha, \beta} \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq p).$$

Nous dirons que $(\mathcal{S}\mathcal{L})_{\alpha, \beta}$ est le **système linéaire associé à (\mathcal{L}) dans les bases α, β** . Ce système est équivalent à (\mathcal{L}) . Il est clair que $(\mathcal{S}\mathcal{L})_{\alpha, \beta}$ et (\mathcal{L}) sont homogènes ou non simultanément. Et cela nous donne une idée pour une méthode de résolution effective de (\mathcal{L}) : puisqu'on a le choix de α et de β , pourquoi ne pas essayer de les choisir de telle sorte que le système associé $(\mathcal{S}\mathcal{L})_{\alpha, \beta}$ soit d'une forme aussi simple que possible ? Bien entendu la même remarque s'applique au système $(\mathcal{S}\mathcal{L})$ auquel nous avons associé l'équation vectorielle (2). Pourquoi ne pas associer à (2), par des choix adéquats de nouvelles bases dans K^n et dans K^p , des systèmes linéaires $(\mathcal{S}\mathcal{L})_{\alpha, \beta}$ dont certains pourront se révéler plus simples que $(\mathcal{S}\mathcal{L})$ donné au départ ?

Systèmes de Cramer ⁽¹⁾

DÉFINITION XIV.2.1

Soit $n \in \mathbb{N}^*$. On appelle **système de Cramer** un système linéaire de n équations à n inconnues, du type

$$(\mathcal{S}\mathcal{L}^c) \quad (\forall i \in \llbracket 1, n \rrbracket) \quad \sum_{j=1}^n a_{ij} x_j = b_i$$

dans lequel la matrice $M = [a_{ij}]$ est **inversible**, c'est-à-dire de **rang n** .

Si nous considérons l'endomorphisme $f_M \in \text{Hom}_K(K^n)$ associé à M dans la base canonique $\mathcal{B} = (e_1, \dots, e_n)$ de K^n , nous savons que M est inversible ssi f_M est un automorphisme de K^n . Compte tenu des remarques qui suivent le théorème XIV.2.1 et du fait que M est inversible ssi $\det(M) \neq 0$, on obtient :

THÉORÈME XIV.2.2

Soit $n \in \mathbb{N}^*$. Considérons le système linéaire de n équations à n inconnues

$$(\mathcal{S}\mathcal{L}^c) \quad (\forall i \in \llbracket 1, n \rrbracket) \quad \sum_{j=1}^n a_{ij} x_j = b_i$$

⁽¹⁾ Gabriel Cramer, mathématicien suisse (1704-1752).

dans lequel la matrice $M \in \mathfrak{M}_n(K)$ est fixée. Il y a équivalence entre les six conditions suivantes :

- (I) Le système $(\mathcal{S} \mathcal{L}^c)$ est de Cramer.
- (II) $\det(M) \neq 0$.
- (III) Pour tout choix du second membre $(b_1, \dots, b_n) \in K^n$, le système $(\mathcal{S} \mathcal{L}^c)$ admet **une solution unique**.
- (IV) Le système $(\mathcal{S} \mathcal{L}^c)$ est compatible pour tout choix du second membre $(b_1, \dots, b_n) \in K^n$.
- (V) Pour tout choix du second membre $(b_1, \dots, b_n) \in K^n$, le système $(\mathcal{S} \mathcal{L}^c)$ a **au plus** une solution.
- (VI) La seule solution du système homogène de matrice M est la solution triviale.

Un système $(\mathcal{S} \mathcal{L}^c)$ étant supposé de Cramer, il reste à écrire de façon commode sa solution. Pour cela, prenons la forme matricielle du système

$$(3) \quad MX = B$$

et multiplions les deux membres de (3) à gauche par M^{-1} : on obtient sous forme matricielle l'unique solution de $(\mathcal{S} \mathcal{L}^c)$:

$$(4) \quad X = M^{-1} B .$$

Si l'on sait calculer M^{-1} , on saura par la même occasion, résoudre $(\mathcal{S} \mathcal{L}^c)$. Il y a bien la formule théorique $M^{-1} = (\det(M))^{-1} \tilde{M}$ vue au § XIII.4, mais elle est impraticable pour des applications concrètes à cause de la longueur des calculs dès que n dépasse 3.

En réalité pour résoudre un système numérique $(\mathcal{S} \mathcal{L}^c)$ on opère par substitution, ce qui réduit à chaque étape le nombre d'inconnues d'au moins une unité en ne se servant que des 4 opérations rationnelles, procédé qui a l'avantage en prenant certaines précautions de pouvoir être complètement automatisé. (Voir § XIV-4.)

Le lecteur a également appris dès la classe de Quatrième à résoudre l'équation à 1 inconnue $a_{11} x = b_1$ qui a pour solution $x_1 = \frac{1}{a_{11}} b_1$, avec

$a_{11} \in K^*$, et un système de deux équations à deux inconnues (par substitution, ou par combinaison linéaire, ou même graphiquement).

Remarque 1 : Inversement, supposons qu'on sache trouver une solution (x_1, \dots, x_n) à $(\mathcal{S} \mathcal{L}^c)$ pour un second membre (b_1, \dots, b_n) arbitraire. D'abord cela prouve que le système est de Cramer (théorème XIV.2.2) et l'expression de sa solution en fonction de (b_1, \dots, b_n) fournit justement, par identification avec (4), une expression de la matrice inverse M^{-1} . On obtient ainsi un moyen fort efficace de prouver qu'une matrice est inversible et d'explicitier son inverse.

Exemple 1 : Le corps de base est \mathbb{C} . Montrer que la matrice $M \in \mathfrak{M}_3(\mathbb{C})$ ci-après est inversible, et calculer son inverse :

$$M = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 0 & -2 \end{bmatrix}.$$

Bien sûr, dans un cas aussi simple, il n'est pas interdit de calculer $\det(M) = 1$ et d'appliquer la formule $M^{-1} = (\det(M))^{-1} \tilde{M}$, mais procédons comme il est dit dans la remarque 1 : pour un second membre arbitraire (y_1, y_2, y_3) considérons le système linéaire associé à M , aux inconnues (x_1, x_2, x_3) :

$$(5) \quad \begin{cases} x_1 + x_2 = y_1 \\ 2x_1 + 2x_2 + x_3 = y_2 \\ x_1 - 2x_3 = y_3 \end{cases}$$

Des deux premières équations on déduit $y_2 - 2y_1 = x_3$ (c'est exactement ce que l'on obtient en tirant x_1 de la première équation et en substituant l'expression obtenue dans la seconde !), puis de la troisième on tire

$$x_1 = 2x_3 + y_3 = -4y_1 + 2y_2 + y_3,$$

et enfin de la première

$$x_2 = y_1 - x_1 = 5y_1 - 2y_2 - y_3.$$

Par suite le système (5) a toujours *au plus* la solution

$$\begin{cases} x_1 = -4y_1 + 2y_2 + y_3 \\ x_2 = 5y_1 - 2y_2 - y_3 \\ x_3 = -2y_1 + y_2 \end{cases}$$

Donc d'après le théorème XIV.2.2 il est de Cramer (observation importante : il est donc *inutile* de vérifier que le triplet obtenu est vraiment solution !) et les formules trouvées pour x_1, x_2, x_3 fournissent la matrice inverse recherchée :

$$M^{-1} = \begin{bmatrix} -4 & 2 & 1 \\ 5 & -2 & -1 \\ -2 & 1 & 0 \end{bmatrix}.$$

La réponse serait quasiment la même si, au lieu de considérer M comme une matrice à coefficients dans \mathbb{C} , on avait supposé par exemple

$$M \in \mathfrak{M}_3(\mathbb{Z}/p\mathbb{Z}) \quad (p \in \mathbb{N}^*, p \text{ premier}).$$

Exemple 2 : Le corps de base est \mathbb{C} . On donne $n \in \mathbb{N}^*$ ($n \geq 2$). Montrer, sans utiliser de déterminant, que la matrice alternante

$$\Omega_n = [\omega^{(k-1)(l-1)}]_{(k,l) \in \llbracket 1, n \rrbracket^2}, \quad \text{où } \omega = e^{\frac{2i\pi}{n}},$$

est inversible, et donner son inverse.

Solution : le système linéaire associé à Ω_n est, pour un second membre arbitraire (y_1, \dots, y_n) :

$$(6) \quad \sum_{l=1}^n \omega^{(k-1)(l-1)} x_l = y_k \quad (1 \leq k \leq n).$$

Soit L_k la forme linéaire des (x_l) au premier membre de (6). Pour chaque $r \in \llbracket 1, n \rrbracket$, on déduit de (6) :

$$\sum_{k=1}^n \omega^{(1-r)(k-1)} L_k = \sum_{k=1}^n \omega^{(1-r)(k-1)} y_k,$$

soit en échangeant les sommations sur k et l :

$$(7) \quad \sum_{l=1}^n x_l \sum_{k=1}^n \omega^{(k-1)(l-r)} = \sum_{k=1}^n \omega^{-(k-1)(r-1)} y_k.$$

Mais

$$\sum_{k=1}^n \omega^{(k-1)(l-r)} = \sum_{k=1}^n [\omega^{l-r}]^{k-1} = \begin{cases} n & \text{si } l = r \\ \frac{1 - \omega^{n(l-r)}}{1 - \omega^{(l-r)}} = 0 & \text{si } l \neq r, \end{cases}$$

donc (7) se réduit à

$$x_r = \frac{1}{n} \sum_{k=1}^n \omega^{-(r-1)(k-1)} y_k \quad (1 \leq r \leq n).$$

Par le même raisonnement qu'à l'exemple 1, il s'ensuit que Ω_n est inversible, son inverse étant $\frac{1}{n} [\omega^{-(k-1)(l-1)}]_{(k,l) \in \llbracket 1, n \rrbracket^2}$ qu'on peut noter $\frac{1}{n} \overline{\Omega}_n$, la barre signifiant qu'il s'agit de la matrice formée avec les *conjugués* dans \mathbb{C} des coefficients de Ω_n .

Formules de Cramer

THÉORÈME XIV.2.3

|| Soit un système de Cramer associé à la matrice inversible
 || $M = [a_{ij}] \in \text{GL}(n, K)$ ($n \geq 2$), de second membre (b_1, \dots, b_n) .
 || Pour chaque $i \in \llbracket 1, n \rrbracket$, soit M_i la matrice dont

$$\left\| \begin{array}{l} (\Gamma_{1,i}, \dots, \Gamma_{n,i}) \text{ sont : } \Gamma_{k,i} = \mathcal{C}_k(M) \text{ si } k \neq i, \Gamma_{i,i} = (b_1, \dots, b_n). \\ \text{Alors la solution du système est :} \end{array} \right. \\ (8) \quad \boxed{x_i = \frac{1}{\det(M)} \det(M_i)} \quad (1 \leq i \leq n).$$

Les formules (8) sont appelées **formules de Cramer**.

Démonstration :

Soit (x_1, \dots, x_n) la solution. On a, en écrivant la forme matricielle du système :

$$x_1 \mathcal{C}_1(M) + \dots + x_n \mathcal{C}_n(M) = b.$$

Remplaçons $b = \Gamma_{i,i}$ par $\sum x_k \mathcal{C}_k(M)$ dans la matrice M_i ($1 \leq i \leq n$, i fixé), et développons $\det(M_i)$ par linéarité en sa i -ième colonne. On obtient :

$$(9) \quad \det(M_i) = \sum_{k=1}^n x_k \det(M_{k,i}),$$

où, pour chaque $k \in \llbracket 1, n \rrbracket$, les colonnes $(\Delta_{k,i,l})_{1 \leq l \leq n}$ de $M_{k,i}$ sont $\Delta_{k,i,l} = \mathcal{C}_l(M)$ si $l \neq i$ et $\Delta_{k,i,i} = \mathcal{C}_k(M)$. On voit que $M_{k,i}$ a deux colonnes égales si $k \neq i$, d'où alors $\det(M_{k,i}) = 0$, tandis que si $k = i$, $M_{i,i}$ n'est autre que M . Donc (9) donne

$$\det(M_i) = x_i \det(M) \quad (1 \leq i \leq n),$$

d'où (8) puisque $\det(M) \neq 0$ ■.

Avec les notations de (4), le lecteur vérifiera que les formules de Cramer ne sont qu'une façon *élégante* d'écrire le produit matriciel $M^{-1}B$, en le réduisant au calcul de $n+1$ *déterminants*, alors que l'expression de M^{-1} sous la forme $M^{-1} = \frac{1}{\det(M)} \tilde{M}$ nécessite le calcul de $n^2 + 1$ déterminants.

Malgré leur élégance, les formules de Cramer se révèlent d'un usage peu pratique dès que n dépasse 3, toujours à cause du grand nombre d'opérations nécessitées pour le calcul des déterminants. Donnons cependant un exemple où elles peuvent être utiles :

Exemple 3 : Le corps de base est quelconque. On donne trois éléments distincts a, b, c dans K tels que $\sigma_1 = a + b + c \neq 0$, et un se

$(u, v, w) \in K^3$. Montrer que le système :

$$(10) \quad \begin{cases} x_1 + ax_2 + a^3 x_3 = u \\ x_1 + bx_2 + b^3 x_3 = v \\ x_1 + cx_2 + c^3 x_3 = w \end{cases}$$

est de Cramer, et donner sa solution.

Réponse : La matrice M du système (10) est $\begin{bmatrix} 1 & a & a^3 \\ 1 & b & b^3 \\ 1 & c & c^3 \end{bmatrix}$. On calcule aisé-

ment $\det(M) = (b - c)(c - a)(a - b) \sigma_1 \neq 0$ avec les hypothèses faites, donc le système est de Cramer. On obtient immédiatement la solution (x_1, x_2, x_3) en utilisant les formules de Cramer, ce qui donne :

$$x_i = \frac{1}{\det(M)} D_i \quad (1 \leq i \leq 3), \text{ avec :}$$

$$D_1 = \det \begin{bmatrix} u & a & a^3 \\ v & b & b^3 \\ w & c & c^3 \end{bmatrix} = - [ubc(b^2 - c^2) + vca(c^2 - a^2) + wab(a^2 - b^2)]$$

$$D_2 = \det \begin{bmatrix} 1 & u & a^3 \\ 1 & v & b^3 \\ 1 & w & c^3 \end{bmatrix} = u(b^3 - c^3) + v(c^3 - a^3) + w(a^3 - b^3)$$

$$D_3 = \det \begin{bmatrix} 1 & a & u \\ 1 & b & v \\ 1 & c & w \end{bmatrix} = - [u(b - c) + v(c - a) + w(a - b)] .$$

Exercice 1 : Le corps de base est \mathbb{C} . On donne $n \in \mathbb{N}$ ($n \geq 3$). Calculer l'inverse des matrices

$$M = \begin{bmatrix} 1 & n & (n-1) & \dots & 3 & 2 \\ 2 & 1 & n & & & 3 \\ 3 & 2 & 1 & & & n \\ & & & & & \\ n & \dots & \dots & \dots & 2 & 1 \end{bmatrix} \quad \text{et} \quad N = \begin{bmatrix} -2 & 1 & 3 & 1 & -2 \\ -2 & -2 & 1 & 3 & 1 \\ 1 & -2 & -2 & 1 & 3 \\ 3 & 1 & -2 & -2 & 1 \\ -1 & 3 & 1 & -2 & -2 \end{bmatrix} .$$

Exercice 2 : Le corps de base est \mathbb{C} . Soit $n \in \mathbb{N}$ ($n \geq 2$) et $P \in \mathbb{C}[X]$ de degré $n - 1$. Démontrer (sans déterminants) que la matrice $[P(i + j)]_{(i, j) \in \llbracket 1, n \rrbracket^2}$ est inversible.

Exercice 3 : Soit $n \in \mathbb{N}$, $n \geq 2$. Le corps de base est \mathbb{C} . Étudier le système linéaire

$$x_1 + x_2 = b_1, x_2 + x_3 = b_2, \dots, x_{n-1} + x_n = b_{n-1}, x_n + x_1 = b_n,$$

et lorsqu'il est de Cramer, inverser sa matrice.

Application : Construire un polygone à n sommets connaissant les milieux des côtés. Donner une construction géométrique effective pour un pentagone plan.

Exercice 4 : Le corps de base est quelconque. Soit $n \in \mathbb{N}$ ($n \geq 2$), et $(\alpha_1, \dots, \alpha_n)$, $(\beta_1, \dots, \beta_n)$ et (b_1, \dots, b_n) dans K^n . On suppose les (α_i) tous distincts, les (β_j) tous distincts et $(\forall (i, j) \in \llbracket 1, n \rrbracket^2) \alpha_i + \beta_j \neq 0$. Résoudre le système linéaire

$$\sum_{j=1}^n \frac{1}{\alpha_i + \beta_j} x_j = b_i; \quad 1 \leq i \leq n,$$

et en déduire un calcul de la matrice inverse de $M = \left[\frac{1}{\alpha_i + \beta_j} \right]$.

Indication : Penser aux fractions rationnelles.

Application : Soit $M = \left[\frac{1}{i+j-1} \right]$. Vérifier que M^{-1} est à coefficients dans \mathbb{Z} .

Exercice 5 : Le corps K est une extension de \mathbb{C} . On donne $n \in \mathbb{N}$ ($n \geq 2$).

a) On considère la matrice circulante (cf. § XIII.5, exemple 4) :

$$M = \Gamma(a_0, \dots, a_{n-1}) = [a_{j-i}] \quad (i, j) \in \llbracket 1, n \rrbracket^2,$$

où $(a_0, \dots, a_{n-1}) \in K^n$. On suppose $\det(M) \neq 0$. Inverser M par la méthode du système linéaire.

b) Appliquer aux cas : $K = \mathbb{C}(X)$ et $a_i = X^i$ ($0 \leq i \leq n-1$) et K extension quelconque de \mathbb{C} , avec $a_k = \alpha + k\beta$, α et β donnés ($0 \leq k \leq n-1$).

Exercice 6 : On donne $n \in \mathbb{N}$ ($n \geq 2$) et $(a_1, \dots, a_n) \in K^n$, $(b_1, \dots, b_n) \in K^n$, les (a_i) étant tous distincts. Résoudre le système linéaire

$$\sum_{k=1}^n (a_i)^{k-1} x_k = b_i, \quad 1 \leq i \leq n.$$

Achever numériquement le calcul pour $n = 6$ avec $a_i = i - 1$ ($1 \leq i \leq 6$) et (b_i) quelconque.

Exercice 7 : Résoudre les systèmes linéaires suivants à 4 inconnues :

$$\begin{cases} 4x_1 + 7x_2 + 3x_3 - 2x_4 = 9 \\ 2x_1 - x_2 - 4x_3 + 3x_4 = 13 \\ 3x_1 + 2x_2 - 7x_3 - 4x_4 = 12 \\ 5x_1 - 3x_2 + x_3 + 5x_4 = 13 \end{cases} \quad \begin{cases} 3x_1 + 2x_2 + 4x_3 - x_4 = 13 \\ 5x_1 + x_2 - x_3 + 2x_4 = 9 \\ 2x_1 + 3x_2 - 7x_3 + 3x_4 = 14 \\ 4x_1 - 4x_2 + 3x_3 - 5x_4 = 4 \end{cases}$$

Exercice 8 : On donne $n \in \mathbb{N}$ ($n \geq 2$) et $(a_1, \dots, a_n) \in K^n$, $b \in K$, tels que a_1, \dots, a_n soient distincts. Résoudre le système linéaire

$$\sum_{i=1}^n a_i^k x_i = b^k \quad (0 \leq k \leq n-1).$$

Exercice 9 : On donne a, b, c, d, e dans K . Résoudre les systèmes linéaires suivants et inverser la matrice de chacun des deux systèmes :

$$a) \begin{cases} x_1 - x_2 + x_4 + x_5 = a \\ x_1 + x_2 - x_3 + x_5 = b \\ x_1 + x_2 + x_3 - x_4 = c \\ x_2 + x_3 + x_4 - x_5 = d \\ -x_1 + x_3 + x_4 + x_5 = e \end{cases} \quad b) \begin{cases} -x_2 + x_4 + x_5 = a \\ x_1 - x_3 + x_5 = b \\ x_1 + x_2 - x_4 = c \\ x_2 + x_3 - x_5 = d \\ -x_1 + x_3 + x_4 = e \end{cases}$$

Exercice 10 : On considère $\lambda_1, \dots, \lambda_n$ distincts dans K^* ($n \geq 2$). On posera

$$\Lambda = \prod_{i=1}^n \lambda_i \text{ et } \Lambda_i = \frac{\Lambda}{\lambda_i} \text{ et on suppose que } D = \Lambda + \sum_{j=1}^n \Lambda_j \neq 0.$$

a) Résoudre le système linéaire $\sum_{j=1}^n (1 + \delta_{ij} \lambda_i) x_j = \Lambda_i$ où $(\Lambda_1, \dots, \Lambda_n) \in K^n$ est donné

(δ = symbole de Kronecker). Donner l'inverse de la matrice M du système

b) Soit X une indéterminée sur K . On pose $\Phi(X) = \prod_{i=1}^n (\lambda_i - X)$ et on note $\mathcal{D}(X)$ le polynôme obtenu à partir des $\lambda_i + X$ de la même façon que D est obtenu à partir des λ_i . Enfin soit $\Delta(X) = \det [1 + \delta_{ij}(\lambda_i + X)]$ avec $(i, j) \in \llbracket 1, n \rrbracket^2$. Montrer que :

$$(\forall i) \quad \Delta(-\lambda_i) = \mathcal{D}(-\lambda_i) = -\Phi'(\lambda_i).$$

En déduire $\mathcal{D}(X) = \Delta(X)$, et en particulier : $\det(M) = D$. Retrouver ce résultat par un calcul direct.

c) En appliquant les formules de Cramer, déduire de ce qui précède la valeur, pour $i \in \llbracket 1, n \rrbracket$, du déterminant D_i de la matrice dont les colonnes Γ_j sont $\mathcal{C}_j(M)$ si $j \neq i$ et $\Gamma_i = (A_1, \dots, A_n)$.

Exercice 11 : a) Soit $M = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 0 & 0 \\ 2 & 0 & 1 \end{bmatrix}$. Calculer M^2 et montrer que $M^3 = -I$. En

déduire $\det(M)$ et M^{-1} .

b) Montrer que I , M , et M^2 sont linéairement indépendantes dans $\mathfrak{M}_3(\mathbb{C})$ et dans $\mathfrak{M}_3(\mathbb{R})$.

c) Montrer que la matrice $aI + bM + cM^2$ est régulière ssi $a^3 - b^3 + c^3 + 3abc \neq 0$. En supposant cette condition réalisée, chercher l'inverse de $aI + bM + cM^2$ sous la forme $xI + yM + zM^2$, et donner les valeurs de x, y, z en utilisant les formules de Cramer.

Exercice 12 : Calculer l'inverse de la matrice

$$M = \begin{bmatrix} 1 & 0 & -2 & -2 & 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & -2 & 1 & 1 & 1 & 1 \\ -2 & -2 & 1 & 0 & 1 & 1 & 1 & 1 \\ -2 & -2 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & -2 & -2 \\ 1 & 1 & 1 & 1 & 0 & 1 & -2 & -2 \\ 1 & 1 & 1 & 1 & -2 & -2 & 1 & 0 \\ 1 & 1 & 1 & 1 & -2 & -2 & 0 & 1 \end{bmatrix}.$$

Indication : Vu la forme de M , il semble indiqué de l'écrire $M = I_8 - 2H_1 + H_2 + H_3$ et de chercher M^{-1} sous la forme $I + xH_1 + yH_2 + zH_3$ (Réponse : $x = \frac{-2}{5}$, $y = \frac{-3}{35}$, $z = \frac{1}{7}$).

Exercice 13 : Montrer que la matrice

$$M_k = \begin{bmatrix} 1 & \frac{1}{1!} & \cdots & \frac{1}{k!} \\ \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{(k+1)!} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{k!} & \frac{1}{(k+1)!} & \cdots & \frac{1}{(2k)!} \end{bmatrix}$$

est inversible pour tout $k \in \mathbb{N}^*$.

Exercice 14 : On donne 4 nombres réels a_1, a_2, a_3, a_4 distincts deux à deux. Résoudre dans \mathbb{R}^4 le système

$$\begin{cases} |a_1 - a_2| x_2 + |a_1 - a_3| x_3 + |a_1 - a_4| x_4 = 1 \\ |a_2 - a_1| x_1 + |a_2 - a_3| x_3 + |a_2 - a_4| x_4 = 1 \\ |a_3 - a_1| x_1 + |a_3 - a_2| x_2 + |a_3 - a_4| x_4 = 1 \\ |a_4 - a_1| x_1 + |a_4 - a_2| x_2 + |a_4 - a_3| x_3 = 1 \end{cases}$$

Exercice 15 : Calculer l'inverse de la matrice suivante :

$$M = \begin{bmatrix} 0 & -1 & 1 & 3 \\ -1 & 0 & 3 & 1 \\ 1 & 3 & 0 & -1 \\ 3 & 1 & -1 & 0 \end{bmatrix}$$

en la décomposant en 4 blocs, ou par toute autre méthode.

§ XIV.3 ÉQUATIONS LINÉAIRES SUR UN CORPS : CAS GÉNÉRAL

Considérons un système linéaire

$$(\mathcal{S}\mathcal{L}) \quad \sum_{j=1}^n a_{ij} x_j = b_i, \quad 1 \leq i \leq p,$$

de matrice

$$M = [a_{ij}] \in \mathfrak{M}_{p,n}(K) \quad (n \geq 1, p \geq 1).$$

Soit r son rang. Si $r = 0$, c'est-à-dire $M = 0$, la discussion de $(\mathcal{S}\mathcal{L})$ est évidente (le système est compatible ssi $b = (b_1, \dots, b_n) = 0$, et s'il l'est, l'ensemble de ses solutions est K^n).

Nous supposons donc $r \geq 1$. Alors, tout mineur de M : $\Delta_{I,J} = \det(\mathcal{M}_{I,J}(M))$ tel que $I \subset \llbracket 1, p \rrbracket$, $J \subset \llbracket 1, n \rrbracket$, $\text{card}(I) = \text{card}(J) = r$ et $\Delta_{I,J} \neq 0$ est appelé un **déterminant principal** du système. Pour un tel mineur non nul, les équations d'indice $i \in I$ s'appellent les **équations principales** correspondant à ce mineur principal ; les $(x_j)_{j \in J}$ s'appellent les **inconnues principales** correspondant à ce mineur principal.

Système homogène

Supposons maintenant $(\mathcal{S}\mathcal{L})$ homogène et de rang r . Soit $f_M \in \text{Hom}_K(K^n, K^p)$ de matrice M dans les bases canoniques. Nous savons que l'ensemble \mathcal{S} des solutions de $(\mathcal{S}\mathcal{L})$ est $\text{Ker}(f_M)$: c'est un sous- K -ev de K^n dont la dimension est $n - \text{rg}(f_M) = n - r$. On a donc ce premier résultat.

THÉORÈME XIV.3.1

|| Si le système $(\mathcal{S}\mathcal{L})$ est **homogène et de rang r** , il est compatible et l'ensemble de ses solutions est un sous- K -ev de dimension $n - r$ de K^n .

Ce théorème est particulièrement intéressant si le nombre p des équations est $< n$. Il s'agit maintenant de calculer effectivement toutes ces solutions, dans l'hypothèse où $r \geq 1$. C'est l'objet du théorème suivant :

THÉORÈME XIV.3.2

|| Supposons le système $(\mathcal{S}\mathcal{L})$ homogène de rang $r \geq 1$.
 (I) Le système est **équivalent au système des équations principales** associées à un choix fixé arbitrairement de déterminant principal.
 (II) Supposons en outre $r < n$: on obtient les solutions de $(\mathcal{S}\mathcal{L})$ en faisant choix d'un déterminant principal, et

inconnues non principales correspondantes des valeurs arbitraires, et pour chaque choix de ces valeurs, en résolvant le système des équations principales par rapport aux inconnues principales, qui est alors un système de Cramer de rang r .

(III) Enfin si $r = n$, alors le système $(\mathcal{S}\mathcal{L})$ admet une solution unique, et c'est la solution nulle.

Démonstration :

Faisons choix d'un déterminant principal. Le système $(\mathcal{S}\mathcal{L}')$ des r équations principales est encore de rang r . Notons \mathcal{S} (resp. \mathcal{S}') le K -ev des solutions de $(\mathcal{S}\mathcal{L})$ (resp. de $(\mathcal{S}\mathcal{L}')$). Il est évident que $\mathcal{S} \subset \mathcal{S}'$. Mais, d'après le théorème XIV.3.1, $\dim(\mathcal{S}) = \dim(\mathcal{S}') = n - r$. Donc $\mathcal{S} = \mathcal{S}'$, ce qui prouve l'équivalence énoncée par (I). Notons au passage que si $r = n$ (ce qui sous-entend $p \geq n$), alors $n - r = 0$, et la solution unique est la solution nulle : $\mathcal{S} = \{(0, 0, \dots, 0)\}$. Supposons maintenant $r < n$. Parmi les déterminants d'ordre r non nuls, nous faisons choix d'un déterminant principal $\Delta_{I,J} = \det(\mathcal{M}_{I,J}(M))$. Considérons les inconnues $(x_j)_{j \in \llbracket 1, n \rrbracket \setminus J}$ non principales comme des *paramètres* (dans la pratique cela revient à faire passer les termes qui les contiennent au second membre et à leur donner des valeurs « arbitraires »). Le système $(\mathcal{S}\mathcal{L}'')$ dont les seules inconnues sont les r inconnues principales $(x_j)_{j \in J}$ et formé par les équations principales admet alors pour matrice $\mathcal{M}_{I,J}(M) \in GL_r(K)$ (de déterminant $\neq 0$), donc le système $(\mathcal{S}\mathcal{L}'')$ est de Cramer. Ce système de Cramer $(\mathcal{S}\mathcal{L}'')$ fournit donc une solution unique pour les inconnues principales $(x_j)_{j \in J}$. On pourra donc compléter chaque système de valeurs arbitraires données aux inconnues non principales $(x_j)_{j \in \llbracket 1, n \rrbracket \setminus J}$ en affectant aux inconnues principales $(x_j)_{j \in J}$ les valeurs qui résultent de la résolution du système $(\mathcal{S}\mathcal{L}'')$ et l'on obtient ainsi une solution du système $(\mathcal{S}\mathcal{L}')$ des équations principales où les inconnues sont $(x_j)_{1 \leq j \leq n}$. Il est évident que toute solution de \mathcal{S}' peut être définie de cette manière, et comme $\mathcal{S}' = \mathcal{S}$ le théorème en résulte. ■

Dans la pratique on a intérêt à renuméroter les équations du système $(\mathcal{S}\mathcal{L})$ et à renuméroter également les inconnues pour faire en sorte que le déterminant principal devienne

$$\Delta_r = \Delta_{\llbracket 1, r \rrbracket \times \llbracket 1, r \rrbracket} = \det(M_r)$$

avec $M_r = \mathcal{M}_{\llbracket 1, r \rrbracket, \llbracket 1, r \rrbracket}(M)$ dont nous rappelons qu'elle doit être inversible. Nous supposons dans ce qui suit que cette opération préliminaire a été faite, et bien entendu on choisit Δ_r comme déterminant principal. Supposons maintenant $1 \leq r \leq n - 1$. Fixons $j \in \llbracket r + 1, n \rrbracket$, et donnons aux inconnues non principales (x_j) les valeurs $(\delta_{kj})_{r+1 \leq k \leq n}$, où δ = symbole de Kronecker. Résolvons les équations principales par rapport aux inconnues principales $(x_i)_{i \in \llbracket 1, r \rrbracket}$. On obtient une solution $Z_j = (z_{1j}, \dots, z_{nj})$ de $(\mathcal{S}\mathcal{L})$ dans laquelle $z_{kj} = \delta_{kj}$ pour $r +$

où le vecteur colonne Z'_j de coordonnées (z_{1j}, \dots, z_{rj}) est, d'après la formule (4) du § XIV.2, $Z'_j = -M_r^{-1} C_j$, en notant C_j la matrice colonne de composantes (a_{1j}, \dots, a_{rj}) .

D'autre part, soit \mathcal{S} le K -ev des solutions de $(\mathcal{S}\mathcal{L})$. L'application linéaire

$$\Phi : \mathcal{S} \rightarrow K^{n-r}, (x_1, \dots, x_n) \mapsto (x_{r+1}, \dots, x_n)$$

est *surjective* (puisqu'on peut donner des valeurs arbitraires aux inconnues non principales), donc *bijective*. Les vecteurs $\varepsilon_j = (\delta_{kj})_{r+1 \leq k \leq n}$ forment, quand on donne à j les valeurs de $r+1$ à n , une base de K^{n-r} (indexé par $\llbracket r+1, n \rrbracket$), et la définition des Z_j prouve que $\Phi(Z_j) = \varepsilon_j$ pour tout j . Finalement on a prouvé que (Z_{r+1}, \dots, Z_n) est une base du K -ev \mathcal{S} des solutions de $(\mathcal{S}\mathcal{L})$.

Voici maintenant un cas particulier très important de système homogène :

THÉORÈME XIV.3.3

Soit n un entier ≥ 2 . Supposons que le système homogène

$$(\mathcal{S}\mathcal{L}) \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad 1 \leq i \leq n-1$$

à $n-1$ équations et n inconnues, de matrice $M = [a_{ij}]$ soit **de rang $n-1$** . L'ensemble \mathcal{S} de ses solutions est alors une **droite vectorielle**, engendrée par le vecteur (D_1, D_2, \dots, D_n) où, pour tout $i \in \llbracket 1, n \rrbracket$,

$$D_i = (-1)^i \det(\mathcal{M}_{\llbracket 1, n-1 \rrbracket, \llbracket 1, n \rrbracket \setminus \{i\}}(M)).$$

Démonstration :

On sait que \mathcal{S} est une droite vectorielle par application du théorème XIV.3.1. Pour tout $i \in \llbracket 1, n-1 \rrbracket$, $S_i = \sum_{j=1}^n a_{ij} D_j$ est, au signe près, le développement par rapport à la i -ième ligne du déterminant de la matrice de lignes

$$(\mathcal{L}_1(M), \mathcal{L}_2(M), \dots, \mathcal{L}_{n-1}(M), \mathcal{L}_i(M))$$

qui a deux lignes égales, donc $S_i = 0$.

Donc (D_1, \dots, D_n) est une solution de $(\mathcal{S}\mathcal{L})$; et puisque le rang de M est $(n-1)$, on a $(D_1, \dots, D_n) \neq (0, \dots, 0)$. ■

Cas général : étude d'un système compatible

Supposons le système : $(\mathcal{S}\mathcal{L}) \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq p)$ compatible et de

rang $r \geq 1$. Faisons le choix d'un déterminant principal et soit $(\mathcal{S} \mathcal{L}')$ le système des équations principales correspondantes. Nous noterons \mathcal{S} l'ensemble des solutions de $(\mathcal{S} \mathcal{L})$, \mathcal{S}' celui des solutions de $(\mathcal{S} \mathcal{L}')$ et \mathcal{S}_0 l'ensemble des solutions de $(\mathcal{S} \mathcal{L}_0)$, système homogène associé à $(\mathcal{S} \mathcal{L})$. Nous savons que \mathcal{S} est une s.v.l.a. de K^n , de dimension $n - r$, d'espace directeur \mathcal{S}_0 .

THÉOREME XIV.3.4 (théorème de Rouché-Fontené)

Adoptons les notations et hypothèses ci-dessus (le système $(\mathcal{S} \mathcal{L})$ est supposé compatible).

(I) $\mathcal{S} = \mathcal{S}'$, i.e. les solutions de $(\mathcal{S} \mathcal{L})$ sont les solutions du système des équations principales.

(II) Pour résoudre $(\mathcal{S} \mathcal{L}')$, on donne des valeurs arbitraires aux inconnues non principales, et on résout par rapport aux inconnues principales le système obtenu, qui est alors un système linéaire de Cramer.

Démonstration :

(I) Il est évident que $\mathcal{S} \subset \mathcal{S}'$. Mais $(\mathcal{S} \mathcal{L}')$ est un système linéaire à n inconnues de rang r , donc \mathcal{S}' est une s.v.l.a. de K^n de dimension $n - r$ (la même dimension que \mathcal{S}). Donc $\mathcal{S} = \mathcal{S}'$.

(II) Le raisonnement est identique à celui du théorème XIV.3.2. ■

A la suite de ce théorème XIV.3.2, nous avons montré comment obtenir une base du K -ev \mathcal{S}_0 . Ayant calculé une telle base, il suffira donc pour connaître \mathcal{S} , d'avoir déterminé une solution particulière de $(\mathcal{S} \mathcal{L})$.

Etude de la compatibilité

Il nous reste à savoir reconnaître si un système linéaire donné est ou non compatible. Reprenons donc le système linéaire

$$(\mathcal{S} \mathcal{L}) \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad 1 \leq i \leq p \quad (n \geq 1, p \geq 1),$$

de matrice $M = [a_{ij}]$, supposé de rang $r \geq 1$, avec

$$\Delta_r = \det (\mathcal{M}_{[1, r], [1, r]}(M)).$$

Ce déterminant Δ_r est supposé $\neq 0$, et nous ferons choix de Δ_r comme déterminant principal du système.

Examinons d'abord le cas simple où $r = p$. Dans ce cas où toutes les équations sont principales (ce qui sous-entend $p \leq n$), soit (z_1, \dots, z_p) la solution du système de Cramer $\sum_{j=1}^p a_{ij} x_j = b_i \quad 1 \leq i \leq p$: il est clair que

$(z_1, \dots, z_p, 0, \dots, 0)$ est une solution de $(\mathcal{S} \mathcal{L})$, donc $(\mathcal{S} \mathcal{L})$ est évidemment compatible.

Etudions maintenant le cas où $r < p$. Dans ce cas, soit M_r la matrice $\mathcal{M}_{[\![1, r]\!], [\![1, r]\!]}(M)$; formons la matrice N dont les vecteurs colonnes sont $(\mathcal{C}_1(M), \mathcal{C}_2(M), \dots, \mathcal{C}_r(M), b)$ où $b = {}^t(b_1, \dots, b_n)$.

On a donc $N \in \mathfrak{M}_{p, r+1}(K)$. Dans la matrice N , la sous-matrice M_r possède exactement $(p - r)$ matrices bordantes $\Gamma_1, \dots, \Gamma_{p-r}$, où Γ_i est la matrice carrée d'ordre $r + 1$ dont les lignes sont $\mathcal{L}_1(N), \mathcal{L}_2(N), \dots, \mathcal{L}_r(N), \mathcal{L}_{r+i}(N)$. On pose :

DÉFINITION XIV.3.1

⎧ Avec les notations ci-dessus, le système $(\mathcal{S} \mathcal{L})$ étant de rang r ,
 ⎧ $1 \leq r \leq p - 1$, la suite $(\det(\Gamma_1), \det(\Gamma_2), \dots, \det(\Gamma_{p-r}))$ est
 ⎧ appelée suite des **déterminants caractéristiques** du système $(\mathcal{S} \mathcal{L})$
 ⎧ relative au choix du déterminant principal Δ_r .

THÉORÈME XIV.3.5

|| Avec les notations et hypothèses de la définition XIV.3.1, pour que
 || le système $(\mathcal{S} \mathcal{L})$ soit **compatible** il faut et il suffit que **tous ses**
 || **déterminants caractéristiques soient nuls**.

Démonstration :

Soit $f_M: K^n \rightarrow K^p$ l'application linéaire de matrice M dans les bases canoniques; $\text{Im}(f_M)$ est le sous- K -ev de K^p engendré par les vecteurs colonnes $\mathcal{C}_1(M), \dots, \mathcal{C}_n(M)$, et aussi par les vecteurs colonnes $\mathcal{C}_1(M), \dots, \mathcal{C}_r(M)$ qui en forment une base, puisque $\Delta_r \neq 0$. Pour que $(\mathcal{S} \mathcal{L})$ soit compatible, il faut et il suffit que $b \in \text{Im}(f_M) = \text{Vect}(\mathcal{C}_1(M), \dots, \mathcal{C}_r(M))$. Cette condition équivaut à dire que la matrice N de colonnes $(\mathcal{C}_1(M), \dots, \mathcal{C}_r(M), b)$ est de rang r . Or $\Delta_r \neq 0$ et les seuls bordants de Δ_r dans la matrice N sont précisément les déterminants caractéristiques: $\det(\Gamma_1), \dots, \det(\Gamma_{p-r})$. Le théorème résulte donc du théorème XI.4.3. ■

On pourrait énoncer plus brièvement que le système $(\mathcal{S} \mathcal{L})$ est compatible ssi les matrices de colonnes

$$(\mathcal{C}_1(M), \dots, \mathcal{C}_n(M)) \text{ et } (\mathcal{C}_1(M), \dots, \mathcal{C}_n(M), b)$$

ont exactement le même rang r et l'étude qui précède donne une méthode, une fois connu le rang de M , pour expliciter les conditions exprimant que $b \in \text{Im}(f_M)$.

Conservons les mêmes hypothèses. Dans la matrice bordante Γ_i $1 \leq i \leq p - r$ soit $D_{i,1}, \dots, D_{i,r+1}$ les cofacteurs des termes b_1, \dots, b_r, b_{r+i} . On a: $D_{i,r+1} = \Delta_r$. Soit φ_i la forme linéaire sur K^n définie par $\varphi_i(y_1, \dots, y_n) = \sum_{k=1}^n D_{i,k} \cdot y_k + y_{r+i} \cdot \Delta_r$.

Le développement par rapport à sa dernière colonne du déterminant caractéristique $\det(\Gamma_i)$ montre que $\det(\Gamma_i) = \varphi_i(b_1, \dots, b_n)$. Par conséquent, si le système est compatible, on a $\varphi_i(b_1, \dots, b_n) = 0$ pour tout i . Autrement dit, les formes linéaires $(\varphi_i)_{1 \leq i \leq p-r}$ s'annulent sur $\text{Im}(f_M)$. On a bien mieux :

THÉORÈME XIV.3.6

|| Avec les notations et hypothèses ci-dessus, les formes linéaires $\varphi_1, \dots, \varphi_{p-r}$ forment une base de l'orthogonal $(\text{Im}(f(M)))^0$ de $\text{Im}(f(M))$ dans le dual de K^p .

Démonstration :

On a vu que, pour $b \in K^p$, on a $b \in \text{Im}(f)$ ssi $\varphi_i(b) = 0$ pour tout i (théorème XIV.3.5). Donc, notant W le K -ev Vect $(\varphi_1, \dots, \varphi_{p-r})$, on a : ${}^0W = \text{Im}(f_M)$. Or, nous savons que $({}^0W)^0 = W$ (théorème XII.2.5, corollaire 1) ; donc $(\text{Im}(f_M))^0 = W$.

De plus, en vertu du théorème XII.2.5, $\dim W = p - \dim(\text{Im}(f_M)) = p - r$. C'est le nombre des φ_i . Donc $(\varphi_1, \dots, \varphi_{p-r})$ est une base de W . ■

Bien entendu, on pourrait vérifier directement sans peine l'indépendance linéaire des formes (φ_i) .

Le théorème XIV.3.6 peut s'énoncer ainsi : toute relation linéaire liant les b_i qui rendent le système (\mathcal{SL}) compatible est conséquence de celles obtenues en annulant les déterminants caractéristiques.

Exemple 1 : Le corps de base est \mathbb{C} . a est un paramètre dans \mathbb{C} . Etudier le système linéaire aux inconnues (x, y, z) :

$$(\mathcal{SL}) \quad \begin{cases} 2(a-1)x + 2y - z = 2(a+1) \\ 2x + 2ay + 2z = 4a^2 + 3 \\ 4ax + 2(2a+1)y + (2a+1)z = 16a^3 - 2a^2 - a + 5. \end{cases}$$

Solution : Il s'agit d'un système linéaire de 3 équations à 3 inconnues avec second membre. Son déterminant est

$$\Delta(a) = \begin{vmatrix} 2(a-1) & 2 & -1 \\ 2 & 2a & 2 \\ 4a & 2(2a+1) & 2a+1 \end{vmatrix} = \det(M(a)).$$

On commence par calculer ce déterminant, et l'on trouve sans peine $\Delta(a) = 4a(a-1)(2a-1)$.

On est donc conduit à considérer l'ensemble $E = \left\{0, 1, \frac{1}{2}\right\}$.

Si $a \notin E$, les formules de Cramer donnent $x = \frac{A}{\Delta(a)}$, $y = \frac{B}{\Delta(a)}$, $z = \frac{C}{\Delta(a)}$ avec

$$A = \begin{vmatrix} 2(a+1) & 2 & -1 \\ 4a^2+3 & 2a & 2 \\ 16a^3-2a^2-a+5 & 2(2a+1) & (2a+1) \end{vmatrix} =$$

$$= 2a(a-1)(16a^2+34a+19)$$

$$B = \begin{vmatrix} 2(a-1) & 2(a+1) & -1 \\ 2 & 4a^2+3 & 2 \\ 4a & 16a^3-2a^2-a+5 & 2a+1 \end{vmatrix} =$$

$$= 12a(a-1)(1-2a)(1+2a)$$

$$C = \begin{vmatrix} 2(a-1) & 2 & 2(a+1) \\ 2 & 2a & 4a^2+3 \\ 4a & 2(2a+1) & 16a^3-2a^2-a+5 \end{vmatrix} =$$

$$= 4a(a-1)(16a^3-10a^2-17a-11)$$

d'où

$$x = \frac{16a^2+34a+19}{2(2a-1)}, \quad y = -3(2a+1),$$

$$z = \frac{16a^3-10a^2-17a-11}{2a-1}.$$

Il reste à étudier les cas particuliers.

Si $a = 0$, $M(a) = \begin{bmatrix} -2 & 2 & -1 \\ 2 & 0 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ est une matrice de rang 2.

Les lignes de $M(a)$ vérifient $L_3 = L_1 + L_2$ et le second membre ($b_1 = 2, b_2 = 3, b_3 = 5$) vérifie $b_3 = b_1 + b_2$, donc (théorème XIV.3.5) *le système est compatible*. Ses solutions s'obtiennent en prenant x et y pour inconnues principales, d'où :

$$x = \frac{3}{2} - \lambda, \quad y = \frac{5}{2} - \frac{\lambda}{2}, \quad z = \lambda, \quad \text{avec } \lambda \text{ arbitraire.}$$

Si $a = 1$, $M(a) = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 2 & 2 \\ 4 & 6 & 3 \end{bmatrix}$ est encore de rang 2.

Les lignes de $M(a)$ vérifient $L_3 = L_1 + 2L_2$, et le second membre ($b_1 = 4, b_2 = 7, b_3 = 18$) est tel que $b_3 = b_1 + 2b_2$. Donc, à nouveau, *le système est compatible*, et ses solutions s'obtiennent en prenant encore x et y pour inconnues principales, d'où :

$$x = \frac{3}{2} - 3\lambda, \quad y = 2 + \frac{\lambda}{2}, \quad z = \lambda, \quad \text{avec } \lambda \text{ arbitraire.}$$

Si $a = \frac{1}{2}$, $M(a) = \begin{bmatrix} -1 & 2 & -1 \\ 2 & 1 & 2 \\ 2 & 4 & 2 \end{bmatrix}$ est encore de rang 2.

Les lignes de $M(a)$ vérifient $5 L_3 = 6 L_1 + 8 L_2$, mais le second membre ($b_1 = 3, b_2 = 4, b_3 = 6$) est tel que $5 b_3 \neq 6 b_1 + 8 b_2$, donc le système est incompatible.

Remarquons que dans les deux cas particuliers où le système reste compatible, les formules du cas $a \notin E$ en donnent une solution particulière (celle qui correspond à $\lambda = 11$ si $a = 0$, celle qui correspond à $\lambda = 22$ si $a = 1$) tandis que pour $a = \frac{1}{2}$ les formules du cas $a \notin E$ n'ont plus de sens.

Exemple 2 : Déterminant de Sylvester :

Donnons-nous deux entiers m et n non nuls, et deux polynômes dans $K[X]$:

$$P = a_0 X^m + a_n X^{m-1} + \dots + a_m \in K_m[X] ,$$

$$Q = b_0 X^n + \dots + b_n \in K_n[X] , \quad \text{avec } a_0 b_0 \neq 0 ,$$

dont nous noterons D le pgcd normalisé ; cherchons les couples

$$(U, V) \in K_{n-1}[X] \times K_{m-1}[X]$$

tels que

$$(1) \quad UP + VQ = 0 .$$

Prenons pour inconnues les coefficients de $U = x_0 X^{n-1} + \dots + x_{n-1}$ et $V = y_0 X^{m-1} + \dots + y_{m-1}$, pris dans l'ordre $(x_0, x_1, \dots, x_{n-1}, y_0, y_1, \dots, y_{m-1})$. L'équation (1) équivaut au système linéaire et homogène de $m+n$ équations en ces $m+n$ inconnues, dont la matrice $\mathcal{E}(a, b) \in \mathfrak{M}_{m+n}(K)$ est :

$$\mathcal{E}(a, b) = \left[\begin{array}{cccc|cccc} a_0 & 0 & \dots & 0 & b_0 & 0 & \dots & 0 \\ & a_0 & \dots & 0 & b_1 & \dots & 0 & \\ & a_1 & \dots & 0 & \vdots & \dots & \vdots & \\ & \vdots & \dots & \vdots & b_{n-1} & \dots & \vdots & \\ & a_m & \dots & a_0 & b_n & \dots & \vdots & \\ & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \\ 0 & a_m & \dots & \vdots & \vdots & \dots & \vdots & \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \\ \vdots & 0 & \dots & \vdots & 0 & \dots & \vdots & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & b_n \end{array} \right] \left. \vphantom{\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array}} \right\} m \text{ lignes}$$

La matrice $\mathcal{E}(a, b)$ s'appelle *matrice de Sylvester* de (P, Q) ; son déterminant $R(a, b)$ s'appelle *résultant* de P et Q (ou *déterminant de Sylvester*). Par conséquent, il y a une solution différente de

l'équation (1) ssi $R(a, b) = 0$. Pour en savoir plus, considérons l'application linéaire $\Phi : K_{n-1}[X] \times K_{m-1}[X] \rightarrow K_{m+n-1}[X]$, $(U, V) \mapsto UP + VQ$. La matrice de Φ quand on prend pour bases, $(X^{m+n-1}, X^{m+n-2}, \dots)$ dans $K_{m+n-1}[X]$, et

$$((X^{n-1}, 0), (X^{n-2}, 0), (1, 0), (0, X^{m-1}), (0, X^{m-2}), \dots, (0, 1))$$

dans $K_{n-1}[X] \times K_{m-1}[X]$, est précisément $\mathcal{E}(a, b)$. Posons $P_1 = P/D$, $Q_1 = Q/D$ et $d = \deg(D)$. Le théorème de Gauss VII.2.5 montre : si $d = 0$, $\text{Ker}(\Phi) = \{0\}$; et si $d \geq 1$, $\text{Ker}(\Phi)$ est l'ensemble des couples $(FQ_1, -FP_1)$, où $F \in K_{d-1}[X]$; dans ce dernier cas, $\text{Ker}(\Phi)$ est isomorphe à $K_{d-1}[X]$ par l'application $F \mapsto (FQ_1, -FP_1)$, donc est de dimension d .

Il résulte de cette étude, d'abord que $R(a, b) \neq 0$ ssi $D = 1$ (i.e. P et Q sont premiers entre eux). Ensuite, en appliquant la formule du rang à l'application Φ , que $\text{rg}(\mathcal{E}(a, b)) + d = m + n$, c'est-à-dire :

$$d = m + n - \text{rg}(\mathcal{E}(a, b)).$$

En particulier, si K est algébriquement clos, $R(a, b) = 0$ est une condition nécessaire et suffisante pour que P et Q aient au moins une racine commune, car alors cela signifie que $d \geq 1$.

Exercice 1 : Etudier les systèmes linéaires suivants, le corps de base étant \mathbb{C} :

$$\begin{array}{ll} \text{a) } \begin{cases} ax_2 + bx_3 + cx_4 = a + b + c \\ ax_1 + \quad + cx_3 + bx_4 = a \\ bx_1 + cx_2 + \quad + ax_4 = b \\ cx_1 + bx_2 + ax_3 = c \end{cases} & \text{b) } \begin{cases} x + \lambda y + 2\lambda z = a \\ \lambda x + y + \lambda z = b \\ 2\lambda x + 2\lambda y + z = c \\ (2\lambda + 1)x + 3\lambda y + (2\lambda + 1)z = d \end{cases} \\ \quad \quad \quad (a, b, c \text{ donnés}) & \quad \quad \quad (\lambda, a, b, c, d \text{ donnés}) \\ \text{c) } \begin{cases} 2(a+1)x + 3y + az = a+4 \\ (4a-1)x + (a+1)y + (2a-1)z = 2a+2 \\ (5a-4)x + (a+1)y + (3a-4)z = a-1 \end{cases} & \text{d) } \begin{cases} x - ay + a^2z = a \\ ax - a^2y + az = 1 \\ ax + y - a^3z = 1 \end{cases} \\ \quad \quad \quad (a \text{ donné}) & \quad \quad \quad (a \text{ donné}) \end{array}$$

Exercice 2 : Montrer que le système homogène

$$\begin{cases} x = by + cz + dt \\ y = cz + dt + ax \\ z = dt + ax + by \\ t = ax + by + cz \end{cases} \quad (a, b, c, d \text{ donnés, } \neq -1)$$

admet des solutions non nulles ssi les coefficients a, b, c, d vérifient la relation

$$\frac{a}{a+1} + \frac{b}{b+1} + \frac{c}{c+1} + \frac{d}{d+1} = 1.$$

Exercice 3 : Soit (a, b, c, d) une liste des racines dans \mathbb{C} du polynôme

$$P(X) = X^4 - \frac{3}{2}X^2 - X - \lambda \quad (\lambda \in \mathbb{C}).$$

Discuter et résoudre le système linéaire

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 2 \\ ax_1 + bx_2 + cx_3 + dx_4 = 1 \\ a^2x_1 + b^2x_2 + c^2x_3 + d^2x_4 = -1 \\ a^3x_1 + b^3x_2 + c^3x_3 + d^3x_4 = -2 \end{cases}$$

Exercice 4 : Résoudre et discuter le système

$$\frac{ax + by + h}{c} = \frac{bx + cy + h}{a} = \frac{cx + ay + h}{b}$$

(a, b, c, h donnés, les trois premiers non nuls).

Exercice 5 : Le système homogène

$$\begin{cases} ax + by + cz + dt = \lambda x \\ dx + ay + bz + ct = \lambda y \\ cx + dy + az + bt = \lambda z \\ bx + cy + dz + at = \lambda t \end{cases} \quad (a, b, c, d \text{ donnés})$$

admet des solutions non nulles pour certaines valeurs de $\lambda \in \mathbb{C}$. Lesquelles ? Déterminer λ pour que, ζ étant une racine quatrième de 1, ($\zeta, \zeta^2, \zeta^3, \zeta^4$) soit solution.

Exercice 6 : Soit le système linéaire

$$a_k x_1 + b_k x_2 + c_k x_3 + d_k x_4 = e_k \quad 1 \leq k \leq 4,$$

supposé de Cramer. Montrer que, pour que sa solution $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ soit telle que

$\alpha_1 = \alpha_2 = 0$ il faut et il suffit que la matrice $\begin{bmatrix} c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \\ e_1 & e_2 & e_3 & e_4 \end{bmatrix}$ soit de rang ≤ 2 . Généraliser.

Exercice 7 : Soit $n \in \mathbb{N}^*$. On considère un système linéaire de Cramer :

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad 1 \leq i \leq n,$$

dont on note $(\alpha_1, \dots, \alpha_n)$ la solution, et Δ le déterminant. Pour $(c_1, \dots, c_n) \in K^n$, démontrer que :

$$c_1 \alpha_1 + \dots + c_n \alpha_n = -\frac{1}{\Delta} D, \quad \text{avec } D = \det \begin{bmatrix} A & \begin{matrix} b_1 \\ \vdots \\ b_n \end{matrix} \\ c_1 \dots c_n & 0 \end{bmatrix},$$

où $A = [a_{ij}] \in \mathfrak{M}_n(K)$.

Exercice 8 : On reprend l'exercice 10 du § XIV.2 en y supposant $D = 0$.

a) Si $\sum_{i=1}^n A_i \Lambda_i \neq 0$, le système est incompatible.

b) On suppose $\sum_{i=1}^n A_i \Lambda_i = 0$ et les λ_i tous $\neq 0$. Montrer que le système est alors de rang $n - 1$ et qu'il est compatible. Donner ses solutions.

Exercice 9 : Reprendre les exercices 4 et 6 du § XIV.2 lorsque ces systèmes ne sont pas de Cramer et donner une discussion complète.

Exercice 10 : (méthode de Frobenius)

On donne deux entiers n et p : $1 \leq p < n$, et une matrice $M = [a_{ij}] \in \mathfrak{M}_{p,n}(K)$ supposée de rang p . On considère le système linéaire homogène de matrice M :

$$(\mathcal{S}) \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad 1 \leq i \leq p.$$

a) Montrer qu'on peut choisir des $(a_{i,j})$ dans K ($p+1 \leq i \leq n$, $1 \leq j \leq n$) tels que la matrice $N = [a_{ij}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ soit inversible.

b) En supposant choisis de tels (a_{ij}) , pour chaque $(i, j) \in \llbracket p+1, n \rrbracket \times \llbracket 1, n \rrbracket$, soit A_{ij} le cofacteur de a_{ij} dans N . Démontrer que pour chaque $i \in \llbracket p+1, n \rrbracket$, la suite $S_i = (A_{i1}, \dots, A_{in})$ est une solution de (\mathcal{SL}) . Montrer que les $(S_i)_{p+1 \leq i \leq n}$ sont linéairement indépendantes dans K^n , et en déduire qu'elles forment une base du K -ev des solutions de (\mathcal{SL}) .

Exercice 11 : Soit L une extension du corps K supposé infini, et n un entier ≥ 2 . On considère deux matrices M et N dans $\mathfrak{M}_n(K)$, supposées distinctes et semblables dans $\mathfrak{M}_n(L)$. Il s'agit de prouver qu'elles sont alors semblables dans $\mathfrak{M}_n(K)$. Pour cela on note $[x_{ij}]$ n^2 inconnues, éléments de L , $(i, j) \in \llbracket 1, n \rrbracket^2$, et on considère le système linéaire et homogène (\mathcal{S}) de n^2 équations en ces n^2 inconnues qui s'écrit $PM = NP$, où $P = [x_{ij}] \in \mathfrak{M}_n(L)$.

a) Montrer que (\mathcal{S}) est de rang ρ tel que $1 \leq \rho < n^2$.

b) Soit \mathcal{E} le L -ev des solutions dans L^{n^2} du système (\mathcal{S}) . Montrer que \mathcal{E} admet une base formée d'éléments de K^{n^2} (cf. exercice 10). On notera (S_1, S_2, \dots, S_ν) une telle base ($\nu = n^2 - \rho$).

c) Soit $f: \mathcal{E} \rightarrow L$, $[x_{ij}] \mapsto \det [x_{ij}]$. Vérifier que f est non nulle, polynomiale, et que, si $(\lambda_1, \dots, \lambda_\nu)$ sont les coordonnées de $[x_{ij}]$ dans la base (S_1, \dots, S_ν) , $f(\sum \lambda_i S_i)$ est un polynôme en les λ_i à coefficients dans K . Appliquer enfin le corollaire 1 du théorème X.1.2 et conclure : M et N sont semblables dans $\mathfrak{M}_n(K)$.

d) On suppose que $\dim_K(L)$ est finie, égale à $e \geq 2$. Soit ξ_1, \dots, ξ_e une base du K -cv L . Vérifier que $\mathfrak{M}_n(L) = \bigoplus_{i=1}^e \xi_i \mathfrak{M}_n(K)$. On considère une matrice $P \in \mathfrak{M}_n(L)$ inversible telle

que $N = PMP^{-1}$. On l'écrit $P = \sum_{i=1}^e \xi_i P_i$, avec $P_i \in \mathfrak{M}_n(K)$ pour tout i . En considérant la

fonction $L^e \rightarrow L$, $(\lambda_1, \dots, \lambda_e) \mapsto \det(\lambda_1 P_1 + \dots + \lambda_e P_e)$, donner dans ce cas une nouvelle solution de la question posée au début.

Exercice 12 : On donne $(a, b) \in K^2$, $a \neq b$ et $n \in \mathbb{N}$, $n \geq 3$. Discuter et résoudre le système linéaire $ax_i + bx_{n+1-i} + \sum_{j \notin \{i, n+1-i\}} x_j = C_i$ ($1 \leq i \leq n$).

Exercice 13 : Soit (a_1, a_2, a_3, a_4) , (b_1, b_2, b_3, b_4) et (c_1, c_2, c_3, c_4) trois éléments de \mathbb{R}^4 tels que $(c_1, c_2, c_3, c_4) \neq (0, 0, 0, 0)$ et $\text{rg} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix} = 2$. On considère le système linéaire (\mathcal{S}) :

$$(\mathcal{S}) \begin{cases} (a_2 b_3 - a_3 b_2) x_1 + (a_3 b_1 - a_1 b_3) x_2 + (a_1 b_2 - a_2 b_1) x_3 = c_4 \\ (a_2 b_4 - a_4 b_2) x_1 + (a_4 b_1 - a_1 b_4) x_2 + (a_1 b_2 - a_2 b_1) x_4 = -c_3 \\ (a_3 b_4 - a_4 b_3) x_1 + (a_4 b_1 - a_1 b_4) x_3 + (a_1 b_3 - a_3 b_1) x_4 = c_2 \\ (a_3 b_4 - a_4 b_3) x_2 + (a_4 b_2 - a_2 b_4) x_3 + (a_2 b_3 - a_3 b_2) x_4 = -c_1. \end{cases}$$

Utiliser $M = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ x_1 & x_2 & x_3 & x_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix}$ pour trouver la CNS de compatibilité de (\mathcal{S}) .

Si cette CNS est vérifiée, donner deux équations, équivalant ensemble à (\mathcal{S}) .

Exercice 14 : Dans un plan euclidien orienté muni d'un repère orthonormé direct (O, \vec{i}, \vec{j}) , on donne trois droites D , D' et D'' non concourantes et deux à deux non parallèles d'équations respectives $ux + vy + w = 0$, $u'x + v'y + w' = 0$ et $u''x + v''y + w'' = 0$.

note A, A', A'' les sommets, définis respectivement par (D', D'') , (D'', D) et (D, D') . On pose :

$$W = \begin{vmatrix} u' & v' \\ u'' & v'' \end{vmatrix}, \quad W' = \begin{vmatrix} u'' & u \\ v'' & v \end{vmatrix}, \quad W'' = \begin{vmatrix} u & v \\ u' & v' \end{vmatrix} \quad \text{et} \quad \Delta = \begin{vmatrix} u & v & w \\ u' & v' & w' \\ u'' & v'' & w'' \end{vmatrix}.$$

a) Montrer que l'aire algébrique du triangle $(AA'A'')$ est $\frac{1}{2WW'W''} \Delta^2$ (on sait que l'aire est comptée positivement si les sommets (A, A', A'') sont parcourus dans le sens direct).

b) Généraliser à $n + 1$ hyperplans en position générale d'un espace euclidien de dimension n , et au volume du polyèdre borné qu'ils délimitent.

Exercice 15 : Le corps de base est \mathbb{C} . Résoudre le système aux inconnues x et y :

$$\begin{cases} (Ax + By)x = Lx + L'y + L'' \\ (Ax + By)y = Mx + M'y + M'' \end{cases},$$

où $A, B, L, L', L'', M, M', M''$ sont donnés.

Indication : En posant $\rho = Ax + By$ on obtient 3 équations linéaires aux 2 inconnues x et y . En écrivant la condition de compatibilité, on obtient une équation de degré 3 en ρ . On pourra généraliser avec plus de 2 inconnues.

Exercice 16 : Discuter sur \mathbb{C} , et résoudre, les systèmes d'équations ci-après, à trois inconnues x, y, z .

$$a) \begin{cases} x^2 + y^2 + z^2 = a \\ a'x + b'y + c'z = \beta \\ a''x + b''y + c''z = \gamma \end{cases}; \quad b) \begin{cases} x^2 - yz = \alpha \\ y^2 - zx = \beta \\ z^2 - xy = \gamma. \end{cases}$$

Indication : Pour b) on interprétera les premiers membres à l'aide d'une matrice carrée d'ordre 3 convenable.

Exercice 17 : Déterminer tous les n -uples (a_1, a_2, \dots, a_n) de réels tels que $a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0$ pour tout n -uplet (x_1, \dots, x_n) tel que $x_1 + \dots + x_n = 0$.

Exercice 18 : Soit le système

$$(\mathcal{S}) \quad \begin{cases} y(x^4 - y^2 + x^2) = x \\ x(x^4 - y^2 + x^2) = 1. \end{cases}$$

En faisant la combinaison $xL_1 - yL_2$ on obtient $y = x^2$. Vérifier que le système (\mathcal{S}) et le système (\mathcal{S}') $\begin{cases} y(x^4 - y^2 + x^2) = x \\ y = x^2 \end{cases}$ n'ont pas les mêmes solutions. Pourquoi ?

§ XIV.4 MÉTHODES DIRECTES DE RÉOLUTION; PIVOT PARTIEL

Dans ce paragraphe, nous supposons que le corps de base K est \mathbb{R} ou \mathbb{C} .

De nombreux problèmes concrets se ramènent à la résolution de systèmes linéaires, ou au calcul d'inverses de matrices, le nombre d'inconnues (resp. la dimension de la matrice) pouvant être très élevé, parfois de l'ordre de plusieurs centaines. Or, l'application directe des formules de Cramer, dès qu'il y a plus de 5 inconnues, défie la puissance des petites calculatrices, et sollicite inutilement les grosses : en effet, si l'on voulait calculer un déterminant d'ordre n à partir de sa formule de définition (§ XIII.4), il faudrait additionner $n!$ termes, chacun égal au produit de n fa

nécessiterait $n! - 1$ additions précédées de $(n - 1)n!$ multiplications. Les formules de Cramer demandant le calcul de $n + 1$ déterminants et de n quotients, cela ferait en tout

$$n + (n + 1) [n! - 1 + (n - 1)n!] = n(n + 1)! - 1 \text{ opérations.}$$

Pour $n = 5$ ce nombre vaut déjà 3 599 ; pour $n = 10$, il vaut près de 400 millions (399 167 999). Même un gros ordinateur, au bout de tant d'opérations, accumulera les erreurs d'arrondi au point parfois de donner un résultat non fiable.

Il a donc fallu mettre au point des méthodes itératives variées pour inverser les matrices ou résoudre des systèmes linéaires, méthodes beaucoup plus économes en nombre d'opérations, et parfois autocorrectives, ce qui améliore la fiabilité. Nous nous bornerons à exposer le *principe* des plus simples d'entre elles, n'ayant nullement l'ambition de nous substituer aux nombreux traités spécialisés (et très techniques) sur ces questions.

Compléments sur les opérations élémentaires

Nous reprendrons ci-dessous les notations du § XI.2. Commençons par compléter la proposition XI.5.1. Soit $n \in \mathbb{N}$, $n \geq 2$. Notons (E_{ij}) la base canonique du K -ev $\mathfrak{M}_n(K)$. Fixons $i \in \llbracket 1, n \rrbracket$ et considérons $T \in \mathfrak{M}_n(K)$ du type suivant :

$$(1) \quad T = I_n + \sum_{j \neq i} \lambda_j E_{ji},$$

où les $\lambda_j \in K$ sont quelconques (T est la matrice dont la i -ième colonne admet les composantes $(\lambda_1, \dots, \lambda_{i-1}, 1, \lambda_{i+1}, \dots, \lambda_n)$ et dont la j -ième colonne, pour $j \neq i$, est $(\delta_{kj})_{1 \leq k \leq n}$, δ = symbole de Kronecker).

Alors, pour toute matrice $M \in \mathfrak{M}_n(K)$, la matrice TM s'obtient en remplaçant $\mathcal{L}_j(M)$ par $\mathcal{L}_j(M) + \lambda_j \mathcal{L}_i(M)$ pour tout $j \neq i$. Il suffit pour le voir d'appliquer la règle de multiplication de deux matrices. Cela peut aussi se déduire de la proposition XI.5.1 en remarquant que T est le produit commutatif des $(I_n + \lambda_j E_{ji})_{j \neq i}$.

On en déduit aussitôt, et cela revêt une certaine importance pour la suite, que $\det(T) = 1$, et que l'inverse de T est donné par :

$$(2) \quad T^{-1} = I_n - \sum_{j \neq i} \lambda_j E_{ji}.$$

De la même manière, si l'on fixe $i \in \llbracket 1, n \rrbracket$, et si l'on considère $U \in \mathfrak{M}_n(K)$ du type

$$(3) \quad U = I_n + \sum_{j \neq i} \mu_j E_{ij},$$

où les $\mu_j \in K$ sont quelconques, on vérifie que, pour toute matrice $M \in \mathfrak{M}_n(K)$, le produit MU s'obtient, à partir de M en y rem

tout $j \neq i$, $\mathcal{C}_j(M)$ par $\mathcal{C}_j(M) + \mu_j \mathcal{C}_i(M)$ (et en gardant la colonne $\mathcal{C}_i(M)$) ; d'où $\det(U) = 1$, et U^{-1} donnée par

$$(4) \quad U^{-1} = I_n - \sum_{j \neq i} \mu_j E_{ij}.$$

Mineurs principaux

Pour toute $M \in \mathfrak{M}_n(K)$ ($n \geq 1$), nous noterons $(\mathcal{M}_r(M))_{1 \leq r \leq n}$ les matrices principales :

$$\mathcal{M}_r(M) = \mathcal{M}_{\llbracket 1, r \rrbracket, \llbracket 1, r \rrbracket}(M) \quad \text{et} \quad (\Delta_r(M))_{1 \leq r \leq n}$$

leurs déterminants que nous avons appelés les **mineurs principaux** de M ($\Delta_r(M) = \det(\mathcal{M}_r(M))$). En multipliant M par des matrices *triagonales*, on obtient des résultats très simples :

PROPOSITION XIV.4.1

Soit U, M, T dans $\mathfrak{M}_n(K)$ avec U trigonale inférieure et T trigonale supérieure. Alors, pour tout $r \in \llbracket 1, n \rrbracket$ on a :

$$(5) \quad \Delta_r(UM) = \Delta_r(U) \Delta_r(M)$$

$$(6) \quad \Delta_r(MT) = \Delta_r(M) \Delta_r(T).$$

En particulier, si U est **unipotente inférieure** (cf. § XI.2), et si T est **unipotente supérieure**, on a, pour tout $r \in \llbracket 1, n \rrbracket$:

$$(7) \quad \Delta_r(UM) = \Delta_r(M), \quad \Delta_r(MT) = \Delta_r(M).$$

Démonstration :

Il suffit d'effectuer le *produit par blocs* associés à la suite $(r, n - r)$ (pour $1 \leq r \leq n$) pour constater que :

$$\mathcal{M}_r(UM) = \mathcal{M}_r(U) \mathcal{M}_r(M) \quad \text{et} \quad \mathcal{M}_r(MT) = \mathcal{M}_r(M) \mathcal{M}_r(T). \quad \blacksquare$$

En particulier, si les mineurs principaux de M sont tous non nuls, et si U et T sont *inversibles* (pas de coefficient nul sur la diagonale), on voit que les mineurs principaux de UM et de MT sont encore tous non nuls. On a même un résultat plus précis : notons $\mathcal{C}(n, K)$ l'ensemble des $M \in \mathfrak{M}_n(K)$ dont tous les mineurs principaux sont $\neq 0$.

THÉORÈME XIV.4.1

Soit $\Gamma_+(n, K)$ (resp. $\Gamma_-(n, K)$) le groupe des matrices trigonales supérieures (resp. inférieures) inversibles. Les applications :

$$\mathcal{U}_-(n, K) \times \Gamma_+(n, K) \rightarrow \mathcal{C}(n, K), (U, T) \mapsto UT$$

$$\text{et} \quad \Gamma_-(n, K) \times \mathcal{U}_+(n, K) \rightarrow \mathcal{C}(n, K), (V, W) \mapsto VW$$

sont des bijections.

Démonstration :

Si $U \in \mathcal{U}_-(n, K)$ et $T \in \Gamma_+(n, K)$, il est clair que $UT \in \mathcal{C}(n, K)$ à cause de la proposition XIV.4.1 : $\Delta_r(UT) = \Delta_r(T) \neq 0$.

Pour prouver que la première application est injective, considérons U_1 et U_2 dans $\mathcal{U}_-(n, K)$, et T_1 et T_2 dans $\Gamma_+(n, K)$ telles que $U_1 T_1 = U_2 T_2$. Il s'ensuit

$$U_2^{-1} U_1 = T_2 T_1^{-1} \in \mathcal{U}_-(n, K) \cap \Gamma_+(n, K) = \{I_n\},$$

d'où $U_1 = U_2$ et $T_1 = T_2$. Il reste à prouver la surjectivité. Considérons donc une matrice $M = M^{(n)} = [a_{ij}^{(n)}] \in \mathcal{C}(n, K)$. Si $n = 1$, c'est terminé. Si $n \geq 2$, puisque $a_{11}^{(n)} \neq 0$, on peut poser

$$U_n = I_n - \sum_{j=2}^n \frac{a_{j,1}^{(n)}}{a_{1,1}^{(n)}} E_{j,1},$$

d'où $U_n \in \mathcal{U}_-(n, K)$ et son inverse est $U_n^{-1} = I_n + \sum_{j=2}^n \frac{a_{j,1}^{(n)}}{a_{1,1}^{(n)}} E_{j,1}$ dont le calcul est immédiat.

Effectuons le produit par blocs

$$U_n M^{(n)} = \left[\begin{array}{c|ccc} a_{1,1}^{(n)} & a_{1,2}^{(n)} & \dots & a_{1,n}^{(n)} \\ \hline 0 & & & \\ \vdots & & M^{(n-1)} & \\ 0 & & & \end{array} \right],$$

avec

$$M^{(n-1)} = [a_{1+i,1+j}^{(n-1)}] \in \mathfrak{M}_{n-1}(K).$$

Les mineurs principaux de M sont ceux de $U_n M$ (proposition XIV.4.1), d'où :

$$a_{1,1}^{(n)} \Delta_k(M^{(n-1)}) = \Delta_{k+1}(M^{(n)}) \quad (1 \leq k \leq n).$$

En particulier $M^{(n-1)} \in \mathcal{C}(n-1, K)$ et $a_{2,2}^{(n-1)} \neq 0$. On arrive ainsi, au bout de k opérations, à une relation :

$$(8) \quad U_{n-k+1} \cdots U_{n-1} U_n M^{(n)} = \left[\begin{array}{cccc|c} a_{1,1}^{(n)} & a_{1,2}^{(n)} & \dots & \dots & a_{1,n}^{(n)} \\ 0 & a_{2,2}^{(n-1)} & \dots & \dots & a_{2,n}^{(n-1)} \\ \vdots & 0 & & & \\ 0 & \vdots & & a_{k,k}^{(n-k+1)} \dots a_{k,n}^{(n-k+1)} & \\ 0 & 0 & \dots & 0 & \\ 0 & 0 & \dots & 0 & \boxed{M^{(n-k)}} \end{array} \right]$$

avec $U_n, U_{n-1}, \dots, U_{n-k+1}$ dans $\mathcal{U}_-(n, K)$ et

$$M^{(n-k)} = [a_{k+i, k+j}^{(n-k)}]_{(i,j) \in \llbracket 1, n-k \rrbracket^2} \in \mathfrak{M}_{n-k}(K) \quad (n-k \geq 1).$$

Le même raisonnement que ci-dessus prouve alors que, tant que $n-k \geq 2$, $M^{(n-k)} \in \mathcal{C}(n-k, K)$, ses mineurs principaux étant

$$\frac{\Delta_{k+1}(M)}{P_k}, \dots, \frac{\Delta_n(M)}{P_k},$$

avec $P_k = a_{1,1}^{(n)} a_{2,2}^{(n-1)} \dots a_{k,k}^{(n-k+1)} = \Delta_k(M) \neq 0$.

Si l'on pose $U_{n-k} = I_n - \sum_{j=k+2}^n \frac{a_{j,k+1}^{(n-k)}}{a_{k+1,k+1}^{(n-k)}} E_{j,k+1}$, alors

$$U_{n-k} U_{n-k+1} \dots U_n M^{(n)}$$

prend la forme

$$(9) \quad U_{n-k} U_{n-k+1} \dots U_n M^{(n)} = \left[\begin{array}{ccc|ccc} a_{1,1}^{(n)} & \dots & \dots & \dots & \dots & a_{1,n}^{(n)} \\ & \ddots & & & & \vdots \\ 0 & & \ddots & & & \\ & & & a_{k+1,k+1}^{(n-k)} & \dots & a_{k+1,n}^{(n-k)} \\ & & & & \ddots & \\ & & & & 0 & \\ & & & & \vdots & \\ 0 & \dots & \dots & 0 & & \end{array} \right] \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \hline M^{(n-k-1)} \end{array}$$

avec $M^{(n-k-1)} \in \mathcal{C}(n-k-1, K)$ et l'on arrive finalement, par récurrence descendante, à obtenir $U_n, U_{n-1}, \dots, U_3, U_2$, toutes dans $\mathcal{U}_-(n, K)$ telles que :

$$U_2 U_3 \dots U_n M^{(n)} = \left[\begin{array}{cccc|ccc} a_{1,1}^{(n)} & \dots & \dots & \dots & a_{1,n}^{(n)} & & & \\ 0 & a_{2,2}^{(n-1)} & \dots & \dots & a_{2,n}^{(n-1)} & & & \\ \vdots & 0 & & & \vdots & & & \\ \vdots & \vdots & & & \vdots & & & \\ 0 & 0 & \dots & \dots & a_{n,n}^{(1)} & & & \end{array} \right] = T, \quad \text{où } T \in \Gamma_+(n, K)$$

d'où $M = UT$, produit d'une matrice unipotente inférieure U égale à $U_n^{-1} \dots U_3^{-1} U_2^{-1}$ par une matrice trigonale supérieure T inversible, avec les mineurs principaux de T égaux à ceux de M , c'est-à-dire

$$\Delta_k(M) = \prod_{i=1}^k a_{i,i}^{(n-i+1)} \quad \text{pour } 1 \leq k \leq n.$$

Bien entendu, la bijectivité de la deuxième application mentionnée dans le théorème XIV.4.1 peut se démontrer de manière tout à fait analogue (à moins qu'on ne préfère partir de la décomposition précédente pour tM). ■

On peut remarquer que cette démonstration reste valable pour toutes les matrices de $\mathcal{C}(n, K)$ à coefficients dans un corps commutatif quelconque.

Méthode de Gauss

C'est la méthode classique de résolution d'un système de Cramer par substitution, en utilisant un *procédé algorithmique* facile à programmer de manière à être rendu entièrement automatique pour pouvoir être confié à une machine.

Considérons donc une matrice $M = [a_{ij}] \in \mathfrak{M}_n(K)$ ($n \geq 2$). Supposons $M \in \mathcal{C}(n, K)$ ⁽¹⁾, et cherchons la solution du système de Cramer de matrice M , de second membre $B = {}^t(b_1, \dots, b_n)$ donné, système qu'on peut écrire sous sa forme matricielle

$$(10) \quad MX = B$$

où $X = {}^t(x_1, \dots, x_n)$ est le vecteur colonne des inconnues. Suivons pas à pas la démarche utilisée dans la démonstration du théorème XIV.4.1. La première étape a consisté à multiplier la matrice $M = M^{(n)}$ à gauche par U_n pour lui substituer $M^{(n-1)}$. Effectuons la même opération sur la matrice colonne $B = B^{(n)}$. Il va lui être substitué la colonne $U_n B^{(n)} = B^{(n-1)}$. Toute l'astuce de la méthode consiste dans le fait que les coefficients de $M^{(n-1)}$ et de $B^{(n-1)}$ viennent se mettre en mémoire à la place de ceux de $M^{(n)}$ et de $B^{(n)}$. On continue ainsi de proche en proche jusqu'à l'étape finale qui conduit à $M^{(1)} = U_2 U_3 \dots U_n M^{(n)} = T$ et à $B^{(1)} = U_2 U_3 \dots U_n B^{(n)}$. Le système à résoudre $MX = B$ a été finalement remplacé par le système

$$(11) \quad TX = B^{(1)},$$

où l'inconnue X est toujours ${}^t(x_1, \dots, x_n)$, mais où la matrice T est une *matrice triangulaire supérieure* inversible. La résolution du système de Cramer (11) ne présente plus aucune difficulté : on procède par substitutions successives, en commençant par la dernière équation $a_{n,n}^{(1)} x_n = b_n^{(1)}$, et en remontant. Un second avantage pratique de cette méthode est qu'elle permet de résoudre « simultanément » plusieurs systèmes linéaires ayant la même matrice M mais des seconds membres distincts B, B', \dots . En effet la même succession d'opérations (prémultiplication par $U^{(n)}, U^{(n-1)}, \dots$)

⁽¹⁾ Il existe des méthodes numériques permettant de s'assurer à l'avance, dans une large gamme de cas concrets, et assez rapidement, qu'une matrice carrée est inversible. On pourra donc s'assurer préalablement de la validité de l'hypothèse ici postulée.

portera sur les divers seconds membres et l'on arrivera finalement aux systèmes $TX = B^{(1)}$, $TX = B'^{(1)}$, ..., avec la même T , et tous résolubles de la même manière. Un exemple particulièrement frappant consiste dans la recherche pratique de la matrice inverse de M : pour résoudre l'équation matricielle $MM^{-1} = I_n$, il suffit de résoudre successivement les équations $MX = \mathcal{C}_1(I_n)$, $MX = \mathcal{C}_2(I_n)$, ... et l'on obtient ainsi successivement les colonnes de la matrice inverse qui viennent se mettre en mémoire à la place de $\mathcal{C}_1(I_n)$, ..., $\mathcal{C}_n(I_n)$.

Remarque 1 : Cette méthode permet accessoirement de calculer le déterminant de la matrice M , qui est égal à $a_{11}^{(n)} a_{22}^{(n-1)} \dots a_{nn}^{(1)}$ (avec les notations du théorème XIV.4.1). On notera surtout que le nombre d'opérations simples (additions, soustractions, multiplications, divisions) utilisées ici pour le calcul de $\det(M)$ est de très loin inférieur à $nn! - 1$ (cf. exercice 1).

Une variante de la méthode de Gauss consiste à ne conserver que le résultat du théorème XIV.4.1, c'est-à-dire l'existence du couple

$$(U, T) \in \mathcal{U}_-(n, K) \times \Gamma_+(n, K)$$

tel que $M = UT$ et à déterminer ce couple (U, T) par un procédé autre que celui qui a servi dans la démonstration (par exemple en cherchant les n^2 coefficients inconnus par un système de n^2 équations à n^2 inconnues facile à résoudre pas à pas). Cette décomposition de M est appelée décomposition *LR* par les numériciens. Une fois obtenue, le système linéaire (10) se résout en deux étapes : on résout d'abord

$$(12) \quad UY = B$$

à l'inconnue $Y = {}^t(y_1, \dots, y_n)$, qui est un système de Cramer à matrice unipotente inférieure U , dont la résolution est immédiate de proche en proche, en commençant par la première équation $y_1 = b_1$ et en procédant à des substitutions successives.

Puis, Y_0 étant la solution de (12), on résout le système

$$(13) \quad TX = Y_0$$

à l'inconnue $X = {}^t(x_1, \dots, x_n)$ qui n'est autre que le système (11) déjà envisagé, dont la résolution est elle aussi immédiate en commençant par la dernière équation.

Sous les mêmes hypothèses, et avec les mêmes notations, la matrice inverse de M est telle que $M^{-1} = T^{-1} U_2 U_3 \dots U_n$. Le seul calcul non évident est celui de T^{-1} , mais il n'est guère difficile, ne serait-ce qu'en écrivant $T = DV$, où D est diagonale, et $V \in \mathcal{U}_+(n, K)$, donc ${}^tV \in \mathcal{U}_-(n, K)$ et son inverse peut s'obtenir sous forme de produit de matrices en utilisant le même algorithme que celui utilisé dans

tion du théorème XIV.4.1, ou encore par la méthode des « coefficients indéterminés ».

Méthode de Jordan

Reprenons la matrice $M \in \mathcal{C}(n, K)$. Pour éviter des confusions avec la méthode de Gauss nous noterons $M = [a_{ij}]$ sous la forme $N^{(n)} = [b_{ij}^{(n)}]$, avec $n \geq 2$. La première étape est la même que dans l'algorithme précédent : on calcule d'abord $W_n = I_n - \sum_{j=2}^n \frac{b_{j,1}^{(n)}}{b_{1,1}^{(n)}} E_{j,1}$.

Nous allons maintenant définir un nouvel algorithme qui semble conduire plus rapidement au calcul de M^{-1} . L'idée consiste à faire apparaître des zéros non seulement sous la diagonale mais également au-dessus. Pour cela, supposons qu'à l'étape n° k nous ayons trouvé des matrices $W_n, W_{n-1}, \dots, W_{n-k+1}$ du type (1) ⁽¹⁾ telles que

$$(14) \quad W_{n-k+1} \cdots W_{n-1} W_n M = \left[\begin{array}{cccc|c} b_{1,1}^{(n)} & 0 & \cdots & 0 & \\ & \ddots & & & \\ 0 & b_{2,2}^{(n-1)} & \cdots & 0 & \\ \vdots & \vdots & \ddots & & \\ 0 & \vdots & & b_{k,k}^{(n-k+1)} & \\ \vdots & 0 & \cdots & 0 & \end{array} \right] N^{(n-k)}$$

avec

$$N^{(n-k)} = [b_{i,k+j}^{(n-k)}]_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, n-k \rrbracket} \in \mathfrak{M}_{n, n-k}(K),$$

et $k \leq n-1$. Notons $S^{(n-k)}$ la matrice $[b_{k+i, k+j}^{(n-k)}] \in \mathfrak{M}_{n-k}(K)$. On vérifie facilement qu'à chaque étape, l'algorithme décrit ci-dessous *ne modifie pas les mineurs principaux de la matrice du type (14) à laquelle il s'applique*. Donc les mineurs principaux de $W_{n-k+1} \cdots W_n M$ sont respectivement $b_{11}^{(n)}, b_{11}^{(n)} b_{22}^{(n-1)}, \dots, b_{11}^{(n)} \cdots b_{kk}^{(n-k+1)}, P_k \Delta_1(S^{(n-k)}), \dots, P_k \Delta_{n-k}(S^{(n-k)})$, avec

$$P_k = b_{11}^{(n)} b_{22}^{(n-1)} \cdots b_{k,k}^{(n-k+1)} = \Delta_k(M).$$

En particulier $S^{(n-k)} \in \mathcal{C}(n-k, K)$ et $b_{k+1, k+1}^{(n-k)} = \frac{\Delta_{k+1}(M)}{\Delta_k(M)} \neq 0$. Posons

⁽¹⁾ Rappelons que les matrices du type (1) ont leur diagonale formée de 1, une colonne dont les coefficients sont quelconques (sauf celui sur la diagonale qui vaut 1), et tous les autres coefficients nuls.

alors

$$W_{n-k} = I_n - \sum_{j \neq k+1} \frac{b_{j,k+1}^{(n-k)}}{b_{k+1,k+1}^{(n-k)}} E_{j,k+1}.$$

On voit que $W_{n-k} W_{n-k+1} \dots W_{n-1} W_n M$ a exactement la même forme que le second membre de (12), mais à l'ordre $k+1$. Par récurrence, on a donc des matrices W_n, W_{n-1}, \dots, W_1 du type (1) telles que le produit $W_1 W_2 \dots W_n M$ soit une matrice diagonale $\mathcal{D} = \text{Diag}(b_{1,1}^{(n)}, \dots, b_{n,n}^{(1)})$ (avec les notations ci-dessus), telle que $b_{1,1}^{(n)} = a_{1,1} = \Delta_1(M)$, et $b_{k,k}^{(n-k+1)} = \frac{\Delta_k(M)}{\Delta_{k-1}(M)}$ pour $2 \leq k \leq n$.

Le calcul de $M^{-1} = \mathcal{D}^{-1} W_1 W_2 \dots W_n$ est alors immédiat, et l'on en déduit bien sûr théoriquement la solution du système de Cramer

$$(10) \quad MX = B$$

pour n'importe quel second membre B . En réalité, pour résoudre (10) par cette méthode, il est inutile de calculer le produit $W_1 W_2 \dots W_n$. On procède de proche en proche comme dans la méthode de Gauss, le second membre $B = B^{[n]}$ étant remplacé successivement par $W_n B^{[n]} = B^{[n-1]}$, $W_{n-1} B^{[n-1]} = B^{[n-2]}$, ... jusqu'à l'étape finale où le système à résoudre est devenu

$$(15) \quad \mathcal{D}X = B^{[1]}$$

où l'inconnue X est toujours la matrice colonne ${}^t(x_1, \dots, x_n)$, et où la matrice $\mathcal{D} = W_1 W_2 \dots W_n M$ est diagonale.

Comme la méthode de Gauss, celle de Jordan permet de résoudre simultanément plusieurs systèmes linéaires ayant la même matrice M et des seconds membres distincts. Une astuce de programmation consiste à désigner les coordonnées de B par $(b_{1,n+1}, b_{2,n+1}, \dots, b_{n,n+1})$, celles de B' par $(b_{1,n+2}, \dots, b_{n,n+2})$, ..., ce qui permet de rendre le calcul de la suite $(B^{[n]}, B^{[n-1]}, \dots, B^{[1]})$, ..., automatique.

Méthode du pivot partiel de Gauss

Le succès des méthodes précédentes repose sur le fait que les mineurs principaux de la matrice M sont tous non nuls, de sorte que les « pivots » $a_{11}^{(n)}, a_{22}^{(n-1)}, \dots, a_{nn}^{(1)}$ (resp. $b_{11}^{(n)}, \dots, b_{nn}^{(1)}$) situés sur la diagonale sont tous $\neq 0$. Elles tomberaient en défaut si, à une certaine étape, l'un des pivots était nul et perdrait toute précision si l'un des pivots, sans être nul, était trop petit. Il est donc naturel de chercher à les améliorer.

Considérons ainsi une matrice $M \in \mathfrak{M}_n(K)$ qu'on suppose seulement inversible, et le système de Cramer $MX = B$ ($n \geq 2$) que l'on :

résoudre. Nous poserons $M = M^{(n)} = [a_{i,j}^{(n)}] = G_n$. Puisque M est inversible, on a $\text{Max}_{i=1}^n |a_{i,1}| > 0$. Désignons par i_n l'indice (le premier s'il y en a plusieurs) où ce maximum est atteint, c'est-à-dire tel que :

$$|a_{i_n,1}| = \text{Max}_{i=1}^n |a_{i,1}|.$$

Si $i_n = 1$ on garde a_{11} comme pivot pour la première étape de la méthode de Gauss (resp. de Jordan). Dans ce cas posons $P_n = I_n$. Si $i_n > 1$, soit P_n la *matrice de permutation* associée à la transposition $\tau_n = \langle 1, i_n \rangle \in \mathfrak{S}_n$: multiplier $M^{(n)}$ à gauche par P_n revient alors à échanger les lignes 1 et i_n de M (c'est-à-dire à permuter les équations n° 1 et n° i_n du système). Dans tous les cas nous poserons $N^{(n)} = P_n M^{(n)} = [b_{ij}^{(n)}]$. On est donc sûr que $|b_{1,1}^{(n)}| = \text{Max}_{i=1}^n |a_{i,1}| \neq 0$ et $b_{1,1}^{(n)}$ peut être pris comme pivot pour la première

étape de la méthode de Gauss (resp. de Jordan) qui consiste, rappelons-le, à former la matrice $U_n = I_n - \sum_{j=2}^n \frac{b_{j,1}^{(n)}}{b_{1,1}^{(n)}} E_{j,1}$ et à faire le produit $U_n N^{(n)}$ qui est de la forme

$$U_n N^{(n)} = \begin{bmatrix} b_{1,1}^{(n)} & \dots & b_{1,n}^{(n)} \\ 0 & \boxed{M^{(n-1)}} & \\ \vdots & & \\ 0 & & \end{bmatrix} = G_{n-1},$$

et comme $U_n \in \mathcal{U}_-(n, K)$ on a

$$\begin{aligned} \det(U_n N^{(n)}) &= \det(N^{(n)}) = \det(P_n M) = \\ &= b_{11}^{(n)} \det(M^{(n-1)}) = \varepsilon_n \det(M), \end{aligned}$$

avec $\varepsilon_n = +1$ si $i_n = 1$ et $\varepsilon_n = -1$ si $i_n > 1$ (c'est-à-dire s'il y a eu transposition de deux lignes), d'où $\det(M^{(n-1)}) \neq 0$. Le terme $b_{1,1}^{(n)}$ est appelé le *premier pivot partiel* de l'algorithme. Le fait d'avoir choisi pour $b_{1,1}^{(n)}$ le *maximum* des $|a_{i,1}|$ assure non seulement qu'il est non nul mais optimise en quelque sorte la précision du calcul, puisque c'est par lui qu'on divise les $b_{j,1}^{(n)}$.

Il est évident qu'on peut recommencer sur $M^{(n-1)}$ l'opération qui vient d'être effectuée sur $M^{(n)}$, c'est-à-dire choisir d'abord le *second pivot partiel* en prenant l'un des termes de la première colonne de $M^{(n-1)}$ qui a une valeur absolue maximum (donc $\neq 0$). Puis on permute les lignes n° 2 et i_{n-1} de G_{n-1} à l'aide d'une matrice de permutation P_{n-1} , d

pivot vienne sur la diagonale, et enfin on multiplie à gauche la matrice G_{n-1} par une matrice $U_{n-1} \in \mathcal{U}_-(n, K)$ pour obtenir

$$U_{n-1} P_{n-1} G_{n-1} = \begin{bmatrix} b_{1,1}^{(n)} & b_{1,2}^{(n)} & \dots & b_{1,n}^{(n)} \\ 0 & b_{2,2}^{(n-1)} & \dots & b_{2,n}^{(n-1)} \\ \vdots & \vdots & \boxed{M^{(n-2)}} & \\ 0 & 0 & & \end{bmatrix} = G_{n-2}$$

avec $b_{2,2}^{(n-1)} \neq 0$. Bien entendu, on effectue au fur et à mesure sur le second membre de l'équation (10) exactement les mêmes opérations que sur la matrice M , de sorte qu'au bout de $n-1$ étapes on aboutit à

$$(16) \quad TX = U_2 P_2 \dots U_n P_n B = C$$

où T est une matrice trigonale supérieure, donc à un système immédiat à résoudre.

Si l'on utilise des matrices W à la place des matrices U on aboutit de la même façon à

$$\mathcal{D}X = W_2 P_2 \dots W_n P_n B = C',$$

où \mathcal{D} est diagonale.

Quant au déterminant de M , il vaut tout simplement $\varepsilon_2 \dots \varepsilon_n \det(T)$ (resp. $\varepsilon_2 \dots \varepsilon_n \det(\mathcal{D})$), c'est-à-dire $(-1)^t b_{1,1}^{(n)} \dots b_{n,n}^{(1)}$, t désignant le nombre total de transpositions de lignes qu'on a été amené à effectuer.

Méthode du pivot total de Gauss

A la première étape de l'algorithme du pivot partiel nous avons choisi pour premier pivot le terme *de la première colonne* de M qui avait la valeur absolue maximum. Il semble qu'on peut améliorer encore la précision du calcul en choisissant comme premier pivot, le coefficient *de M* qui a la valeur absolue maximum, c'est-à-dire a_{i_n, j_n} tel que

$$|a_{i_n, j_n}| = \max_{(i,j) \in \llbracket 1, n \rrbracket^2} |a_{i,j}|,$$

mais cela oblige à introduire deux matrices de permutation, une pour les lignes, une autre pour les colonnes (ce qui revient à permuter le nom des inconnues). De même, à chaque étape, on choisit comme pivot un élément de valeur absolue maximum parmi *tous* les coefficients de $M^{(n-k)}$, et non pas seulement dans sa première colonne. Cette méthode peut paraître séduisante à première vue, mais en réalité elle augmente le nombre de comparaisons à effectuer ainsi que le nombre de permutations, de sorte que le gain en précision espéré est compensé par une perte de

pourquoi nous n'insisterons pas davantage sur cette méthode du pivot total, le jeu n'en valant pas la chandelle. En pratique, quand la méthode du pivot partiel s'applique (c'est-à-dire pour M inversible), elle suffit et donne de bons résultats.

Valeur des méthodes ci-dessus

Cet ouvrage n'ayant en aucun cas la prétention de se substituer, même partiellement, au moindre traité d'Analyse Numérique, nous nous contenterons des quelques indications ci-après.

On peut être amené, lors des combinaisons d'équations utilisées dans les méthodes de Gauss et de Jordan, à soustraire deux nombres très voisins. Compte tenu du nombre limité de chiffres utilisé, cela conduit à une perte de précision et même à des résultats complètement faux. C'est d'ailleurs l'une des raisons qui a conduit aux méthodes de pivot maximum, plus fiables.

Cependant si par exemple $\det(M)$ est très « petit » par rapport aux coefficients de M , les *erreurs d'arrondi* (impossible de les éliminer !) risquent de faire perdre tout sens aux calculs. Prenons un exemple très simple : soit à résoudre le système

$$(S) \quad \begin{cases} x - (1,9999) y = 3 \\ 2x - 4y = 1 \end{cases}$$

dont le déterminant vaut $-2 \cdot 10^{-4}$. Il a pour solution exacte

$$x = (5 + 5 \cdot 10^{-5}) 10^4, \quad y = (2,5) \cdot 10^4.$$

Modifions « légèrement » le système et résolvons

$$(S') \quad \begin{cases} x - (1,99995) y = 3 \\ 2x - 4y = 1 \end{cases}$$

Cela donne $x = (10 + 5 \cdot 10^{-5}) 10^4, y = 5 \cdot 10^4$.

Une toute petite variation sur *un* des coefficients a suffi pour entraîner le doublement de x et y . C'est en ce sens qu'on peut dire que le système est *mal conditionné*. Dans la pratique un système mal conditionné se reconnaît au fait que les éléments de la matrice M et ceux de la matrice M^{-1} sont d'un ordre de grandeur nettement différent. Ainsi on pourra vérifier que la matrice de Hilbert

$$H = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

est mal conditionnée, en calculant exactement H^{-1} dont les coefficients sont tous entiers : certains de ces coefficients sont de valeur absco

De même, la matrice $M = \begin{bmatrix} 10 & 9 & 1 \\ 9 & 10 & 5 \\ 1 & 5 & 9 \end{bmatrix}$ qui semble anodine, avec sa symétrie, ses coefficients entiers, son déterminant pourtant égal à 1, est mal conditionnée. Le lecteur s'en convaincra s'il essaie de résoudre les systèmes linéaires $MX = B$, avec $B = \begin{bmatrix} -5 \\ 4 \\ 18 \end{bmatrix}$; puis $MX = B'$, avec $B' = \begin{bmatrix} -5 \\ 4,1 \\ 18 \end{bmatrix}$.

Exercice 1 : Les opérations de base étant l'addition, la soustraction, la multiplication et la division :

a) Montrer que pour résoudre un système de Cramer d'ordre n par la méthode de Gauss, il suffit de moins de $\frac{2}{3}n^3 + \frac{3}{2}n^2$ opérations, à comparer avec les $n(n+1)! - 1$ nécessités par les formules de Cramer.

b) Calculer de même un nombre suffisant d'opérations de base pour pouvoir trouver l'inverse de la matrice du système.

c) Même question pour la méthode de Jordan et pour la méthode du pivot partiel (on évaluera aussi le nombre maximum de comparaisons à effectuer pour trouver les pivots maximum).

Exercice 2 : Soit K un corps commutatif et $n \in \mathbb{N}$ ($n \geq 2$). On note $\mathcal{D}^*(n, K)$ le groupe multiplicatif des matrices diagonales D , avec $D = \text{Diag}(\lambda_1, \dots, \lambda_n) \in \text{GL}(n, K)$. Montrer que l'application

$$\mathcal{U}_-(n, K) \times \mathcal{D}^*(n, K) \times \mathcal{U}_+(n, K) \rightarrow \mathcal{C}(n, K),$$

$(U, D, V) \mapsto UDV$ est bijective, et en déduire que l'application $\mathcal{U}_-(n, K) \times \mathcal{U}_+(n, K) \rightarrow \mathcal{S}\mathcal{C}(n, K)$, $(U, V) \mapsto UV$ est bijective, $\mathcal{S}\mathcal{C}(n, K)$ désignant l'ensemble

$$\{M \in \mathcal{C}(n, K) \mid \forall r \in \llbracket 1, n \rrbracket, \Delta_r(M) = 1\}.$$

Exercice 3 : Soit K un corps commutatif, $n \in \mathbb{N}$ ($n \geq 2$), et $M \in \mathfrak{M}_n(K)$ de rang $r \in \llbracket 1, n-1 \rrbracket$. On suppose $\Delta_k(M) \neq 0$ pour $k \in \llbracket 1, r \rrbracket$. Montrer qu'il existe $U \in \mathcal{U}_-(n, K)$ et $V \in \mathcal{U}_+(n, K)$ telles que si

$$D = \text{Diag} \left(\Delta_1(M), \frac{\Delta_2(M)}{\Delta_1(M)}, \dots, \frac{\Delta_r(M)}{\Delta_{r-1}(M)}, 0, \dots, 0 \right),$$

alors $M = UDV$. Y a-t-il unicité ?

Exercice 4 : Résoudre le système $\begin{cases} 7x + 13y - 4z = 16 \\ 13x - 7y - 3z = 3 \\ 8x + 2y - 5z = 5 \end{cases}$ mais en s'obligeant à ne

conserver à chaque étape du calcul que trois chiffres significatifs. Comparer les résultats obtenus selon que l'on utilise la méthode de Gauss, la méthode de Jordan, la méthode du pivot partiel.

Exercice 5 : Répondre aux mêmes questions qu'à l'exercice 4 pour le système

$$\begin{cases} 12,1x + 10,2y + 14,2z = 36,5 \\ 15,1x + 20,1y + 21,4z = 56,6 \\ 13,1x + 19,8y + 19,6z = 52,5 \end{cases}.$$

Montrer que, si l'on conserve dans les calculs quatre chiffres significatifs on aboutit à un résultat bien plus proche de la solution exacte.

Exercice 6 : On considère la matrice de Vandermonde

$$\begin{bmatrix} 1 & x_0 & x_0^2 & x_0^3 & x_0^4 & x_0^5 \\ 1 & x_1 & x_1^2 & & & x_1^5 \\ \vdots & \vdots & \vdots & & & \vdots \\ 1 & x_5 & x_5^2 & & & x_5^5 \end{bmatrix}$$

où l'on a choisi $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$ et $x_5 = 5$. Calculer exactement sa matrice inverse. Ensuite calculer cet inverse par la méthode du pivot partiel de Gauss avec la précision autorisée par votre calculatrice. Constatez-vous des écarts significatifs ? Conclusion ?

Exercice 7 : Le corps de base est \mathbb{R} . Soit L une matrice ligne $[l_1, \dots, l_n]$ et C une matrice colonne $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ non nulles. A quelle condition portant sur le nombre $LC = \sum_{i=1}^n l_i c_i$ la matrice carrée $I_n + CL$ est-elle inversible ? Montrer que son inverse peut alors s'écrire sous la forme $I_n + kCL$ (k à déterminer). Soit $M \in \text{GL}(n, \mathbb{R})$ et $N = M + CL$. CNS sur le nombre $LM^{-1}C$ pour que N soit inversible. Montrer qu'alors

$$N^{-1} = M^{-1} - \frac{1}{1 + LM^{-1}C} M^{-1}CLM^{-1}.$$

Application numérique :

$$M = \begin{bmatrix} 2 & 3 & 6 & 2 \\ 1 & 2 & 3 & 1 \\ -3 & -6 & 0 & -2 \\ 0 & 4 & 8 & 1 \end{bmatrix} \quad \text{et} \quad N = \begin{bmatrix} 2,01 & 3 & 5,99 & 1,98 \\ 1 & 2 & 3 & 1 \\ -3 & -6 & 0 & -2 \\ 0 & 4 & 8 & 1 \end{bmatrix}.$$

On prendra $C = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ et on déterminera L pour que $N - M = CL$, puis on calculera l'inverse de N à partir de celui de M .

Par quel nombre faut-il remplacer $a_{24} = 1$ pour que la matrice M ne soit plus inversible ? Expliquer pourquoi, en pratique, il faut nuancer l'énoncé abrupt : « toute matrice carrée est, soit inversible, soit non inversible ».

Chapitre XV

RÉDUCTION D'ENDOMORPHISMES OU DE MATRICES CARRÉES

Dans tout le chapitre, K désignera un corps commutatif. Si E est un K -ev de dimension finie $n \geq 1$, l'expression « *base ordonnée de E* » désignera une base de E indexée par $\llbracket 1, n \rrbracket$.

§ XV.1 VALEURS PROPRES ET POLYNÔME CARACTÉRISTIQUE

DÉFINITION XV.1.1

Soit E un K -ev non nul et $u \in \text{Hom}_K(E)$. On appelle **valeur spectrale** de u tout $\lambda \in K$ tel que $u - \lambda \text{Id}_E$ soit non bijectif, et **valeur propre** de u tout $\lambda \in K$ tel que $\text{Ker}(u - \lambda \text{Id}_E) \neq \{0\}$. Si $\lambda \in K$ est valeur propre de u , le sous- K -ev $\text{Ker}(u - \lambda \text{Id}_E)$ s'appelle **sous-espace propre** associé à λ , et les vecteurs x **non nuls** tels que $x \in \text{Ker}(u - \lambda \text{Id}_E)$, c'est-à-dire tels que $u(x) = \lambda x$, s'appellent les **vecteurs propres** associés à la valeur propre λ .
Un vecteur $x \in E$ est dit **propre** pour u ssi : il est **non nul**, et il existe $\lambda \in K$ tel que $u(x) = \lambda x$.

L'ensemble des valeurs spectrales de u s'appelle le *spectre* de u . Toute valeur propre de u est spectrale, mais en général une valeur spectrale n'est pas valeur propre.

Exemple 1 : Soit E le \mathbb{R} -ev $\mathcal{C}([0, 1], \mathbb{R})$ des fonctions continues de $[0, 1]$ dans \mathbb{R} . A toute $f \in E$ associons $u(f) = g$ telle que $g(x) = \int_0^x f(t) dt$ pour $x \in [0, 1]$. Alors $u \in \text{Hom}_{\mathbb{R}}(E)$, et u est injective, donc 0 n'est pas valeur propre de u . Cependant u est très loin d'être bijective puisque $u(f)$ est de classe C^1 pour toute $f \in E$. Donc 0 est valeur spectrale de u .

Cependant, lorsque la dimension de E est finie, l'injectivité et la bijectivité de $u \in \text{Hom}_K(E)$ sont équivalentes (théorème IX.6.1). Par suite :

PROPOSITION XV.1.1

|| Si E est un K -ev de dimension **finie**, le spectre de tout $u \in \text{Hom}_K(E)$ est égal à l'ensemble de ses valeurs propres.

Soit x un vecteur propre de $u \in \text{Hom}_K(E)$, le K -ev étant à nouveau non nul quelconque. Comme $x \neq 0_E$, il y a un et *un seul* $\lambda \in K$ tel que $u(x) = \lambda x$, et ce scalaire λ est évidemment une valeur propre de u , pour laquelle x est vecteur propre associé : ce scalaire s'appelle **valeur propre associée au vecteur propre x** .

Notons que 0 est valeur propre de u ssi $\text{Ker}(u) \neq \{0_E\}$, et que les vecteurs propres associés à la valeur propre 0 sont les éléments de $\text{Ker}(u) \setminus \{0_E\}$.

Enfin remarquons que si $F = \text{Ker}(u - \lambda \text{Id}_E)$ est un sous-espace propre pour l'endomorphisme u associé à la valeur propre λ , F est u -stable, et $u|_F = \lambda \text{Id}_F$.

Polynôme caractéristique

LEMME 1

|| Soit E un K -ev de dimension **finie** $n \geq 1$ et $u \in \text{Hom}_K(E)$. Si \mathcal{B} est une base ordonnée de E , le polynôme

$$\det [\text{Mat}_{\mathcal{B}}(u) - XI_n] \in K[X]$$

|| ne dépend que de u et non du choix de \mathcal{B} .

Démonstration :

On considère deux bases ordonnées \mathcal{B} et \mathcal{B}' . Soit P la matrice de passage de \mathcal{B} à \mathcal{B}' , et $M = \text{Mat}_{\mathcal{B}}(u)$, $M' = \text{Mat}_{\mathcal{B}'}(u)$. On sait que $M' = P^{-1}MP$ (corollaire du théorème XI.3.4).

Raisonnons dans la K -algèbre $\mathfrak{M}_n(K[X])$, qui contient $\mathfrak{M}_n(K)$ (et qu'on peut considérer comme sous- K -algèbre de l'algèbre des matrices carrées $\mathfrak{M}_n(L)$, où L est le corps $K(X)$). On a :

$$M' - XI_n = P^{-1}(M - XI_n)P$$

(X est un scalaire dans $\mathfrak{M}_n(L)$), d'où

$$\begin{aligned} \det(M' - XI_n) &= \det(P^{-1}) \det(M - XI_n) \det(P) = \\ &= \det(P^{-1}) \det(P) \det(M - XI_n) = \det(P^{-1}P) \det(M - XI_n) \\ &= \det(I_n) \det(M - XI_n) = \det(M - XI_n). \quad \blacksquare \end{aligned}$$

DÉFINITION XV.1.2

Soit $u \in \text{Hom}_K(E)$, où E est un K -ev de dimension finie $n \geq 1$. On appelle **polynôme caractéristique de u** le polynôme, élément de $K[X]$, que nous noterons $\chi_u(X)$, tel que

$$\chi_u(X) = \det (\text{Mat}_{\mathcal{B}}(u) - XI_n)$$

pour toute base ordonnée \mathcal{B} de E .

Exemple 2 : Si $u = \lambda \text{Id}_E$ ($\lambda \in K$), on a $\text{Mat}_{\mathcal{B}}(u) = \lambda I_n$ pour toute base \mathcal{B} de E , d'où :

$$\chi_u(X) = (\lambda - X)^n.$$

Remarquons incidemment que, pour u , tout vecteur non nul de E est vecteur propre associé à l'unique valeur propre λ .

THÉORÈME XV.1.1

Les notations étant celles de la définition XV.1.2, on a pour tout $\lambda \in K$:

$$\chi_u(\lambda) = \det (u - \lambda \text{Id}_E).$$

En particulier, les **racines de $\chi_u(X)$ dans K sont les valeurs propres de u** .

Démonstration :

On choisit une base ordonnée \mathcal{B} de E . Alors

$$\chi_u(\lambda) = \det (\text{Mat}_{\mathcal{B}}(u) - \lambda I_n) = \det (u - \lambda \text{Id}_E)$$

(théorème XIII.4.2). ■

Par définition, la **multiplicité d'une valeur propre λ de u** est sa multiplicité en tant que racine du polynôme caractéristique $\chi_u(X)$. Nous allons maintenant développer $\chi_u(X)$ (les notations étant toujours celles de la définition XV.1.2). Après avoir choisi une base ordonnée $\mathcal{B} = (e_1, \dots, e_n)$ de E , et en posant $M = \text{Mat}_{\mathcal{B}}(u) = [a_{ij}]$, on a :

$$\chi_u(X) = \det (M - XI_n) = \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ \vdots & \ddots & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - X \end{vmatrix}.$$

Soit, pour tout $i \in \llbracket 1, n \rrbracket$, $\Gamma_i \in \mathfrak{M}_{n,1}(K)$ la colonne de coordonnées $(\delta_{ji} 1_K)_{1 \leq j \leq n}$ (δ = symbole de Kronecker). Les colonnes de I

$\mathcal{C}_1(M) - X\Gamma_1, \dots, \mathcal{C}_n(M) - X\Gamma_n$. Développons alors $\chi_u(X)$ par n -linéarité en ces colonnes. On obtient :

$$(1) \quad \chi_u(X) = (-1)^n X^n + \sum_{\substack{I \subset \llbracket 1, n \rrbracket \\ I \neq \emptyset}} (-1)^{n - \text{card}(I)} X^{n - \text{card}(I)} \det(M_I),$$

où M_I est la matrice dont les colonnes sont $\mathcal{C}_i(M_I) = \mathcal{C}_i(M)$ si $i \in I$ et $\mathcal{C}_i(M_I) = \Gamma_i$ si $i \notin I$. Un calcul immédiat montre que $\det(M_I)$ est le mineur centré $\Delta_{I, I}(M)$. De sorte qu'en regroupant, au second membre de (1), les parties I de même cardinal, on a le développement cherché :

$$(2) \quad \chi_u(X) = (-1)^n X^n + \sum_{k=1}^n (-1)^k \tau_k(u) X^{n-k},$$

avec

$$(3) \quad \tau_k(u) = \sum_{\substack{I \subset \llbracket 1, n \rrbracket \\ \text{card}(I) = k}} \Delta_{I, I}(M) \quad \text{pour tout } k \in \llbracket 1, n \rrbracket.$$

Les coefficients $\tau_k(u)$ ne dépendent bien que de u puisqu'on sait que $\chi_u(X)$ est indépendant du choix de la base.

On reconnaît en particulier

$$\tau_1(u) = \sum_{i=1}^n a_{ii} = \text{Tr}(M) = \text{Tr}(u) \quad \text{et} \quad \tau_n(u) = \det(M) = \det(u).$$

On retrouve le fait que u est injective ssi 0 n'est pas valeur propre, c'est-à-dire ssi $\det(u) \neq 0$.

Les relations (3) montrent que les autres coefficients $\tau_k(u)$ sont des fonctions polynomiales de u , homogènes de degré k .

Une conséquence simple mais très importante de (2) est :

THÉORÈME XV.1.2

|| Soit E un K -ev non nul de dimension finie, et $u \in \text{Hom}_K(E)$. Alors $\chi_u(X)$ est de degré $\dim_K(E) \geq 1$. En conséquence, si K est algébriquement clos, u admet au moins une valeur propre.

Rassemblons dans le théorème suivant des résultats qui seront d'un usage quasi permanent dans la suite :

THÉORÈME XV.1.3

|| Soit E un K -ev non nul de dimension finie, et $u \in \text{Hom}_K(E)$.
 || (I) u et ${}^t u$ ont le même polynôme caractéristique.
 || (II) Si F est un sous- K -ev de E , distinct de $\{0\}$ et de E , et u -stable (i.e. $u(F) \subset F$), et si $v = u|_F$, alors $\chi_v(X)$ divise $\chi_u(X)$.

|| (III) Si F et G sont des sous- K -ev non nuls de E , u -stables, et supplémentaires dans E , alors, en posant

|| $v = u|_F$ et $w = u|_G$: $\chi_u(X) = \chi_v(X) \chi_w(X)$.

Démonstration :

(I) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $M = \text{Mat}_{\mathcal{B}}(u)$. Si $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ est la base duale de \mathcal{B} , on sait que $\text{Mat}_{\mathcal{B}}({}^t u) = {}^t M$ (théorème XII.2.2). Donc

$$\chi_u(X) = \det({}^t M - XI_n) = \det({}^t(M - XI_n)) = \det(M - XI_n) = \chi_u(X).$$

(II) On choisit une base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $\mathcal{C} = (e_1, \dots, e_p)$ soit une base de F . On a alors, par blocs :

$$M = \text{Mat}_{\mathcal{B}}(u) = \begin{bmatrix} P & R \\ \hline 0 & Q \end{bmatrix}, \text{ avec } P = \text{Mat}_{\mathcal{C}}(v), \quad Q \in \mathfrak{M}_{n-p}(K)$$

et $R \in \mathfrak{M}_{p, n-p}(K)$, d'où, grâce au théorème XIII.4.6 :

$$\chi_u(X) = \det(M - XI_n) = \det(P - XI_p) \det(Q - XI_{n-p}) = \chi_v(X) \varphi(X),$$

$$\text{avec } \varphi(X) = \det(Q - XI_{n-p}) \in K[X].$$

(III) On choisit une base $\mathcal{B} = (e_1, \dots, e_n)$ telle que $\mathcal{C} = (e_1, \dots, e_p)$ et $\mathcal{D} = (e_{p+1}, \dots, e_n)$ soient des bases de F et G respectivement, d'où

$$M = \text{Mat}_{\mathcal{B}}(u) = \begin{bmatrix} P & 0 \\ \hline 0 & Q \end{bmatrix} \text{ avec } P = \text{Mat}_{\mathcal{C}}(v) \text{ et } Q = \text{Mat}_{\mathcal{D}}(w).$$

Par suite :

$$\begin{aligned} \chi_u(X) &= \det(M - XI_n) = \det(P - XI_p) \det(Q - XI_{n-p}) = \\ &= \chi_v(X) \chi_w(X). \quad \blacksquare \end{aligned}$$

La propriété (III) du théorème XV.1.3 s'étend immédiatement au cas où $E = \bigoplus_{i=1}^p F_i$, les F_i étant des sous- K -ev non nuls et u -stables. En posant

$u_i = u|_{F_i}$ on obtient : $\chi_u(X) = \prod_{i=1}^p \chi_{u_i}(X)$. Ajoutons que deux endomorphismes semblables ont le même polynôme caractéristique.

Extension aux matrices carrées

Donnons-nous $n \in \mathbb{N}^*$. A chaque matrice $M \in \mathfrak{M}_n(K)$ associons comme à l'accoutumée l'endomorphisme $u_M \in \text{Hom}_K(K^n)$ tel que Mat

en notant \mathcal{C} la base canonique de K^n . Par définition, le **polynôme caractéristique de M** (noté $\chi_M(X)$) est celui de u_M , c'est donc $\det(M - XI_n)$.

Les **valeurs propres de M** sont celles de u_M : ce sont donc les *racines dans K* de $\chi_M(X)$.

Les **vecteurs propres de M** sont ceux de u_M .

On peut être tenté d'identifier M et u_M , mais il faut être très prudent avant de le faire car justement, quand on manipule des valeurs propres, on est presque toujours amené à faire des changements de base. Or, dans une nouvelle base de K^n , le même endomorphisme u_M est défini par une matrice M' généralement différente de M , l'un des buts recherché étant précisément que M' ait une forme « plus simple » que M . Une identification trop hâtive expose donc à des confusions...

Exercice 1 : Calculer les polynômes caractéristiques des matrices suivantes, le corps de base étant \mathbb{C} .

a) $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$ ($n \geq 2$) avec $a_{ij} = a$ si $i > j$, $a_{ij} = b$ si $i < j$ et $a_{ii} = 0$ pour tout i .

$$b) M = \begin{bmatrix} 0 & \dots & \dots & 0 & a_0 \\ & \ddots & & \vdots & a_1 \\ 1 & & & & \\ & \ddots & & & \vdots \\ 0 & & & 0 & a_{n-2} \\ \vdots & & & & \\ 0 & \dots & 0 & 1 & a_{n-1} \end{bmatrix} \quad (n \geq 3) \quad c) M = \begin{bmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{bmatrix}$$

d) $M = [a_{ij}]$ avec $a_{ij} = a$ pour $i + j \neq n + 1$ et $a_{ij} = b$ pour $i + j = n + 1$.

$$e) M = \begin{bmatrix} a & c & b \\ c & a+b & c \\ b & c & a \end{bmatrix} \quad f) M = \Gamma(a_0, \dots, a_{n-1}) = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}$$

(matrice circulante d'ordre n)

$$g) M = \begin{bmatrix} 0 & \dots & \dots & 0 & a_n \\ \vdots & & & \vdots & a_{n-1} \\ \vdots & & & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \vdots \\ a_n & a_{n-1} & \dots & a_2 & a_1^2 \end{bmatrix} \quad h) M = \begin{bmatrix} 3 & -5 & 2 & -6 \\ 0 & 5 & 0 & 4 \\ -2 & 7 & -1 & 11 \\ 0 & -4 & 0 & -3 \end{bmatrix}$$

Exercice 2 : Soit $n \in \mathbb{N}^*$. On donne (a_1, \dots, a_n) et (b_1, \dots, b_n) dans \mathbb{R}^n tels que $a_1 < a_2 < \dots < a_n$ et $(\forall i) b_i > 0$. Montrer que le polynôme caractéristique de la matrice

$$M = \begin{bmatrix} a_1 + b_1 & b_1 & \dots & b_1 \\ b_2 & a_2 + b_2 & \dots & b_2 \\ \vdots & \vdots & \ddots & \vdots \\ b_n & b_n & \dots & a_n + b_n \end{bmatrix}$$

est dissocié et à facteurs simples dans $\mathbb{R}[X]$.

Exercice 3 : Soit n et p dans \mathbb{N}^* , $n \geq p$, et deux matrices $M \in \mathfrak{M}_{n,p}(K)$ et $N = \mathfrak{M}_{p,n}(K)$. Démontrer (cf. exercice 25 du § XIII.5) :

$$\det(MN - XI_n) = (-1)^{n-p} X^{n-p} \det(NM - XI_p).$$

En particulier, si $n = p$, $\chi_{MN}(X) = \chi_{NM}(X)$.

Indication : On complète M et N avec des zéros pour les rendre carrées. Dans le cas où $n = p$, on commencera par supposer d'abord M ou N inversible.

Exercice 4 : Soit E le \mathbb{C} -ev des fonctions continues de $[0, 1]$ dans \mathbb{C} . On considère $u \in \text{Hom}_{\mathbb{C}}(E)$ qui associe, à $f \in E$, la fonction $g : [0, 1] \rightarrow \mathbb{C}$, $x \mapsto \int_0^1 \text{Min}(x, t) f(t) dt$. Valeurs propres et vecteurs propres de u ?

Exercice 5 : Soit E le \mathbb{C} -ev des fonctions continues de \mathbb{R} dans \mathbb{C} . On considère $u \in \text{Hom}_{\mathbb{C}}(E)$ qui associe, à $f \in E$, la fonction $g : \mathbb{R} \rightarrow \mathbb{C}$, $x \mapsto \int_0^\pi \sin(x-t) f(t) dt$. Valeurs propres et vecteurs propres de u ?

Exercice 6 : Soit $A \in \mathfrak{M}_n(K)$ ($n \geq 1$). On considère $B \in \mathfrak{M}_{2n}(K)$ définie ainsi par blocs : $B = \begin{bmatrix} 0 & I_n \\ A & 0 \end{bmatrix}$. Démontrer que $\chi_B(X) = (-1)^n \chi_A(X^2)$.

Exercice 7 : Soit n, p deux entiers ≥ 2 . On donne A_0, A_1, \dots, A_{p-1} dans $\mathfrak{M}_n(\mathbb{C})$ et on considère la matrice *circulante par blocs* $M \in \mathfrak{M}_{np}(\mathbb{C})$:

$$M = \begin{bmatrix} A_0 & A_1 & \dots & A_{p-1} \\ A_{p-1} & A_0 & \dots & A_{p-2} \\ \dots & \dots & \dots & \dots \\ A_1 & A_2 & \dots & A_0 \end{bmatrix}.$$

a) Calculer $\chi_M(X)$ sous forme d'un produit de facteurs de degré p (cf. exercice 19 du § XIII.5).

b) Soit $\alpha \in \mathbb{C}$, $\alpha \neq 1$. On suppose $A_k = \alpha^k A_0$ pour tout k . Donner une expression de $\chi_M(X)$ sous forme du déterminant d'une matrice carrée d'ordre n .

Exercice 8 : On donne $n \in \mathbb{N}$. On pose $\Delta_0 = 1$ et $\Delta_1 = -X$. Pour $n \geq 2$, $\Delta_n(X)$ désigne le polynôme caractéristique de la matrice de $\mathfrak{M}_n(\mathbb{C})$.

$$M_n = \begin{bmatrix} 0 & a_2 & 0 & \dots & \dots & 0 \\ -c_1 & 0 & a_3 & \dots & \dots & \vdots \\ 0 & -c_2 & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots & a_n \\ 0 & \dots & \dots & 0 & -c_{n-1} & 0 \end{bmatrix}$$

Les suites $(a_k)_{k \geq 2}$ et $(c_k)_{k \geq 1}$ sont données, à valeurs dans \mathbb{R}_+^* .

a) Établir la relation de récurrence :

$$(\forall n \geq 2) \quad \Delta_n(X) = -X \Delta_{n-1}(X) + a_n c_{n-1} \Delta_{n-2}(X).$$

En déduire que $(\forall p \in \mathbb{N}^*) \quad \Delta_{2p}(X) = F_p(X^2)$ et $\Delta_{2p+1}(X) = -X G_p(X^2)$, où F_p et G_p sont des polynômes en X à coefficients dans \mathbb{R}_+^* .

b) Montrer que : si n est pair, $\Delta_n(x)$ possède n racines imaginaires pures distinctes ; et si n est impair, $\Delta_n(x)$ admet $n-1$ racines imaginaires pures distinctes en plus de la racine simple 0.

Exercice 9 : Soit E un K -ev non nul de dimension finie.

a) Soit u et v deux endomorphismes de E *permutables* ($uv = vu$).

Montrer que tout sous-espace propre de l'un est stable par l'autre. Si K est algébriquement clos, en déduire que u et v ont au moins un vecteur propre commun.

b) Soit u et $v \in \text{Hom}_K(E)$ et $\beta \in K^*$ tel que $uv - vu = \beta u$. On suppose K de caractéristique 0. Démontrer que u est nilpotent (i.e. $\exists p \in \mathbb{N}^*$ tel que $u^p = 0$).

Indication : Parmi les très nombreuses solutions possibles, en voici une : soit $T_v \in \text{Hom}_K(\text{Hom}_K(E))$ défini par $w \mapsto wv - vw$. Montrer que si, pour $p \in \mathbb{N}^*$, $u^p \neq 0$, alors u^p est vecteur propre de T_v , préciser pour quelle valeur propre, et conclure.

c) Avec les hypothèses du b) mais en supposant K algébriquement clos, démontrer que u et v ont au moins un vecteur propre commun.

d) Sous les hypothèses du c), prouver que u et v sont simultanément tri-

Exercice 10 : Soit E un K -ev non nul de dimension finie. On donne $\alpha \in \text{Hom}_K(E)$. Soit S et T éléments de $\text{Hom}_K(\text{Hom}_K(E))$ tels que $S : u \mapsto \alpha u$ et $T : u \mapsto u\alpha$. Montrer que

$$\chi_S(X) = \chi_T(X) = (\chi_\alpha(X))^{\dim(E)}.$$

Exercice 11 : Trouver les valeurs propres et les vecteurs propres de l'endomorphisme de $\mathbb{C}[X]$ défini par :

$$P(X) \mapsto (X^2 - 1)P''(X) + (2X + 1)P'(X).$$

Exercice 12 : Soit a et b distincts dans \mathbb{R} , et E le \mathbb{R} -ev $\mathbb{R}_{2n}[X]$. Donner les valeurs propres et les vecteurs propres de l'endomorphisme u de E tel que :

$$u(P(X)) = (X - a)(X - b)P'(X) - [2nX - n(a + b)]P(X) \quad (P \in E).$$

Exercice 13 : On suppose K infini. Soit $M \in \mathfrak{M}_n(K)$ ($n \geq 2$). En fonction de $\chi_M(X)$, calculer le polynôme caractéristique de la matrice \tilde{M} des cofacteurs de M , en commençant par le cas où M est inversible.

Indication pour le cas où M n'est pas inversible : on a $\chi_{\tilde{M}}(X) = (-1)^{n-1} X^{n-1}(\tau - X)$ avec $\tau = \sum_{i=1}^n \Delta_{i,i}(M)$, $\Delta_{i,i}(M)$ étant le cofacteur d'indice (i, i) dans M .

Exercice 14 : Soit

$$A = \begin{bmatrix} a_1 & 1 & 0 & \dots & 0 \\ a_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ a_n & 0 & \dots & \dots & 0 \end{bmatrix}$$

une matrice à éléments réels, avec $a_1 > 0$, $a_n > 0$, $a_2, \dots, a_{n-1} \geq 0$. Former le polynôme caractéristique de A . Montrer que A admet une valeur propre $r > 0$, que $r < 1 + \text{Max}(a_1, \dots, a_n)$ et que, si λ est une valeur propre autre que r , $|\lambda| < r$. Montrer qu'il existe un entier k tel que A^k ait des éléments tous strictement positifs.

§ XV.2 TRIGONALISATION

Soit u un endomorphisme d'un K -ev E de dimension $n \geq 1$. Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E dans laquelle

$$\text{Mat}_{\mathcal{B}}(u) = M = [a_{ij}],$$

la matrice M' de u dans la base $\mathcal{B}' = (e'_1, \dots, e'_n)$ telle que

$$e'_i = e_{n+1-i} \quad (1 \leq i \leq n)$$

est $[a'_{ij}]$, où $a'_{i,j} = a_{n+1-i, n+1-j}$ pour tous i, j . En particulier, M est trigonale supérieure (resp. inférieure) ssi M' est trigonale inférieure (resp. supérieure). Or nous avons vu dans les chapitres précédents l'intérêt d'avoir affaire à des matrices trigonales. On peut donc se demander s'il existe des bases \mathcal{B} dans lesquelles $\text{Mat}_{\mathcal{B}}(u)$ soit *trigonale* (sans préciser s'il s'agit de matrice trigonale supérieure ou inférieure).

DÉFINITION XV.2.1

Un endomorphisme $u \in \text{Hom}_K(E)$, où E est un K -ev non nul de dimension finie, est dit **trigonalisable** ssi il existe au moins une base ordonnée \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(u)$ soit trigonale. Toute base \mathcal{B} de ce type est alors dite **trigonalisante** pour u , et on dit que u est **trigonalisé** dans une telle base.

Nous allons voir qu'il est facile de caractériser les endomorphismes trigonalisables par leurs valeurs propres.

Avec les notations ci-dessus, supposons u trigonalisable et soit, par exemple, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E trigonalisante supérieure pour u :

$$M = \text{Mat}_{\mathcal{B}}(u) = [a_{ij}]_{(i,j) \in \llbracket 1, n \rrbracket^2} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{bmatrix}.$$

On en déduit :

$$(1) \quad \chi_u(X) = \chi_M(X) = \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} - X \end{vmatrix} = \prod_{i=1}^n (a_{ii} - X).$$

Donc la liste $(a_{11}, a_{22}, \dots, a_{nn})$ est une liste des racines de $\chi_M(X)$ qui est donc, on le voit, **dissocié** dans $K[X]$. On constate également que le produit des valeurs propres $a_{11}, a_{22}, \dots, a_{nn}$ est $\det(M) = \det(u)$.

Nous allons prouver que la condition nécessaire pour $\chi_M(X)$ d'être dissocié dans $K[X]$ est aussi suffisante pour que u soit trigonalisable. Bien entendu, la définition XV.2.1 s'étend immédiatement aux matrices carrées : une matrice M de $\mathfrak{M}_n(K)$ est dite **trigonalisable** ssi l'endomorphisme $u_M \in \text{Hom}_K(K^n)$ tel que $\text{Mat}_{\mathcal{B}}(u_M) = M$ l'est. Cela implique l'existence d'une base \mathcal{B} de K^n dans laquelle la matrice de u_M est trigonale, autrement dit M est trigonalisable ssi elle est *semblable à une matrice trigonale*.

Drapeaux

Dans le K -ev E de dimension $n \geq 1$, on appelle **drapeau** toute suite (E_0, E_1, \dots, E_n) de sous- K -ev de E tels que

$$E_0 \subset E_1 \subset \dots \subset E_n \quad \text{et} \quad \dim E_i = i \quad \text{pour} \quad 0 \leq i \leq n.$$

(D'où $E_0 = \{0\}$ et $E_n = E$). Si $u \in \text{Hom}_K(E)$, un drapeau $(E_i)_{0 \leq i \leq n}$ est dit **u -stable** ssi $u(E_i) \subset E_i$ pour tout i , i.e. ssi chaque E_i est u -stable.

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Posant $E_0 = \{0\}$ et

$$E_i = \text{Vect}(e_1, \dots, e_i) \quad \text{pour} \quad i \geq 1,$$

(E_0, \dots, E_n) est un drapeau de E , dit **associé à la base \mathcal{B}** .

THÉORÈME XV.2.1

|| Soit E un K -ev de dimension $n \geq 1$ et $u \in \text{Hom}_K(E)$. Pour que u soit trigonalisable, il faut et il suffit qu'il existe dans E un drapeau u -stable.

Démonstration :

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base trigonalisante supérieure pour u , il est immédiat que le drapeau associé à \mathcal{B} est u -stable, car $u(e_j) \in \text{Vect}(e_1, \dots, e_j)$ pour tout $j \in \llbracket 1, n \rrbracket$.

Réciproquement, supposons trouvé un drapeau u -stable (E_0, E_1, \dots, E_n) . Appliquant successivement le théorème de la base incomplète, on construit à partir de $e_1 \in E_1 \setminus \{0\}$ une base $\mathcal{B} = (e_1, \dots, e_n)$ de E dont le drapeau associé est (E_0, E_1, \dots, E_n) , et alors

$$u(e_j) \in E_j = \text{Vect}(e_1, \dots, e_j) \quad \text{pour tout } j \in \llbracket 1, n \rrbracket,$$

donc $\text{Mat}_{\mathcal{B}}(u)$ est trigonale supérieure. ■

Caractérisation des endomorphismes trigonalisables

THÉORÈME XV.2.2

|| Avec les notations du théorème XV.2.1, pour que u soit **trigonalisable**, il faut et il suffit que le polynôme caractéristique $\chi_u(X)$ soit **dissocié** dans $K[X]$. En particulier, si K est **algébriquement clos**, u est **toujours trigonalisable**.

Démonstration :

On sait déjà que la condition est nécessaire. Pour voir qu'elle est suffisante, raisonnons par récurrence sur $n = \dim(E)$; la propriété étant évidente pour $n = 1$, supposons la vraie dans les K -ev de dimension $n - 1 \geq 1$, avec $\dim(E) = n$. Considérons le transposé ${}^t u$ de u , d'où ${}^t u \in \text{Hom}_K(E^*)$. Puisque $\chi_{{}^t u}(X) = \chi_u(X)$ (théorème XV.1.3), et puisque $\chi_u(X)$ est dissocié sur K , ${}^t u$ admet un vecteur propre φ , relatif à une valeur propre λ . Puisque $\varphi \neq 0$, $\text{Ker}(\varphi) = H$ est un hyperplan de E . D'autre part $\text{Ker}(\varphi)$ est u -stable, car $x \in E$ et $\varphi(x) = 0$ entraînent

$$\varphi \circ u(x) = [{}^t u(\varphi)](x) = \lambda \varphi(x) = 0.$$

Posons $v = u|_H$; $\chi_v(X)$ divise $\chi_u(X)$ (théorème XV.1.3), donc est dissocié sur K . L'hypothèse de récurrence s'applique à v car $\dim(H) = n - 1$. Soit $(E_0, E_1, \dots, E_{n-1})$ un drapeau v -stable dans H . Alors $(E_0, E_1, \dots, E_{n-1}, E)$ est un drapeau u -stable de E , ce qui achève la démonstration, compte tenu du théorème XV.2.1. ■

En passant aux matrices, il s'ensuit :

COROLLAIRE

|| Pour qu'une matrice $M \in \mathfrak{M}_n(K)$ soit **semblable** dans $\mathfrak{M}_n(K)$ à une matrice **trigonale**, il faut et il suffit que $\chi_M(X)$ soit **dissocié** dans $K[X]$. En particulier, si K est algébriquement clos, toute matrice carrée d'ordre n est semblable à une matrice trigonale.

Si K n'est pas algébriquement clos, la fin de cet énoncé tombe en défaut ; ainsi, il existe dans $\mathfrak{M}_n(\mathbb{Q})$ et $\mathfrak{M}_n(\mathbb{R})$ des matrices non trigonalisables. C'est le cas par exemple de la matrice de rotation $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$, où $\theta \in \mathbb{R}$, $\theta \not\equiv 0 \pmod{\pi}$.

Lorsque u est trigonalisable, il reste à résoudre le problème de la détermination effective de bases trigonalisantes. C'est un problème délicat que nous n'aborderons pas pour le moment ⁽¹⁾. Mais il convient de remarquer que le théorème XV.2.2 n'en a pas moins une vaste portée, car dans bien des questions, le seul fait d'être assuré de l'existence d'une base trigonalisante fournit à lui seul d'importants renseignements.

Exemple 1 : Polynômes d'endomorphismes.

Supposons u trigonalisable et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base trigonalisante supérieure, d'où

$$M = \text{Mat}_{\mathcal{B}}(u) = \begin{bmatrix} a_{11} & \dots & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{bmatrix}.$$

Posons $\lambda_i = a_{ii}$ ($1 \leq i \leq n$), d'où $\chi_u(X) = \prod_{i=1}^n (\lambda_i - X)$. Soit

$$P \in K[X], \quad P = X^p + c_1 X^{p-1} + \dots + c_p \quad (p \geq 1).$$

Pour $k \in \mathbb{N}^*$, on a :

$$M^k = \text{Mat}_{(\mathcal{B})}(u^k) = \begin{bmatrix} \lambda_1^k & \times & \times & \times \\ 0 & \ddots & & \times \\ \vdots & \ddots & \ddots & \times \\ 0 & \dots & 0 & \lambda_n^k \end{bmatrix}$$

(les termes désignés par des croix importent peu).

⁽¹⁾ Un algorithme de trigonalisation est proposé dans l'exercice n° 21 du

$$\begin{aligned} \text{Mat}_{(\mathcal{B})} P(u) &= P(M) = M^p + c_1 M^{p-1} + \dots + c_p I_n = \\ &= \begin{bmatrix} P(\lambda_1) & \times & \times & \times \\ 0 & \ddots & \ddots & \times \\ \vdots & \ddots & \ddots & \times \\ 0 & \dots & 0 & P(\lambda_n) \end{bmatrix}. \end{aligned}$$

Nous voyons donc que $P(u)$ est *trigonalisé dans la base \mathcal{B}* , et que le polynôme caractéristique de $P(u)$ est $\prod_{i=1}^n (P(\lambda_i) - X)$, d'où par exemple :

$$\text{Tr}(P(u)) = \sum_{i=1}^n P(\lambda_i) \quad \text{et} \quad \det(P(u)) = \prod_{i=1}^n P(\lambda_i).$$

Exemple 2 : Exponentielle d'endomorphisme.

Supposons ici $K = \mathbb{R}$ ou \mathbb{C} . On verra dans le *cours d'Analyse* (tome 2) que, dans le K -ev $\text{Hom}_K(E)$, la série $\sum_{k \geq 0} \frac{1}{k!} u^k$ converge, sa somme étant notée $\exp(u)$ et appelée **exponentielle de u** .

De même, pour $M \in \mathfrak{M}_n(K)$, la série $\sum_{k \geq 0} \frac{1}{k!} M^k$ converge, sa somme étant notée $\exp(M)$ et appelée **exponentielle de M** . Ces notations ne doivent cependant pas faire illusion. Ce n'est que dans le cas où u et v commutent ($uv = vu$) que l'on est sûr de retrouver la propriété

$$\exp(u + v) = \exp(u) \exp(v)$$

(démonstration dans le tome 2). On en déduit, à partir de $\exp(0) = \text{Id}_E$ que $\exp(u)$ est inversible et que $(\exp(u))^{-1} = \exp(-u)$. Enfin, si $t \in K$, $\exp(t \text{Id}_E) = e^t \text{Id}_E$.

Si \mathcal{B} est une base ordonnée de E et $M = \text{Mat}_{\mathcal{B}}(u)$, on a : $\text{Mat}_{\mathcal{B}}(\exp(u)) = \exp(M)$. Supposons alors u trigonalisable, et soit \mathcal{B} une base trigonalisante supérieure. Adoptons les notations de l'exemple 1. On voit que $\exp(M)$ est *trigonale supérieure*, sa diagonale principale étant

$$\left(\sum_{k \geq 0} \frac{\lambda_1^k}{k!}, \dots, \sum_{k \geq 0} \frac{\lambda_n^k}{k!} \right), \quad \text{c'est-à-dire} \quad (e^{\lambda_1}, \dots, e^{\lambda_n}).$$

Il s'ensuit que $\exp(u)$ est trigonalisable, et que

$$\chi_{\exp(u)}(X) = \prod_{k=1}^n (e^{\lambda_k} - X).$$

On en déduit en particulier que

$$\det(\exp(u)) = \prod_{k=1}^n e^{\lambda_k} = e^{\sum_{k=1}^n \lambda_k} = e^{\text{Tr}(u)}.$$

Dans les exercices ci-après, E est un K -ev de dimension finie $n \geq 1$.

Exercice 1 : Soit α et β deux endomorphismes trigonalisables dans E , de polynômes caractéristiques $\prod_{i=1}^n (\lambda_i - X)$ et $\prod_{i=1}^n (\mu_i - X)$.

a) On considère $T \in \text{Hom}_K(\text{Hom}_K(E))$ tel que $T(u) = \alpha u \beta$ pour tout $u \in \text{Hom}_K(E)$. Montrer que T est trigonalisable, et que

$$\chi_T(X) = \prod_{(i,j) \in \llbracket 1, n \rrbracket^2} (\lambda_i \mu_j - X).$$

b) Etudier l'endomorphisme $S \in \text{Hom}_K(\text{Hom}_K(E))$ tel que $S(u) = \alpha u - u \beta$ pour tout $u \in \text{Hom}_K(E)$.

Exercice 2 : a) Soit u et v dans $\text{Hom}_K(E)$ *permutables* et trigonalisables. Montrer qu'il existe une base \mathcal{B} de E telles que $\text{Mat}_{\mathcal{B}}(u)$ et $\text{Mat}_{\mathcal{B}}(v)$ soient toutes deux trigonales supérieures.

b) Soit \mathcal{H} une partie non vide de $\text{Hom}_K(E)$ telle que : pour tous u et v dans \mathcal{H} , on ait $uv = vu$, et que chaque $u \in \mathcal{H}$ soit trigonalisable. Montrer qu'il existe une base \mathcal{B} de E telle que $(\forall u \in \mathcal{H}) \text{Mat}_{\mathcal{B}}(u)$ est trigonale supérieure.

Exercice 3 : a) Si $u \in \text{Hom}_K(E)$ est trigonalisable et inversible, et si $\chi_u(X) = \prod_{i=1}^n (\lambda_i - X)$, déterminer $\chi_{u^{-1}}(X)$.

b) On suppose $u \in \text{Hom}_K(E)$ trigonalisable, et $\chi_u(X) = \prod_{i=1}^n (\lambda_i - X)$. Soit $\Phi \in K(X)$ une fraction rationnelle donc aucun λ_i n'est pôle. Montrer comment on peut définir $\Phi(u) = v$ et montrer qu'on a :

$$\chi_v(X) = \prod_{i=1}^n (\Phi(\lambda_i) - X).$$

Exercice 4 : On suppose $K = \mathbb{C}$ (ou plus généralement, K de caractéristique 0 et algébriquement clos). Soit $M \in \mathfrak{M}_n(K)$ ($n \geq 1$). Prouver :

$$\chi_M(X) = (-1)^n X^n \Leftrightarrow (\forall k \in \mathbb{N}^*) \text{Tr}(M^k) = 0 \Leftrightarrow M \text{ est nilpotente.}$$

Exercice 5 : On suppose $K = \mathbb{C}$ et on donne $M \in \mathfrak{M}_n(\mathbb{C})$ ($n \geq 1$) de polynôme caractéristique $\chi_M(X) = \prod_{i=1}^n (\lambda_i - X)$. Montrer que les λ_i sont tous distincts ssi

$$\det(\text{Tr}(M^{i+j-2}))_{(i,j) \in \llbracket 1, n \rrbracket^2} \neq 0.$$

Exercice 6 : Soit M et N dans $\mathfrak{M}_n(\mathbb{C})$ ($n \geq 1$). Pour $k \in \mathbb{N}^*$, comparer $\text{Tr}((MN)^k)$ et $\text{Tr}((NM)^k)$. En déduire que $\chi_{MN}(X) = \chi_{NM}(X)$, en appliquant les résultats du § X.4.

Exercice 7 : Soit M et N dans $\mathfrak{M}_n(\mathbb{C})$ ($n \geq 1$). Montrer que, pour que M et N n'aient aucune valeur propre commune, il faut et il suffit que $\chi_M(N)$ soit inversible.

Exercice 8 : Le corps K est supposé de caractéristique $p \geq 3$. Soit $(a_0, a_1, \dots, a_{p-1}) \in K^p$. On considère la matrice circulante

$$M = \Gamma(a_0, \dots, a_{p-1}) \in \mathfrak{M}_p(K), \quad M = \begin{bmatrix} a_0 & a_1 & \dots & a_{p-1} \\ a_{p-1} & a_0 & \dots & a_{p-2} \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}.$$

Calculer le polynôme caractéristique et le déterminant de M , et retrouver ainsi le résultat de l'exercice n° 24 du § XIII.5.

Indication : Soit $S = \Gamma(0, 1, 0, \dots, 0)$. Remarquer que $M = \sum a_k S^k$ est donc trigonalisable.

Exercice 9 : Soit F un ensemble d'endomorphismes de E . On dit que F est trigonalisable (ou que la *réduction simultanée* à la forme trigonale est possible pour F) s'il existe

dans laquelle la matrice de tout $u \in F$ soit trigonale supérieure. Soit E' un sous- K -ev de E , F -stable (i.e. tel que $u(E') \subset E'$ pour tout $u \in F$). Tout u induit donc un endomorphisme u' de E' et un endomorphisme u de E/E' . Soit F' l'ensemble des u' et \bar{F} l'ensemble des \bar{u} .

On suppose F' trigonalisable dans E' et \bar{F} trigonalisable dans E/E' . Montrer que F est trigonalisable dans E . En déduire une autre démonstration du théorème XV.2.2.

Retrouver également les résultats de l'exercice 2.

Exercice 10 : Soit K un corps de caractéristique 0 algébriquement clos. Pour qu'une matrice U de $\mathfrak{M}_n(K)$ soit unipotente (i.e. pour que $I_n - U$ soit nilpotente), il faut et il suffit que la seule valeur propre de U soit 1_K . Quel est le polynôme caractéristique de U ?

Si K est un corps de caractéristique $p \neq 0$, montrer que U est unipotente ssi il existe un entier $\alpha \geq 0$ tel que $U^{p^\alpha} = I_n$.

Exercice 11 : Soit $n \in \mathbb{N}^*$; on donne A et B dans $\mathfrak{M}_n(\mathbb{C})$. Montrer :

$$(\exists P \in \mathfrak{M}_n(\mathbb{C}) \setminus \{0\} \mid AP = PB) \Leftrightarrow (A \text{ et } B \text{ ont une valeur propre commune}).$$

§ XV.3 SOUS-ESPACES PROPRES

Indépendance des sous-espaces propres

THÉORÈME XV.3.1

Soit u un endomorphisme d'un K -ev E et \mathcal{V} l'ensemble des valeurs propres de u , supposé non vide. Pour $\lambda \in \mathcal{V}$ soit E_λ le sous-espace propre pour u associé à λ (i.e. $E_\lambda = \text{Ker}(u - \lambda \text{Id}_E)$). Alors les sous- K -ev $(E_\lambda)_{\lambda \in \mathcal{V}}$ sont linéairement indépendants.

Démonstration :

Il s'agit de montrer que si Λ est une partie finie quelconque non vide de \mathcal{V} , si $x_\lambda \in E_\lambda$ pour tout $\lambda \in \Lambda$, et si $\sum_{\lambda \in \Lambda} x_\lambda = 0$,

alors $x_\lambda = 0$ pour tout $\lambda \in \Lambda$.

Raisonnons par l'absurde en supposant qu'il existe $q \in \mathbb{N}^*$ et des vecteurs propres x_1, \dots, x_q de u , associés à des valeurs propres distinctes $\lambda_1, \dots, \lambda_q$ respectivement, et tels que $x_1 + \dots + x_q = 0$. Un tel entier q est nécessairement ≥ 2 . Soit alors ν le minimum de tous ces entiers q , et soit y_1, \dots, y_ν des vecteurs propres de u associés respectivement aux valeurs propres μ_1, \dots, μ_ν toutes distinctes, et tels que $y_1 + \dots + y_\nu = 0$. On a donc $\nu \geq 2$. De plus $u(y_1 + \dots + y_\nu) = 0 = \sum_{i=1}^{\nu} \mu_i y_i$, et aussi

$$\mu_1(y_1 + \dots + y_\nu) = 0 = \sum_{i=1}^{\nu} \mu_1 y_i, \text{ d'où par différence :}$$

$$(1) \quad \sum_{i=2}^{\nu} (\mu_i - \mu_1) y_i = 0.$$

Mais $\mu_i - \mu_1 \neq 0$ pour $i \in \llbracket 2, \nu \rrbracket$, et y_2, \dots, y_ν sont *non nuls*, d'où $(\mu_i - \mu_1) y_i \neq 0$ pour $i \in \llbracket 2, \nu \rrbracket$ et donc (1) contredit le fait que ν était le plus petit des entiers q . Cette contradiction achève la démonstration. ■

Dimension d'un sous-espace propre

THÉORÈME XV.3.2

|| Soit E un K -ev de dimension finie $n \geq 1$, $u \in \text{Hom}_K(E)$ et λ une valeur propre de u , de **multiplicité** α . Alors

$$\dim (\text{Ker} (u - \lambda \text{Id}_E)) \leq \alpha .$$

Démonstration :

Soit $F = \text{Ker} (u - \lambda \text{Id}_E)$, d'où $F \neq \{0\}$. F est u -stable et $v = u|_F = \lambda \text{Id}_E$. Nous savons que $\chi_v(X)$ divise $\chi_u(X)$ (théorème XV.1.3) et que $\chi_v(X) = (\lambda - X)^{\dim(F)}$, d'après l'exemple 2 du § XV.1. Donc, $(\lambda - X)^{\dim(F)}$ divise $\chi_u(X)$, d'où $\dim(F) \leq \alpha$. ■

Avec les notations de ce théorème XV.3.2, la *formule du rang* donne immédiatement une expression de $\dim F$ ($F = \text{Ker} (u - \lambda \text{Id}_E)$) :

$$(2) \quad \boxed{\dim(F) = n - \text{rg}(u - \lambda \text{Id}_E)} .$$

Si l'on connaît λ , la relation (2) ramène le calcul de $\dim(F)$ à celui du rang de $u - \lambda \text{Id}_E$, donc, après avoir fait choix d'une base $\mathcal{B} = (e_1, \dots, e_n)$ de E , à celui du rang de la matrice $M - \lambda I_n$, où $M = \text{Mat}_{\mathcal{B}}(u)$:

$$(3) \quad \boxed{\dim(F) = n - \text{rg}(M - \lambda I_n)} .$$

Avec cette base, la détermination du sous-espace propre F lui-même se ramène à la *résolution du système linéaire et homogène*

$$(4) \quad \boxed{(M - \lambda I_n) X = 0}$$

où $X \in \mathfrak{M}_{n,1}(K)$ est la matrice-colonne des inconnues x_1, \dots, x_n : les éléments de F sont exactement les $\sum_{i=1}^n x_i e_i$ pour (x_1, \dots, x_n) solution de (4).

Dans la pratique, u sera connu par sa matrice M dans une base $\mathcal{B} = (e_1, \dots, e_n)$, et donc on pourra rechercher le sous-espace propre associé à la valeur propre λ en résolvant (4), la dimension de ce sous-espace étant donnée par (3) (*méthode du système linéaire*). Notons que le théorème XV.3.2 fournit une information *a priori* sur le rang

Exemple 1 : On suppose $n = 3$, u étant donné par sa matrice M dans une base $\mathcal{B} = (e_1, e_2, e_3)$ de E :

$$M = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

avec $a \in K$. On a $\chi_n(X) = -X(X-1)^2$. Soit

$$F = \text{Ker}(u) \quad \text{et} \quad G = \text{Ker}(u - \text{Id}_E)$$

les sous-espaces propres relatifs respectivement aux valeurs propres 0 et 1. Il est évident que $F = Ke_3$.

On a

$$M - I_n = \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \text{d'où} \quad \text{rg}(M - I_n) = \begin{cases} 2 & \text{si } a \neq 0 \\ 1 & \text{si } a = 0 \end{cases}.$$

Par suite, $\dim(G) = 1$ si $a \neq 0$ et $\dim(G) = 2$, si $a = 0$. Dans chacun de ces cas, la résolution du système (4) est immédiate et fournit $G = Ke_1$ si $a \neq 0$ et $G = \text{Vect}(e_1, e_2)$ si $a = 0$. On vérifie bien que $F \cap G = \{0\}$.

Endomorphismes diagonalisables

Nous avons vu que les sous-espaces propres, s'il en existe, de $u \in \text{Hom}_K(E)$ sont linéairement indépendants, mais la somme (qui est donc *directe*) de ces sous-espaces n'est en général pas égale à E . Il est évident que la situation où cette somme est exactement E est particulièrement remarquable par sa simplicité, et c'est, statistiquement parlant, le cas le plus « fréquent » et qui mérite en tout cas qu'on s'y arrête.

PROPOSITION XV.3.1

|| Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. Les conditions suivantes sont équivalentes :

- (I) E est somme vectorielle des sous-espaces propres de u .
- (II) E est somme directe interne des sous-espaces propres de u .
- (III) E admet au moins une base formée de vecteurs propres de u .
- (IV) Il existe une base de E dans laquelle la matrice u est diagonale.
- (V) u est une somme directe interne d'homothéties.

Démonstration (abrégée) :

D'abord (I) \Leftrightarrow (II) est conséquence du théorème XV.3.1. Il est clair que (III) \Leftrightarrow (IV).

Si F est un sous-espace propre de u , associé à la valeur propre λ , on sait que $u|_F = \lambda \text{Id}_F$, d'où (II) \Rightarrow (III) et (II) \Rightarrow (V).

Il est enfin évident que (III) \Rightarrow (I), et aussi que (V) \Rightarrow

DÉFINITION XV.3.1

Les notations étant celles de la proposition XV.3.1 :
 (I) u est dit **diagonalisable** ssi u vérifie les conditions équivalentes de cette proposition.
 (II) Une matrice $M \in \mathfrak{M}_n(K)$ est dite **diagonalisable** ssi l'endomorphisme u_M qui lui est associé dans la base canonique de K^n est diagonalisable.

Lorsque u est diagonalisable, les bases ordonnées de E dans lesquelles la matrice de u est diagonale sont les bases ordonnées de vecteurs propres de u . Si M est alors matrice de u dans une base ordonnée quelconque de E , la formule du changement de base (formule (11) du § XI.3) montre que M est semblable à une matrice diagonale. Réciproquement, pour la même raison, si M est semblable à une matrice diagonale, u est diagonalisable, d'où :

PROPOSITION XV.3.2

Les notations étant celles de la proposition XV.3.1, soit \mathcal{B} une base ordonnée de E et $M = \text{Mat}_{\mathcal{B}}(u)$.
 Pour que u soit **diagonalisable**, il faut et il suffit que M soit **semblable à une matrice diagonale**. En particulier, une matrice $N \in \mathfrak{M}_n(K)$ est diagonalisable ssi elle est semblable à une matrice diagonale.

Remarquons qu'un endomorphisme diagonalisable est *a fortiori* trigonalisable. Voici un cas remarquable d'endomorphisme diagonalisable :

THÉORÈME XV.3.3

Soit E un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$. Si u possède n valeurs propres distinctes, alors u est **diagonalisable**, et ses sous-espaces propres sont de dimension 1.

Démonstration :

Soit $(\lambda_1, \dots, \lambda_n)$ une liste de ces valeurs propres, et $E_i = \text{Ker}(u - \lambda_i \text{Id}_E)$ pour $i \in \llbracket 1, n \rrbracket$. On sait que $(\forall i) \dim(E_i) \geq 1$, car $E_i \neq \{0\}$. De plus $\dim(E_i) \leq 1$, car les valeurs propres de u sont nécessairement simples ($\chi_u(X) = \prod_{i=1}^n (\lambda_i - X)$). Donc $\dim(E_i) = 1$ pour tout i ; les E_i étant indépendants, il s'ensuit

$$\dim(E_1 \oplus E_2 \oplus \dots \oplus E_n) = \sum_{i=1}^n \dim(E_i) = n = \dim(E),$$

d'où

$$\bigoplus_{i=1}^n E_i = E. \quad \blacksquare$$

Remarque 1 : Le raisonnement précédent prouve, accessoirement, que le sous-espace propre associé à toute valeur propre *simple* d'un endomorphisme d'un K -ev de dimension finie est toujours de dimension 1. Pour reconnaître les autres endomorphismes diagonalisables, on pourra s'aider du :

THÉORÈME XV.3.4

Soit E un K -ev de dimension finie $n \geq 1$, $u \in \text{Hom}_K(E)$ et \mathcal{S} l'ensemble des valeurs propres de u . Pour $\lambda \in \mathcal{S}$, on pose $E_\lambda = \text{Ker}(u - \lambda \text{Id}_E)$ et $m_\lambda =$ multiplicité de la valeur propre λ . Il y a équivalence entre les deux propriétés suivantes :

(I) u est **diagonalisable**.
 (II) $\chi_u(X)$ est dissocié dans $K[X]$, et $(\forall \lambda \in \mathcal{S}) m_\lambda = \dim(E_\lambda)$.

Démonstration :

Montrons d'abord que (I) \Rightarrow (II). Si u est diagonalisable, on a

$$(5) \quad E = \bigoplus_{\lambda \in \mathcal{S}} E_\lambda .$$

Soit $v_\lambda = u|_{E_\lambda}$ ($\lambda \in \mathcal{S}$). On sait que $v_\lambda = \lambda \text{Id}_{E_\lambda}$, d'où

$$\chi_{v_\lambda}(X) = (\lambda - X)^{\dim(E_\lambda)} .$$

Le théorème XV.1.3 et l'égalité (5) montrent alors que

$$(6) \quad \chi_u(X) = \prod_{\lambda \in \mathcal{S}} (\lambda - X)^{\dim(E_\lambda)} .$$

Cela prouve que $\chi_u(X)$ est dissocié sur K et il résulte de la définition même des m_λ que

$$(7) \quad \chi_u(X) = \prod_{\lambda \in \mathcal{S}} (\lambda - X)^{m_\lambda} .$$

En rapprochant (6) et (7) on obtient (II).

Montrons ensuite que (II) \Rightarrow (I). Supposons (II) vrai. Puisque $\chi_u(X)$ est dissocié sur K , on a

$$\chi_u(X) = \prod_{\lambda \in \mathcal{S}} (\lambda - X)^{m_\lambda} ,$$

$$\text{d'où} \quad n = \deg(\chi_u(X)) = \sum_{\lambda \in \mathcal{S}} m_\lambda = \sum_{\lambda \in \mathcal{S}} \dim(E_\lambda) .$$

Puisque les E_λ sont indépendants (théorème XV.3.1), on a :

$$n = \sum_{\lambda \in \mathcal{S}} \dim(E_\lambda) = \dim \left(\bigoplus_{\lambda \in \mathcal{S}} E_\lambda \right) = \dim(E).$$

Donc $\bigoplus_{\lambda \in \mathcal{S}} E_\lambda = E$, d'où (I). ■

Les théorèmes obtenus dans ce paragraphe sont déjà très performants, et leurs applications sont nombreuses.

Exemple 2 : Le corps de base est \mathbb{C} . On donne $(a, b, c) \in \mathbb{C}^3$ et la matrice $M \in \mathfrak{M}_3(\mathbb{C})$ suivante :

$$M = \begin{bmatrix} a & c & b \\ c & a+b & c \\ b & c & a \end{bmatrix}.$$

Montrer que M est diagonalisable. Trouver $P \in GL(3, \mathbb{C})$ telle que PMP^{-1} soit diagonale, et calculer M^n pour $n \in \mathbb{N}^*$.

Solution : Nous noterons \mathcal{C} la base canonique (e_1, e_2, e_3) du \mathbb{C} -ev \mathbb{C}^3 et u l'endomorphisme de \mathbb{C}^3 tel que $\text{Mat}_{\mathcal{C}}(u) = M$. On a d'abord

$$\chi_u(X) = \chi_M(X) = (\lambda - X)(\mu - X)(\nu - X),$$

$$\text{avec } \lambda = a - b, \quad \mu = a + b + c\sqrt{2}, \quad \nu = a + b - c\sqrt{2}.$$

Ces racines sont distinctes ssi $\Delta \neq 0$, en posant $\Delta(a, b, c) = c(2b^2 - c^2)$. Etudions d'abord le cas où $\Delta \neq 0$. Alors u est diagonalisable puisqu'il a 3 valeurs propres distinctes. Cherchons une base de vecteurs propres de u , en appliquant dans la base \mathcal{C} la méthode du système linéaire :

$$\begin{aligned} M - \lambda I_3 &= \begin{bmatrix} b & c & b \\ c & 2b & c \\ b & c & b \end{bmatrix}; \\ M - \mu I_3 &= \begin{bmatrix} -b - c\sqrt{2} & c & b \\ c & -c\sqrt{2} & c \\ b & c & -b - c\sqrt{2} \end{bmatrix}; \\ M - \nu I_3 &= \begin{bmatrix} -b + c\sqrt{2} & c & b \\ c & c\sqrt{2} & c \\ b & c & -b + c\sqrt{2} \end{bmatrix}. \end{aligned}$$

On sait à l'avance que ces trois matrices sont de rang 2, car les valeurs propres de u sont simples. Les systèmes linéaires et homogènes qu'elles définissent se résolvent immédiatement et fournissent la base

propres de $u : \mathcal{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$, où

$$\varepsilon_1 = e_1 - e_3, \quad \varepsilon_2 = e_1 + \sqrt{2} e_2 + e_3, \quad \varepsilon_3 = e_1 - \sqrt{2} e_2 + e_3 ;$$

d'où la matrice de passage

$$P = P_{\mathcal{E}, \mathcal{B}} : P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ -1 & 1 & 1 \end{bmatrix}.$$

On a donc, les vecteurs propres $\varepsilon_1, \varepsilon_2, \varepsilon_3$ étant respectivement associés à λ, μ, ν , l'égalité

$$(8) \quad M = P D P^{-1},$$

où $D = \text{Diag}(\lambda, \mu, \nu)$. On en déduit, pour $n \in \mathbb{N}^*$

$$(9) \quad \boxed{M^n = P D^n P^{-1}} \quad (M^n = \text{Mat}_{\mathcal{E}}(u^n), D^n = \text{Mat}_{\mathcal{B}}(u^n) = \text{Diag}(\lambda^n, \mu^n, \nu^n)).$$

Le calcul demandé de M^n s'en déduit en effectuant le produit des trois matrices P, D^n et P^{-1} , qui ne présente guère de difficulté, sauf le calcul de P^{-1} qui oblige à résoudre un système linéaire, heureusement simple. Mais dans ce cas précis, on peut même éviter d'explicitier P^{-1} . En effet on remarque que

$$(10) \quad a = \frac{2\lambda + \mu + \nu}{4}, \quad b = \frac{-2\lambda + \mu + \nu}{4}, \quad c = \frac{\sqrt{2}(\mu - \nu)}{4},$$

et que la matrice P ne dépend pas de λ, μ, ν . Pour obtenir M^n , il suffit donc de remplacer, dans les expressions (10) de a, b, c , les nombres λ, μ et ν par λ^n, μ^n et ν^n : si l'on pose $M = M_1(\lambda, \mu, \nu)$, on aura donc $M^n = M_1(\lambda^n, \mu^n, \nu^n)$, ce qui donne directement

$$(11) \quad M^n = \frac{1}{4} \begin{bmatrix} 2\lambda^n + \mu^n + \nu^n & \sqrt{2}(\mu^n - \nu^n) & -2\lambda^n + \mu^n + \nu^n \\ \sqrt{2}(\mu^n - \nu^n) & 2(\mu^n + \nu^n) & \sqrt{2}(\mu^n - \nu^n) \\ -2\lambda^n + \mu^n + \nu^n & \sqrt{2}(\mu^n - \nu^n) & 2\lambda^n + \mu^n + \nu^n \end{bmatrix}.$$

L'égalité (8) s'écrit aussi $MP = PD$. On peut vérifier directement qu'elle reste vraie sans aucune restriction sur a, b, c , et par conséquent nous n'avons pas à faire une étude particulière pour $\Delta = 0$ puisqu'on est sûr que u est encore diagonalisable, que la base $\mathcal{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ reste une base de vecteurs propres, et que les formules (9) et (11) donnant M^n restent valables.

La méthode résumée par la formule (9) donnée dans cet exemple s'appelle le **calcul de M^n par diagonalisation**. Elle s'applique

avec toute matrice **diagonalisable**, mais ce n'est pas toujours la méthode la plus rapide et l'on ne perdra pas de vue que pour certaines matrices diagonalisables, un calcul direct de M^n beaucoup plus simple est parfois possible. Par exemple soit $H \in \mathfrak{M}_n(\mathbb{C})$ ($n \geq 2$) la matrice dont tous les coefficients sont égaux à 1. Il est clair que $H^2 = nH$, d'où $H^k = n^{k-1} H$ pour $k \in \mathbb{N}^*$. Alors, pour tous a, b dans \mathbb{C} , la matrice $M = aI_n + bH$ est diagonalisable, mais il serait très maladroit de diagonaliser M pour calculer M^k . En effet, du fait que aI_n et bH sont permutables, $M^k = (aI_n + bH)^k$ se calcule immédiatement par la formule du binôme :

$$M^k = a^k I_n + \sum_{j=1}^k \binom{k}{j} a^{k-j} b^j n^{j-1} H.$$

Exemple 3 : Calculer l'exponentielle de la matrice M de l'exemple 2.

Solution : On sait que $\text{Mat}_{\mathcal{B}}(\exp(u)) = \exp(M)$, et $\text{Mat}_{\mathcal{B}}(\exp(u)) = \exp(D)$, d'où

$$(12) \quad \boxed{\exp(M) = P [\exp(D)] P^{-1}}, \text{ avec } \exp(D) = \text{Diag}(e^\lambda, e^\mu, e^\nu).$$

Donc $\exp(M)$ s'en déduit par le produit des 3 matrices, P , $\exp(D)$ et P^{-1} . Mais il est encore possible ici d'économiser le calcul de P^{-1} . En effet la relation (notations de l'exemple 2)

$$M_1(\lambda, \mu, \nu) = P \text{Diag}(\lambda, \mu, \nu) P^{-1}$$

où P est indépendante de (λ, μ, ν) est vraie avec tous ces triplets. Elle reste donc vraie en y remplaçant (λ, μ, ν) par $(e^\lambda, e^\mu, e^\nu)$, d'où avec un minimum d'efforts :

$$\begin{aligned} \exp(M) &= M_1(e^\lambda, e^\mu, e^\nu) = \\ &= \frac{1}{4} \begin{pmatrix} 2e^\lambda + e^\mu + e^\nu & \sqrt{2}(e^\mu - e^\nu) & -2e^\lambda + e^\mu + e^\nu \\ \sqrt{2}(e^\mu - e^\nu) & 2(e^\mu + e^\nu) & \sqrt{2}(e^\mu - e^\nu) \\ -2e^\lambda + e^\mu + e^\nu & \sqrt{2}(e^\mu - e^\nu) & 2e^\lambda + e^\mu + e^\nu \end{pmatrix}. \end{aligned}$$

Le principe de la méthode exposé dans cet exemple et contenu dans la formule (12) est évidemment valable avec toute matrice (resp. tout endomorphisme) **diagonalisable** et permet d'obtenir son exponentielle. Mais ici encore il peut arriver qu'un calcul direct plus rapide soit possible. Donnons-en un exemple : soit $\theta \in \mathbb{R}$ et

$$M = \begin{bmatrix} \text{ch } \theta & \text{sh } \theta \\ \text{sh } \theta & \text{ch } \theta \end{bmatrix}$$

qui est diagonalisable sur \mathbb{R} . Alors, pour tout $k \in \mathbb{N}$, il est facile d'obtenir

$$M^k = \begin{bmatrix} \operatorname{ch} k\theta & \operatorname{sh} k\theta \\ \operatorname{sh} k\theta & \operatorname{ch} k\theta \end{bmatrix} = \frac{1}{2} e^{k\theta} U + \frac{1}{2} e^{-k\theta} V,$$

avec
$$U = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{et} \quad V = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

d'où immédiatement, sans avoir besoin de diagonaliser M ,

$$\exp(M) = \frac{1}{2} (e^{e^\theta} U + e^{e^{-\theta}} V).$$

Exemple 4 : Relations de récurrence linéaires à coefficients constants.

Donnons-nous un entier $k \geq 1$ et des éléments $\zeta_0, \zeta_1, \dots, \zeta_{k-1}$ de K , avec $\zeta \neq 0$. Notons \mathcal{E} le K -ev $K^{\mathbb{N}}$ des suites à valeurs dans K , et E le sous-ensemble de \mathcal{E} formé des suites $(u_n)_{n \in \mathbb{N}}$ telles que

$$(13) \quad (\forall n \geq 0) \quad u_{n+k} = \zeta_{k-1} u_{n+k-1} + \dots + \zeta_0 u_n.$$

La relation (13) est appelée une **relation de récurrence linéaire à coefficients constants** (les coefficients étant les ζ_i).

On vérifie d'abord que E est un sous- K -ev de \mathcal{E} .

En raisonnant par récurrence, on voit ensuite que si l'on donne

$$\mathbf{a} = (a_1, \dots, a_k) \in K^k,$$

on a une, et une seule, suite $\Phi(\mathbf{a}) = (u_n)_{n \in \mathbb{N}}$ appartenant à E , telle que

$$u_0 = a_1, u_1 = a_2, \dots, u_{k-1} = a_k.$$

De plus l'application $\Phi : K^k \rightarrow E, \mathbf{a} \mapsto \Phi(\mathbf{a})$ est K -linéaire, et elle est évidemment surjective puisque si $v = (v_n) \in E$, on a $v = \Phi(\mathbf{a})$, avec $\mathbf{a} = (v_0, v_1, \dots, v_{k-1})$.

On obtient donc cette propriété très simple :

THÉORÈME XV.3.5

|| Le K -ev E des solutions dans $K^{\mathbb{N}}$ de la relation de récurrence (13) est de dimension finie, égale à k . L'application $\Phi : K^k \rightarrow E$ associant, à chaque $\mathbf{a} = (a_1, \dots, a_k) \in K^k$, l'unique suite $(u_n) \in E$ telle que $u_i = a_{i+1}$ pour $i \in \llbracket 0, k-1 \rrbracket$ est un isomorphisme de K -ev.

Afin d'étudier E , observons que si $u = (u_n) \in E$, la suite $\tau(u) = (v_n)$ telle que $v_n = u_{n+1}$ pour tout $n \in \mathbb{N}$ appartient encore à E . Il est clair que τ ainsi définie est K -linéaire, donc $\tau \in \operatorname{Hom}_K(E)$.

Cherchons les valeurs propres de τ : un scalaire $\lambda \in K$ est valeur propre ssi, pour au moins une suite $(u_n) \in E \setminus \{0\}$, on a :

$$(\forall n) \quad u_{n+1} = \lambda u_n, \quad \text{c'est-à-dire} \quad (\forall n) \quad u_n = \lambda^n u_0.$$

L'appartenance à $E \setminus \{0\}$ d'une telle suite se traduit par les conditions :

$$u_0 \neq 0 \quad \text{et} \quad (\forall n) \quad (\lambda^{n+k} - \zeta_{k-1} \lambda^{n+k-1} - \dots - \zeta_0 \lambda^n) u_0 = 0,$$

ce qui, compte tenu du fait que $\zeta_0 \neq 0$, entraîne $\lambda \neq 0$ (faire $n = 0$), et donc, après division par $\lambda^n u_0$:

$$(14) \quad \lambda^k - \zeta_{k-1} \lambda^{k-1} - \dots - \zeta_0 = 0,$$

la réciproque étant évidente.

En procédant à la synthèse, on a donc le résultat suivant : les valeurs propres de τ sont les racines dans K du polynôme (dit *caractéristique* pour (13))

$$P(X) = X^k - \zeta_{k-1} X^{k-1} - \dots - \zeta_0$$

et les vecteurs propres associés à une valeur propre λ sont les suites géométriques $(u_n) = (c\lambda^n)_{n \in \mathbb{N}}$, où $c \in K^*$ est arbitraire. On constate que le sous-espace propre associé à la valeur propre λ est ici toujours de dimension 1.

Soit $\mathcal{C} = (e_1, \dots, e_k)$ la *base canonique* de K^k et \mathcal{B} la base de E image \mathcal{C} par Φ . On a :

$$\text{Mat}_{\mathcal{B}}(\tau) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \vdots \\ \zeta_0 & \zeta_1 & \dots & \dots & \zeta_{k-1} \end{bmatrix}$$

d'où, facilement, le polynôme caractéristique

$$\chi_{\tau}(X) = (-1)^k P(X),$$

ce qui explique et confirme le résultat obtenu ci-dessus.

Plaçons-nous dans l'hypothèse où $P(X)$ est **dissocié et à racines simples** dans $K[X]$: $P(X) = \prod_{i=1}^k (\lambda_i - X)$. Dans ce cas, par application du théorème XV.3.3, on a une *base de* E en prenant les k suites géométriques

$$s^{(i)} = (\lambda_i^n)_{n \in \mathbb{N}} \quad (1 \leq i \leq k).$$

On peut se servir de cette base pour en déduire la suite

vérifiant $u_0 = a_1, \dots, u_{k-1} = a_k$ pour (a_1, \dots, a_k) donné dans K^k : les coordonnées c_1, c_2, \dots, c_k de u dans la base $(s^{(i)})_{1 \leq i \leq k}$ sont déterminées par le système linéaire de Cramer :

$$(15) \quad c_1(\lambda_1)^{j-1} + c_2(\lambda_2)^{j-1} + \dots + c_k(\lambda_k)^{j-1} = a_j \quad (1 \leq j \leq k),$$

qui exprime que les termes de rang $\leq k-1$ de u sont a_1, \dots, a_k , et dont la matrice $[(\lambda_i)^{j-1}]_{(i,j) \in \llbracket 1, k \rrbracket^2}$ est la matrice de Vandermonde de $(\lambda_1, \dots, \lambda_k)$ dont on sait qu'elle est inversible.

Exemple 5 : Le corps de base est \mathbb{R} . On donne $x \in \mathbb{R}^*$ tel que $|x| < 1$. L'intégrale $I_n(x) = \int_0^\pi \frac{\cos nt}{1 - x \cos t} dt$ vérifie $I_0(x) = \frac{\pi}{\sqrt{1-x^2}}$,
 $I_1(x) = \frac{\pi}{x} \left[\frac{1}{\sqrt{1-x^2}} - 1 \right]$, et

$$(16) \quad (\forall n \geq 2) \quad I_n(x) - \frac{2}{x} I_{n-1}(x) + I_{n-2}(x) = 0.$$

En déduire l'expression générale de $I_n(x)$ (Ecrit Mines M, 1986).

Solution : La relation (16) est linéaire à coefficients constants avec $k = 2$. Son polynôme caractéristique est $P(X) = X^2 - \frac{2}{x}X + 1$. Il a deux racines distinctes ρ_1 et ρ_2 :

$$\rho_1 = \frac{1}{x} (1 + \sqrt{1-x^2}), \quad \rho_2 = \frac{1}{x} (1 - \sqrt{1-x^2}).$$

On a donc $(\forall n) \quad I_n(x) = C_1 \rho_1^n + C_2 \rho_2^n$,

où C_1 et C_2 sont déterminés par le système de Cramer :

$$\begin{cases} C_1 + C_2 = I_0(x) \\ C_1 \rho_1 + C_2 \rho_2 = I_1(x). \end{cases}$$

On trouve $C_1 = 0$, $C_2 = \frac{\pi}{\sqrt{1-x^2}}$, d'où

$$(\forall n) \quad I_n(x) = \frac{\pi}{\sqrt{1-x^2}} \frac{1}{x^n} (1 - \sqrt{1-x^2})^n.$$

Exercice 1 : Trouver les $(a, b, c, d, e, f) \in \mathbb{C}^6$ tels que la matrice $M \in \mathfrak{M}_4(\mathbb{C})$ suivante soit diagonalisable :

$$M = \begin{bmatrix} 1 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & -1 & f \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Exercice 2 : Le corps de base est \mathbb{C} . Montrer que les matrices suivantes sont diagonalisables et calculer leurs puissances k -ièmes pour $k \in \mathbb{N}^*$:

$$a) M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad b) M = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$c) M = \begin{bmatrix} d & a & b & c \\ a & d & c & b \\ b & c & d & a \\ c & b & a & d \end{bmatrix}, (a, b, c, d) \in \mathbb{C}^4$$

$$d) M = \begin{bmatrix} 0 & \sin \varphi & \sin 2\varphi \\ \sin \varphi & 0 & \sin 2\varphi \\ \sin 2\varphi & \sin \varphi & 0 \end{bmatrix}, \varphi \in \mathbb{R} \quad e) M = \begin{bmatrix} \lambda & -1 & -\lambda & 1 \\ -1 & 1 & 1 & 1 \\ -\lambda & -1 & \lambda & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}, \lambda \in \mathbb{C}$$

$$f) M = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \quad g) M = \begin{bmatrix} -1 & -\lambda & -2 & -2 \\ -\lambda & -1 & -2 & -2 \\ 2 & 2 & 1 & \lambda \\ 2 & 2 & \lambda & 1 \end{bmatrix}.$$

Exercice 3 : Le corps de base est \mathbb{C} . Pour chacune des matrices suivantes, dire si elle est ou non diagonalisable, et lorsqu'elle est diagonalisable, en donner une base de vecteurs propres :

$$a) M = \begin{bmatrix} 8 & -1 & -5 \\ -2 & 3 & 1 \\ 4 & -1 & -1 \end{bmatrix} \quad b) M = \begin{bmatrix} a & -d & -c & -d \\ b & a & -d & c \\ c & b & a & -b \\ d & -c & b & a \end{bmatrix}$$

$$c) M = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 \\ 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad d) M = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_2 & \boxed{} & & \\ \vdots & & \ddots & \\ a_n & & & 0 \end{bmatrix} \quad \begin{array}{l} \text{avec } n \geq 2, \\ (a_i) \in \mathbb{C}^n \setminus \{0\} \\ \text{(discuter)} \end{array}$$

$$e) M = \begin{bmatrix} c & a & \dots & a & b \\ a & c & & & \vdots \\ \vdots & & \ddots & & a \\ a & \dots & a & c & b \\ b & \dots & \dots & b & c \end{bmatrix} \in \mathfrak{M}_n(\mathbb{C}) \text{ avec } a \neq b \text{ (discuter)}$$

$$f) \begin{bmatrix} 1 & x + \frac{1}{x} & x^2 + \frac{1}{x^2} \\ x + \frac{1}{x} & 1 & x + \frac{1}{x} \\ x^2 + \frac{1}{x^2} & x + \frac{1}{x} & 1 \end{bmatrix} \quad \text{avec } x \in \mathbb{C}^*.$$

Exercice 4 : Soit $n \in \mathbb{N}^*$ ($n \geq 2$) et $(a_0, a_1, \dots, a_{n-1}) \in \mathbb{C}^n$. On considère la matrice circulante $M = \Gamma(a_0, \dots, a_{n-1}) = [a_{j-i}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ (\bar{k} désigne le reste de k mod (n) pour $k \in \mathbb{Z}$). Montrer que M est diagonalisable, et la diagonaliser en calculant $\Omega M \Omega^{-1}$, où $\Omega = [\omega^{(i-1)(j-1)}]$, $((i, j) \in \llbracket 1, n \rrbracket^2)$ est la matrice *alternante* d'ordre n ($\omega = e^{2i\pi/n}$) (cf. § XIII.5, exemple 4).

Exercice 5 : Soit b et c dans \mathbb{C}^* . On considère la matrice $M \in \mathfrak{M}_n(\mathbb{C})$ suivante, où $n \geq 2$:

$$M = \begin{bmatrix} 0 & c & 0 & \dots & 0 \\ b & & & & \vdots \\ 0 & & & & 0 \\ \vdots & & & & c \\ 0 & \dots & 0 & b & 0 \end{bmatrix}.$$

Montrer que M est diagonalisable, et en calculer une base de vecteurs propres (cf. exercice 14 du § XIII.5).

Exercice 6 : Soit $n \in \mathbb{N}$ ($n \geq 2$) et $(a_1, \dots, a_n) \in \mathbb{C}^n$. Étudier si la matrice

$$M = [a_i a_j]_{(i,j) \in \llbracket 1, n \rrbracket^2}$$

est diagonalisable. Lorsqu'elle l'est, en donner une base de vecteurs propres.

Calculer directement M^k pour $k \in \mathbb{N}^*$, en commençant par $k = 2$.

Exercice 7 : Soit $(a_1, \dots, a_n) \in \mathbb{C}^n \setminus \{0\}$ avec $n \geq 1$. Étudier si la matrice $M \in \mathfrak{M}_{n+1}(\mathbb{C})$ suivante est diagonalisable, et quand elle l'est, en donner une base de vecteurs propres :

$$M = \begin{bmatrix} 1 & -a_1 & -a_2 & \dots & -a_n \\ a_1 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \dots & \ddots & 0 \\ a_n & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Exercice 8 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. On suppose u diagonalisable, de polynôme caractéristique $\chi_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ (les λ_i distincts, les $\alpha_i \geq 1$). On pose $\mathcal{C}_u = \{v \in \text{Hom}_K(E) \mid uv = vu\}$.

a) Vérifier que \mathcal{C}_u est une sous- K -algèbre de $\text{Hom}_K(E)$ (appelée le *commutant* de u).

b) Montrer que \mathcal{C}_u est l'ensemble des $\bigoplus_{i=1}^p v_i$ où $v_i \in \text{Hom}_K(E_i)$ pour tout i , en désignant par E_i le sous-espace propre $\text{Ker}(u - \lambda_i \text{Id}_E)$. En déduire : $\dim_K(\mathcal{C}_u) = \sum_{i=1}^p \alpha_i^2$.

c) On suppose maintenant tous les α_i égaux à 1 (donc $p = n$). Montrer que :

$$\mathcal{C}_u = \{P(u)\}_{P \in K_{n-1}[X]}.$$

Indication : Penser à l'interpolation de Lagrange.

d) Pour chacune des matrices suivantes de $\mathfrak{M}_3(K)$, trouver avec $K = \mathbb{C}$, puis avec $K = \mathbb{R}$, l'ensemble $\mathcal{C}_M = \{A \in \mathfrak{M}_3(K) \mid AM = MA\}$:

$$M = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}; \quad \bar{M} = \begin{bmatrix} 8 & -1 & -5 \\ -2 & 3 & 1 \\ 4 & -1 & -1 \end{bmatrix}.$$

Exercice 9 : On reprend les notations de l'exercice 1 du § XV.2.

a) Montrer que si α et β sont diagonalisables, alors S et T le sont, et en donner une base de vecteurs propres connaissant des bases de vecteurs propres pour α et β .

b) On suppose maintenant $\alpha = \beta$ (et α diagonalisable). Quel est l'ordre de multiplicité de la valeur propre 0 de S ? Interpréter à l'aide du *commutant* de α défini dans l'exercice précédent.

Exercice 10 : Soit E le \mathbb{R} -cv $\mathbb{R}_n[X]$. Trouver les valeurs propres et les vecteurs propres de l'endomorphisme $u \in \text{Hom}_{\mathbb{R}}(E)$ dans chacun des cas suivants, et préciser si u est diagonalisable :

a) $u : P(X) \mapsto \frac{d}{dX}[(\alpha X + \beta)P(X)]$, $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{0\}$ donné.

b) $u : P(X) \mapsto (X - a)(X - b)P'(X) - nXP(X)$, $(a, b) \in \mathbb{R}^2$ donné.

c) $u : P(X) \mapsto (X - \alpha)[P'(X) + P'(\alpha)] - 2[P(x) - P(\alpha)]$, $\alpha \in \mathbb{R}$ donné.

Exercice 11 : On suppose $K = \mathbb{C}$ et on donne $n \in \mathbb{N}^*$. On considère le *discriminant général* d'ordre n défini au § X.6 : $D(X_1, \dots, X_n) = \prod_{1 \leq i < j \leq n} (X_i - X_j)^2$, où (X_1, \dots, X_n) est un système d'indéterminées sur \mathbb{C} .

On notera Φ le polynôme à n lettres à coefficients dans \mathbb{C} tel que

$$D(X_1, \dots, X_n) = \Phi(\sigma_1, \sigma_2, \dots, \sigma_n),$$

où

$$\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} X_{i_1} \dots X_{i_k} \quad (\text{cf. § X.4}).$$

Pour toute matrice $M \in \mathfrak{M}_n(\mathbb{C})$, on note son polynôme caractéristique

$$\tilde{X}_M(X) = (-1)^n (X^n - \tau_1(M) X^{n-1} + \tau_2(M) X^{n-2} + \cdots + (-1)^n \tau_n(M)).$$

a) Montrer que $f(M) = \Phi(\tau_1(M), \dots, \tau_n(M))$ est une fonction polynomiale non nulle sur $\mathfrak{M}_n(\mathbb{C})$, et que $\chi_M(X)$ a toutes ses racines distinctes ssi $f(M) \neq 0$.

b) Soit S l'ensemble des matrices diagonalisables dans $\mathfrak{M}_n(\mathbb{C})$. Montrer que si une fonction polynomiale $g : \mathfrak{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$ est nulle sur S , alors $g = 0$ (cf. théorème X.1.4).

Exercice 12 : a) Soit $n \in \mathbb{N}^*$ et $M \in \mathfrak{M}_n(\mathbb{C})$ une matrice diagonalisable. Démontrer que $\chi_M(M) = 0$.

b) En appliquant le résultat de l'exercice 11b ci-dessus, en déduire que $\chi_M(M) = 0$ pour toute matrice $M \in \mathfrak{M}_n(\mathbb{C})$.

Exercice 13 : On reprend les notations de l'exercice 11 ci-dessus. Une fonction polynomiale $I : \mathfrak{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$ est dite *invariante* ssi :

$$(\forall P \in \text{GL}(n, \mathbb{C}), \forall M \in \mathfrak{M}_n(\mathbb{C})) \quad I(PMP^{-1}) = I(M).$$

a) Vérifier, en revenant à l'étude du § XV.1, que chacune des fonctions $\tau_k : \mathfrak{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$ est invariante ($1 \leq k \leq n$).

b) Les fonctions polynomiales invariantes sur $\mathfrak{M}_n(\mathbb{C})$ forment une sous- \mathbb{C} -algèbre \mathcal{I} de l'algèbre des fonctions polynomiales sur $\mathfrak{M}_n(\mathbb{C})$.

c) La suite $(\tau_1, \tau_2, \dots, \tau_n)$ est algébriquement libre sur \mathbb{C} .

Indication : Réfléchir sur l'exercice 1b du § XV.1.

d) Soit $I : \mathfrak{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$ une fonction polynomiale invariante. Pour $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{C}^n$, on pose $\Psi(\lambda_1, \dots, \lambda_n) = I(\text{Diag}(\lambda_1, \dots, \lambda_n))$. Montrer que Ψ est polynomiale symétrique sur \mathbb{C}^n ; en déduire qu'on a un polynôme unique Θ à n lettres sur \mathbb{C} tel que $\Psi(\lambda_1, \dots, \lambda_n) = \Theta(s_1, \dots, s_n)$, en notant

$$s_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} \lambda_{i_1} \dots \lambda_{i_k} \quad \text{pour } 1 \leq k \leq n.$$

On définit alors $J : \mathfrak{M}_n(\mathbb{C}) \rightarrow \mathbb{C}$, $M \mapsto \Theta(\tau_1(M), \dots, \tau_n(M))$. En utilisant le résultat de l'exercice 11b ci-dessus, montrer que $I = J$. Conclure : quelle est l'algèbre \mathcal{I} ?

Exercice 14 : Soit $u \in \text{Hom}_K(E)$ et $v \in \text{Hom}_K(E)$. On suppose que u et v sont diagonalisables et commutent. Montrer que $v \circ u$ est diagonalisable.

Exercice 15 : Déterminer les suites $(u_n)_{n \geq 0}$ à valeurs dans \mathbb{C} vérifiant la relation de récurrence linéaire $u_{n+2} = u_{n+1} + u_n$.

Application : Calculer le nombre de n -uplets formés de 0 et de 1 mais ne renfermant pas deux 1 consécutifs.

Exercice 16 : Trois suites réelles vérifient les relations de récurrence :

$$u_{n+1} = \frac{v_n + 3w_n}{8}, \quad v_{n+1} = \frac{u_n}{2}, \quad w_{n+1} = \frac{v_n}{2}.$$

On donne u_0, v_0, w_0 . Calculer u_n, v_n, w_n et chercher leur limite quand $n \rightarrow +\infty$.

Exercice 17 : a) Déterminer les suites réelles (u_n) solutions de l'équation de récurrence linéaire : $u_{n+3} = 6u_{n+2} - 11u_{n+1} + 6u_n$.

b) Même question avec $u_{n+3} = \frac{-1}{6}u_{n+2} + \frac{2}{3}u_{n+1} - \frac{1}{6}u_n$.

§ XV.4 POLYNÔMES D'ENDOMORPHISMES OU DE MATRICES

Rappelons quelques faits élémentaires concernant les K -algèbres. Soit \mathcal{A} une K -algèbre, d'élément unité e , et $a \in \mathcal{A}$. Le théorème VII.4.1 nous enseigne qu'on a un, et un seul, homomorphisme de K -algèbres $\varphi_a : K[X] \rightarrow \mathcal{A}$ tel que $\varphi_a(X) = a$. A chaque

$$P = \lambda_0 + \lambda_1 X + \cdots + \lambda_k X^k \in K[X],$$

φ_a associe l'élément $\lambda_0 a^0 + \lambda_1 a + \cdots + \lambda_k a^k$ (où, par convention, $a^0 = e$). L'image de φ_a est une sous- K -algèbre de \mathcal{A} , que l'on notera dans la suite $K[a]$; il est clair que $K[a]$ est une algèbre *commutative*, puisque $K[X]$ l'est.

Appliquons ces considérations au cas où \mathcal{A} est l'algèbre $\mathfrak{M}_n(K)$, où $n \in \mathbb{N}^*$. Si $M \in \mathfrak{M}_n(K)$, l'algèbre $K[M]$ est appelée **algèbre des polynômes de la matrice M** . On peut aussi prendre $\mathcal{A} = \text{Hom}_K(E)$, où E est un K -ev. Si $u \in \text{Hom}_K(E)$, la K -algèbre $K[u]$ sera appelée **algèbre des polynômes de l'endomorphisme u** .

Polynôme minimal

Considérons un K -ev E de dimension finie $n \geq 1$, et fixons $u \in \text{Hom}_K(E)$. Nous savons que le K -ev $\text{Hom}_K(E)$ est de dimension finie, égale à n^2 , alors que le K -ev $K[X]$ est de dimension infinie. Par suite, l'homomorphisme de K -algèbres $\varphi_u : K[X] \rightarrow \text{Hom}_K(E)$, $P \mapsto P(u)$, qui est K -linéaire, ne peut pas être injectif. Donc son noyau est un idéal non nul de $K[X]$. Le théorème VII.2.2 nous montre alors qu'il existe un et un seul polynôme *normalisé* $G \in K[X]$ tel que $\text{Ker}(\varphi_u) = GK[X]$.

Ce polynôme G est nécessairement non constant. Il vérifie $G(u) = 0$ (on dit qu'il est *annulé* par u), et l'ensemble des polynômes annulés par u est $GK[X]$. Le degré de G est tel que $1 \leq d \leq n^2$, puisque la famille $(u^k)_{0 \leq k \leq n^2}$ est forcément liée dans le K -ev $\text{Hom}_K(E)$.

Bien sûr on obtient des résultats analogues en remplaçant $\text{Hom}_K(E)$ par $\mathfrak{M}_n(K)$ et u par une matrice donnée $M \in \mathfrak{M}_n(K)$. Il existe donc un unique polynôme normalisé H , nécessairement non constant, tel que $HK[X]$ soit l'ensemble des polynômes annulés par M .

DÉFINITION XV.4.1

Soit $n \in \mathbb{N}^*$ et E un K -ev de dimension n . Si $u \in \text{Hom}_K(E)$, on appelle **polynôme minimal** de u l'unique polynôme *normalisé* $G \in K[X]$ tel que

$$GK[X] = \{P \in K[X] \mid P(u) = 0\}.$$

Si $M \in \mathfrak{M}_n(K)$, on appelle **polynôme minimal** de M l'unique polynôme **normalisé** $H \in K[X]$ tel que

$$HK[X] = \{P \in K[X] \mid P(M) = 0\}.$$

Le polynôme minimal de u (resp. de M) sera noté ci-dessous $Q_u(X)$, (resp. $Q_M(X)$). On a : $Q_u(X) = X$ (resp. $Q_M(X) = X$) ssi $u = 0$ (resp. $M = 0$). La définition de $Q_u(X)$ est conforme à la définition VII.4.2 quand on prend pour algèbre $\text{Hom}_K(E)$. Fixons une base ordonnée $\mathcal{B} = (e_1, \dots, e_n)$ de E et $u \in \text{Hom}_K(E)$. Soit $M = \text{Mat}_{\mathcal{B}}(u)$. Pour tout polynôme $P \in K[X]$, on a : $\text{Mat}_{\mathcal{B}}(P(u)) = P(M)$, d'où il résulte aussitôt : $P(u) = 0 \Leftrightarrow P(M) = 0$, et en particulier :

$$(1) \quad Q_u(X) = Q_M(X).$$

Si par exemple on part de $N \in \mathfrak{M}_n(K)$ et qu'on note u_N l'endomorphisme de K^n associé à N dans la base canonique, on a :

$$(2) \quad Q_N(X) = Q_{u_N}(X).$$

Exemple 1 : Supposons $n \geq 2$, le corps K n'étant pas de caractéristique 2 ; soit F et G deux sous- K -ev supplémentaires et non nuls de E , et $u = \text{Id}_F \oplus (-\text{Id}_G)$; u est une *involution* de E (cf. l'exemple 2 du § IX.1). Donc le polynôme $X^2 - 1$ est annulé par u .

Mais $u \notin \{\text{Id}_E, -\text{Id}_E\}$ puisque F et G sont supposés non nuls, et que $1_K \neq -1_K$. Donc $Q_u(X) = X^2 - 1$.

Exemple 2 : Supposons $n \geq 2$ et soit u un projecteur de E , de rang $r \in \llbracket 1, n-1 \rrbracket$. Alors $X^2 - X$ est annulé par u (cf. § IX.1), mais $u \neq 0$ et $u \neq \text{Id}_E$, donc : $Q_u(X) = X^2 - X$.

THÉORÈME XV.4.1

|| Soit u un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$. Toute valeur propre de u est racine du polynôme minimal $Q_u(X)$.

Démonstration :

Posons

$$Q_u(X) = X^q + b_1 X^{q-1} + \dots + b_q \quad (q \geq 1),$$

et soit λ une valeur propre de u et $x \in E$ un vecteur propre associé. Alors $u^k(x) = \lambda^k(x)$ pour tout $k \in \mathbb{N}$; d'où, puisque $Q_u(u) = 0$:

$$0 = [Q_u(u)](x) = (\lambda^q + b_1 \lambda^{q-1} + \dots + b_q) x = Q_u(\lambda) x.$$

Comme $x \neq 0$ il s'ensuit bien que $Q_u(\lambda) = 0$. ■

Endomorphismes (ou matrices) nilpotent(e)s

Nous avons déjà rencontré plusieurs fois la notion d'élément nilpotent. Précisons cette notion.

DÉFINITION XV.4.2

Un élément a d'une K -algèbre \mathcal{A} est dit **nilpotent** ssi il existe $k \in \mathbb{N}^*$ tel que $a^k = 0$. S'il en est ainsi, le plus petit de ces entiers k est appelé la **période** de a .

La période vaut 1 ssi $a = 0$.

Exemple 3 : Dans $K_n[X]$ ($n \in \mathbb{N}^*$), soit D l'endomorphisme $P(X) \mapsto P'(X)$; D est élément nilpotent de $\text{Hom}_K(K_n[X])$. Si K est de caractéristique 0, la période de D est $n + 1$.

Exemple 4 : Soit $n \in \mathbb{N}^*$ et $M \in \mathfrak{M}_n(K)$ une matrice trigonale (par exemple trigonale supérieure) $M = [a_{ij}]$. Pour tout $k \in \mathbb{N}^*$, posons $M^k = [a_{ij}^{(k)}]$. Alors $(\forall i) a_{i,i}^{(k)} = (a_{i,i})^k$. Si tous les $a_{i,i}$ sont nuls, un calcul immédiat montre que $a_{i,j}^{(k)} = 0$ pour $i > j - k$. Donc M est nilpotente ssi toutes ses valeurs propres $a_{i,i}$ sont nulles, et si c'est le cas, sa période est $\leq n$.

Bornons-nous pour l'instant à quelques remarques simples sur les endomorphismes nilpotents d'un K -ev $E \neq \{0\}$. D'abord, si $u \in \text{Hom}_K(E)$ est nilpotent, son noyau est $\neq \{0\}$, i.e. 0 est valeur propre de u . De plus, 0 est la seule valeur propre possible ($u^k = 0$, $x \in E \setminus \{0\}$, $\lambda \in K$ et $u(x) = \lambda x$ entraînent $u^k(x) = \lambda^k x = 0$, donc $\lambda = 0$). Si de plus E est de dimension finie, l'entier $r \geq 1$ étant donné, u est nilpotent de période r ssi son polynôme minimal est : X^r . Enfin, on a deux résultats un peu moins évidents :

PROPOSITION XV.4.1

Si E est un K -ev de dimension finie $n \geq 1$ et si $u \in \text{Hom}_K(E)$ est nilpotent, on a : $u^n = 0$, autrement dit : la période r est $\leq n$.

Démonstration :

Prenons $x \in E \setminus \{0\}$ tel que $u^{r-1}(x) \neq 0$. Il suffit de montrer que les r vecteurs $x, u(x), \dots, u^{r-1}(x)$ sont linéairement indépendants, car cela entraînera $r \leq n$. Or, soit $(\lambda_0, \lambda_1, \dots, \lambda_{r-1}) \in K^n$ tel que $\lambda_0 x + \dots + \lambda_{r-1} u^{r-1}(x) = 0$. On applique successivement au premier membre de cette relation : u^{r-1} , puis u^{r-2} , ..., jusqu'à Id_E et cela donne $\lambda_0 = 0$, puis $\lambda_1 = 0$, ..., $\lambda_{r-1} = 0$ car $u^{r-1}(x) \neq 0$. ■

PROPOSITION XV.4.2

Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. Pour que u soit nilpotent, il faut et il suffit que son polynôme caractéristique soit : $(-1)^n X^n$.

Démonstration :

Si $\chi_u(X) = (-1)^n X^n$, u est trigonalisable, donc nilpotent à cause de l'exemple 4 ci-dessus.

Montrons la réciproque par récurrence sur n : la propriété est évidente pour $n = 1$. Supposons-la vraie à l'ordre $n - 1$, avec $n \geq 2$. Alors ${}^t u \in \text{Hom}_K(E^*)$ est aussi nilpotent. Soit $\varphi \in \text{Ker}({}^t u) \setminus \{0\}$ et H l'hyperplan $\text{Ker}(\varphi)$. Il est u -stable, car si $x \in E$, $\varphi(x) = 0$ entraîne $\varphi(u(x)) = ({}^t u(\varphi))(x) = 0$. Posons $v = u|_H$:

v est nilpotent, donc $\chi_v(X) = (-1)^{n-1} X^{n-1}$ d'après l'hypothèse de récurrence. Mais $\chi_v(X)$ divise $\chi_u(X)$ (théorème XV.1.3) ; le quotient $\frac{\chi_u(X)}{\chi_v(X)}$ étant de degré 1, et 0 étant la seule valeur propre possible de u , il s'ensuit : $\chi_u(X) = (-1)^n X^n$. ■

En passant aux matrices, on a :

COROLLAIRE

|| Soit $n \in \mathbb{N}^*$. Toute matrice $M \in \mathfrak{M}_n(K)$ nilpotente est semblable dans $\mathfrak{M}_n(K)$ à une matrice trigonale.

Le théorème de Hamilton-Cayley

Nous venons de voir, en conséquence des propositions XV.4.1 et XV.4.2, que tout endomorphisme nilpotent u d'un K -ev de dimension finie annule son polynôme caractéristique. L'exercice 12 du § XV.3 demandait également de prouver que tout endomorphisme diagonalisable de $\text{Hom}_{\mathbb{C}}(E)$ annule son polynôme caractéristique. Ce sont des cas particuliers d'une propriété générale, que nous allons établir par des méthodes purement matricielles. Nous fixerons $n \in \mathbb{N}^*$ et considérerons $\mathfrak{M}_n(K)$ comme sous- K -algèbre de $\mathfrak{M}_n(K[X])$, elle-même plongée dans $\mathfrak{M}_n(L)$, où L est le corps $K(X)$.

LEMME 1

|| Soit $C_0, C_1, \dots, C_m \in \mathfrak{M}_n(K)$ telles que

$$(3) \quad C_0 + C_1(XI_n) + \dots + C_m(X^m I_n) = 0.$$

|| Alors $C_\alpha = 0$ pour $0 \leq \alpha \leq m$.

Démonstration :

Posons $C_\alpha = [c_{i,j,\alpha}]_{(i,j) \in \llbracket 1, n \rrbracket^2}$ pour $0 \leq \alpha \leq m$. La relation (3) entraîne, pour chaque couple (i, j) :

$$\sum_{\alpha=0}^m c_{i,j,\alpha} X^\alpha = 0, \text{ d'où}$$

$(\forall \alpha) c_{i,j,\alpha} = 0$; et c'est vrai pour tout (i, j) .

Donc $(\forall \alpha) C_\alpha = 0$. ■

THÉORÈME XV.4.2 (Hamilton-Cayley)

|| Pour toute matrice $M \in \mathfrak{M}_n(K)$, on a : $\chi_M(M) = 0$. Autrement dit : le polynôme caractéristique de M est annulé par M .

Démonstration :

Soit N la matrice $M - XI_n \in \mathfrak{M}_n(L)$. Sa matrice complémentaire \tilde{N} vérifie

$$(4) \quad N\tilde{N} = \tilde{N}N = \det(N) I_n = \chi_M(X) I_n$$

(cf. théorème XIII.4.5). L'examen des $(n-1)$ -mineurs de N montre que chaque coefficient de \tilde{N} appartient à $K_{n-1}[X]$. Par conséquent, \tilde{N} s'écrit (de manière unique) :

$$\tilde{N} = A_0 + A_1(XI_n) + \cdots + A_{n-1}(X^{n-1}I_n),$$

où pour tout i $A_i \in \mathfrak{M}_n(K)$.

Développons $\chi_M(X)$ sous la forme : $\chi_M(X) = a_0 + a_1 X + \cdots + a_n X^n$. Alors :

$$(5) \quad \chi_M(X) I_n = a_0 I_n + (a_1 I_n) XI_n + \cdots + (a_n I_n) X^n I_n.$$

En développant $N\tilde{N}$ et $\tilde{N}N$ (l'ordre des termes étant respecté) par distributivité, et en ordonnant suivant les puissances croissantes de X , on obtient :

$$(6) \quad \tilde{N}N = A_0 M + \sum_{k=1}^{n-1} (A_k M - A_{k-1})(X^k I_n) + (-A_{n-1})(X^n I_n),$$

et

$$(7) \quad N\tilde{N} = MA_0 + \sum_{k=1}^{n-1} (MA_k - A_{k-1})(X^k I_n) + (-A_{n-1})(X^n I_n).$$

Par application du lemme 1, nous pouvons identifier les coefficients des $(X^k I_n)_{0 \leq k \leq n}$ aux seconds membres de (5), (6) et (7), d'où :

$$(8) \quad \begin{cases} A_0 M = MA_0 = a_0 I_n ; \\ A_k M - A_{k-1} = MA_k - A_{k-1} = a_k I_n \text{ pour } 1 \leq k \leq n-1 ; \\ \text{et } A_{n-1} = -a_n I_n \\ \text{(en particulier, } MA_k = A_k M \text{ pour } 0 \leq k \leq n-1). \end{cases}$$

On a : $\chi_M(M) = a_0 I_n + \sum_{k=1}^n (a_k I_n) M^k$. En remplaçant, dans cette expression, les $a_k I_n$ par leurs valeurs tirées de (8), on constate que les termes se réduisent deux à deux, et qu'il reste $\chi_M(M) = 0$. ■

Le lecteur aura remarqué que cette démonstration n'utilise que la structure d'anneau commutatif de K , l'intégrité n'intervenant pas. On peut donc énoncer et appliquer la relation $\chi_M(M) = 0$ avec une matrice M à coefficients dans un anneau commutatif arbitraire.

Une conséquence du théorème XV.4.2 est que, pour $M \in \mathfrak{M}_n(K)$, le **polynôme minimal** $Q_M(X)$ **divise le polynôme caractéristique** $\chi_M(X)$, et en particulier, $\deg(Q_M(X)) \in \llbracket 1, n \rrbracket$. En tenant compte de (1), on en déduit :

COROLLAIRE

|| Soit E un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$. Alors $\chi_u(u) = 0$; le **polynôme minimal** $Q_u(X)$ est donc un **diviseur du polynôme caractéristique** $\chi_u(X)$.

On peut avoir : $\deg(Q_u(X)) < n = \deg(\chi_u(X))$, comme le montrent les exemples 1 et 2 ci-dessus lorsque $n \geq 3$.

Remarque 1 : Si l'on sait à coup sûr, au moins théoriquement, expliciter le polynôme caractéristique de $M \in \mathfrak{M}_n(K)$ par un calcul de déterminants (cf. § XV.1), il n'est pas si facile de trouver son polynôme minimal $Q_M(X)$ parmi les diviseurs de $\chi_M(X)$. Un calcul théorique de $Q_M(X)$ est proposé dans l'exercice 6.

Exercice 1 : Prouver directement, par récurrence, le corollaire de la proposition XV.4.2 sur les matrices nilpotentes.

Indication : En prenant comme premier vecteur de base un vecteur de $\text{Ker}(u_M)$, $\text{Mat}_{\mathcal{B}}(u_M)$ prend la forme $\begin{bmatrix} 0 & \times & \times & \times & \times \\ 0 & & & & \\ \vdots & & M_1 & & \\ 0 & & & & \end{bmatrix}$, où M_1 est nilpotente, ce qui permet d'appliquer l'hypothèse de récurrence.

Exercice 2 : Soit $n \in \mathbb{N}^*$ et $(a_0, a_1, \dots, a_{n-1}) \in K^n$. Montrer que pour la matrice $M \in \mathfrak{M}_n(K)$ égale à $\begin{bmatrix} 0 & \dots & \dots & 0 & a_0 \\ 1 & & & \vdots & a_1 \\ 0 & 1 & & 0 & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & a_{n-1} \end{bmatrix}$, on a : $(-1)^n \chi_M(X) = Q_M(X)$.

Exercice 3 : Soit $n \in \mathbb{N}^*$. On note \mathcal{E} l'ensemble des $M \in \mathfrak{M}_n(K)$ telles que $Q_M(X) = (-1)^n \chi_M(X)$ (i.e. $\deg(Q_M(X)) = n$).

a) Montrer qu'on peut trouver des fonctions polynomiales f_1, \dots, f_k sur $\mathfrak{M}_n(K)$ (avec k convenable) telles que : $M \notin \mathcal{E} \Leftrightarrow (\forall i \in \llbracket 1, k \rrbracket f_i(M) = 0)$.

Indication : Interpréter la condition $M \in \mathcal{E}$ à l'aide du rang de la suite de vecteurs (I_n, M, \dots, M^{n-1}) dans $\mathfrak{M}_n(K)$.

b) Si $K = \mathbb{R}$ ou \mathbb{C} , en déduire que \mathcal{E} est un ouvert dense de $\mathfrak{M}_n(K)$.

Indication : S'aider de l'exercice 2 ci-dessus et de l'exercice 1 du § X.1.

Exercice 4 : Soit $n \in \mathbb{N}^*$. On donne A et B éléments de $\mathfrak{M}_n(\mathbb{C})$. Montrer l'équivalence entre les 3 conditions suivantes :

- (I) $\forall C \in \mathfrak{M}_n(\mathbb{C})$, l'équation $AX - XB = C$, $X \in \mathfrak{M}_n(\mathbb{C})$ a une solution unique.
 (II) $(X \in \mathfrak{M}_n(\mathbb{C}) \text{ et } AX - XB = 0) \Rightarrow X = 0$.
 (III) A et B n'ont pas de valeur propre commune (cf. exercice 7 du § XV.2).

Exercice 5 : Utiliser le théorème de Cayley-Hamilton pour calculer l'inverse des matrices suivantes :

$$A = \begin{bmatrix} 5 & -1 & 9 \\ 3 & 4 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 7 & -6 \\ -1 & 4 & -1 \\ 1 & 2 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & -2 & -1 \\ 2 & -1 & 1 \\ 6 & 3 & -2 \end{bmatrix}.$$

Exercice 6 : Soit $n \in \mathbb{N}$, ($n \geq 2$) et $M \in \mathfrak{M}_n(K)$. On pose $N = M - XI_n$ (X est une indéterminée sur K) et $\tilde{N} = [S_{ij}(X)]_{(i,j) \in \llbracket 1, n \rrbracket^2}$. (On sait que $\forall (i, j) S_{ij}(X) \in K_{n-1}[X]$). Soit Δ le pgcd des $(S_{i,j}(X))$ dans $K[X]$, dont on vérifiera qu'il est $\neq 0$. Enfin soit $T = \frac{1}{\Delta} \tilde{N}$.

- a) En utilisant $\tilde{N}N = \chi_M(X) I_n$, prouver que Δ divise $\chi_M(X)$.
 b) Si $G(X) = \frac{1}{\Delta} \chi_M(X)$, montrer que $TN = G(X) I_n$ et en déduire que Q_M divise G .
 c) Soit $D(X) = \frac{\chi_M(X)}{Q_M(X)}$. Montrer qu'on a $B \in \mathfrak{M}_n(K[X])$ telle que $BN = NB = Q_M(X) I_n$ (utiliser : $Q_M(M) = 0$ et le fait que M et XI_n sont permutables).
 d) Démontrer $DBN = \tilde{N}N$. En déduire que $DB = \tilde{N}$, puis, que D divise Δ .
 En déduire enfin l'expression du polynôme minimal : $Q_M(X) = \frac{\chi_M(X)}{\Delta}$.

Exercice 7 : Le corps K est supposé de caractéristique 0. Soit $n \in \mathbb{N}^*$. On donne $A \in \mathfrak{M}_n(K)$ et on pose $B = XI_n - A$, X désignant une indéterminée sur K , $P_A(X) = \det(B)$; \tilde{B} désigne la matrice des cofacteurs de B .

- a) Démontrer : $\text{Tr}(\tilde{B}) = \frac{dP_A(X)}{dX}$.
 b) Montrer qu'on a $(\alpha_1, \dots, \alpha_n) \in K^n$ et des matrices C_0, C_1, \dots, C_{n-1} dans $\mathfrak{M}_n(K)$ telles que $P_A(X) = X^n + \alpha_1 X^{n-1} + \dots + \alpha_n$ et $\tilde{B} = C_0 X^{n-1} + C_1 X^{n-2} + \dots + C_{n-1}$.
 Vérifier $C_0 = I_n$, $-C_{n-1}A = \alpha_n I_n$, et $C_k - C_{k-1}A = C_k - AC_{k-1} = \alpha_k I_n$ pour $1 \leq k \leq n-1$ (cf. la preuve du théorème XV.4.2). En déduire

$$C_k = A^k + \sum_{j=0}^{k-1} \alpha_{k-j} A^j \quad \text{pour } 1 \leq k \leq n-1.$$

- c) Montrer : $\text{Tr}(C_0) = n$, $\text{Tr}(C_k) = (n-k) \alpha_k$ pour $1 \leq k \leq n-1$.
 A l'aide de b) en déduire un système de Cramer qui fournit $\text{Tr}(A)$, $\text{Tr}(A^2)$, ..., $\text{Tr}(A^n)$ en fonction de $\alpha_1, \alpha_2, \dots, \alpha_n$. En déduire que si $\text{Tr}(A^k) = 0$ pour $1 \leq k \leq n$, alors $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.
 d) Déduire de l'étude précédente que si l'on pose $A_0 = A$ et, pour $k \geq 1$, $A_k = A \left(A_{k-1} - \frac{1}{k} \text{Tr}(A_{k-1}) I_n \right)$, alors $A_n = 0$ et
 (*) $\chi_A(X) = (-1)^n \left[X^n - \text{Tr}(A_0) X^{n-1} - \frac{1}{2} \text{Tr}(A_1) X^{n-2} - \dots - \frac{1}{n} \text{Tr}(A_{n-1}) \right],$

l'équation $A_n = 0$ étant équivalente au théorème XV.4.2.

Commentaire : L'expression (*) donne une méthode itérative pour calculer $\chi_A(X)$ (méthode de Faddeev) qui se révèle bien plus rapide, dans les cas numériques, que le calcul direct de $\chi_A(X)$ par déterminants.

Exercice 8 : On donne $n \in \mathbb{N}^*$, $m \in \mathbb{N}^*$. Soit A_0, A_1, \dots, A_m dans $\mathfrak{M}_n(\mathbb{C})$, et $f(X) = \det(A_0 X^m + A_1 X^{m-1} + \dots + A_m)$ (X : indéterminée sur \mathbb{C}).

Montrer : Si $M \in \mathfrak{M}_n(\mathbb{C})$ vérifie $A_0 M^m + A_1 M^{m-1} + \dots + A_m = 0$, alors $f(M) = 0$ (généralisation du théorème XV.4.1).

Exercice 9 : Le corps de base est \mathbb{C} . Pour $n \in \mathbb{N}^*$ on pose $\zeta_n = e^{2i\pi/n}$ et on se propose de calculer la somme de Gauss $G_n = \sum_{k=0}^{n-1} \zeta_n^{k^2}$.

a) Calculer G_1, G_2, G_3, G_4 et G_5 . Calculer $\sum_{k=0}^{n-1} \zeta_n^{kr}$ pour $r \in \mathbb{Z}$.

b) On considère la matrice $A_n = [a_{r,s}]_{(r,s) \in \llbracket 1, n \rrbracket^2}$ telle que $a_{r,s} = \zeta_n^{(r-1)(s-1)}$. Montrer que $G_n = \text{Tr}(A_n)$. Calculer A_n^2 et A_n^4 . En déduire que A_n a, au plus, 4 valeurs propres.

c) Montrer que toute matrice $U \in \mathfrak{M}_n(\mathbb{C})$ telle que $U^2 = I_n$ est diagonalisable. Pour n impair ($n = 2p + 1$) expliciter une base dans laquelle A_n^2 est diagonale.

d) Montrer que $G_n \overline{G_n} = \sum_{0 \leq r \leq n-1} \left(\sum_{0 \leq s \leq n-1} \zeta_n^{(r+s)^2 - r^2} \right)$ et en déduire que, pour $n = 2p + 1$, $|G_n| = \sqrt{n}$.

e) Montrer que A_n est semblable à une matrice diagonale $\text{Diag}(\lambda_1, \dots, \lambda_n)$, où les coefficients (λ_i) prennent leurs valeurs dans $\{\sqrt{n}, -\sqrt{n}, i\sqrt{n}, -i\sqrt{n}\}$.

Soit a, b, c, d le nombre de fois que ces valeurs respectives sont prises. Montrer que si $n = 2p + 1$ ($p \in \mathbb{N}$), on a : $a + b = p + 1$, $c + d = p$, $(a - b)^2 + (c - d)^2 = 1$. Calculer $\det(A_n) = \prod_{k=1}^n \lambda_k$ et en déduire que $(b + d)$ a la même parité que p . Acheter la détermination de G_n lorsque n est impair.

Exercice 10 : Expliquer clairement pourquoi toute matrice nilpotente de $\mathfrak{M}_3(K)$ est semblable à l'une (et bien sûr une seule) des 3 matrices suivantes :

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

§ XV.5 SOUS-ESPACES CARACTÉRISTIQUES

Donnons-nous un K -ev E et $u \in \text{Hom}_K(E)$. Nous allons étudier de manière plus approfondie l'algèbre $K[u]$ des *polynômes de u* . Une propriété essentielle de cette algèbre est sa *commutativité*. Pour faciliter la lecture, si $P \in K[X]$ et $x \in E$, à la place de $[P(u)](x)$ nous écrirons $P(u) \cdot x$.

PROPOSITION XV.5.1

|| Avec les notations ci-dessus, pour tout $P \in K[X]$, $\text{Im}(P(u))$ et $\text{Ker}(P(u))$ sont u -stables.

Démonstration :

Pour $x \in E$, si $P(u) \cdot x = 0$, alors

$$P(u) \cdot u(x) = (P(u)u) \cdot x = (uP(u)) \cdot x = u[P(u) \cdot x] = u(0) = 0,$$

donc $\text{Ker}(P(u))$ est u -stable.

Si $x = P(u) \cdot y$, alors

$$u(x) = (uP(u)) \cdot y = (P(u)u) \cdot y = P(u) \cdot u(y) \in \text{Im}(P(u)).$$

Donc $\text{Im}(P(u))$ est u -stable. ■

THÉORÈME XV.5.1 (« lemme des noyaux »)

On donne un K -ev E , des polynômes $S_1, S_2, \dots, S_p \in K[X]$ **deux à deux premiers entre eux** ($p \geq 2$) et $u \in \text{Hom}_K(E)$. On pose $S = \prod_{i=1}^p S_i$, $\sigma_i = S_i(u)$, $\sigma = S(u)$, $N = \text{Ker}(\sigma)$, $N_i = \text{Ker}(\sigma_i)$ ($1 \leq i \leq p$). Alors :

(I) On a : $N = \bigoplus_{i=1}^p N_i$.

(II) Si $S(u) = 0$ (i.e. $N = E$), les projecteurs g_i , sur N_i parallèlement à $\bigoplus_{j \neq i} N_j$, appartiennent, pour chaque i , à $K[u]$.

Démonstration :

Pour chaque i , soit $T_i = \frac{S}{S_i}$ et $\tau_i = T_i(u)$. L'hypothèse entraîne que les $(T_i)_{1 \leq i \leq p}$ sont premiers entre eux dans leur ensemble. Choisissons donc des polynômes $A_1, \dots, A_p \in K[X]$ tels que $\sum_{i=1}^p A_i T_i = 1$, et posons $\alpha_i = A_i(u)$ pour $1 \leq i \leq p$, d'où :

$$(1) \quad \sum_{i=1}^p \alpha_i \tau_i = \text{Id}_E.$$

(I) Montrons d'abord que $N = \sum_{i=1}^p N_i$. Pour chaque i , l'inclusion $N_i \subset N$ est évidente, car $\sigma = \tau_i \sigma_i$. D'autre part, soit $x \in N$. De (1) on déduit : $x = \sum_{i=1}^p \alpha_i \tau_i \cdot x$, mais pour tout i ,

$$\sigma_i(\alpha_i \tau_i \cdot x) = (\alpha_i(\tau_i \sigma_i)) \cdot x = \alpha_i \cdot (\sigma \cdot x) = 0,$$

donc $\alpha_i \tau_i \cdot x \in N_i$, donc $x \in \sum_{i=1}^p N_i$, et finalement $N = \sum_{i=1}^p N_i$. Il reste à prouver que les (N_i) sont indépendants. Pour cela donnons-nous $x_1 \in N_1, \dots, x_p \in N_p$ tels que $x_1 + x_2 + \dots + x_p = 0$. Pour i et j distincts dans $\llbracket 1, p \rrbracket$, il est sûr que $\tau_i \cdot x_j = 0$ car $\tau_i = \beta_{ij} \sigma_j$ ($\beta_{ij} = \left(\prod_{k \neq \{i\}} S_k \right)(u)$).

Fixons $i \in \llbracket 1, p \rrbracket$. On a :

$$x_i = \sum_{j=1}^p \alpha_j \tau_j x_i = \alpha_i \tau_i \cdot x_i = \alpha_i \tau_i \cdot \left(- \sum_{j \neq i} x_j \right) = 0$$

d'après ce qui précède. L'assertion (I) est bien prouvée.

(II) On vient de voir que si $x \in N$, alors $y_i = \alpha_i \tau_i \cdot x \in N_i$ pour tout i , et $x = y_1 + \dots + y_p$. Donc si $x \in E$, cela prouve que $g_i(x) = (\alpha_i \tau_i) \cdot x$ pour tout x , autrement dit, que $g_i = \alpha_i \tau_i$; d'où le résultat, car $\alpha_i \tau_i \in K[u]$. ■

Remarque 1 : Si les S_i sont de degré 1, $S_i = X - a_i$, il est immédiat de trouver des polynômes A_1, \dots, A_p : il suffit de prendre les A_i constants :
 $(\forall i) A_i = \frac{1}{\prod_{j \neq i} (a_i - a_j)}$ (interpolation de Lagrange !).

Remarque 2 : L'assertion (II) du théorème XV.5.1 permet de voir que les seuls sous-espaces u -stables de E sont les $\bigoplus_{i=1}^p V_i$, où pour chaque i ,

V_i est un sous- K -ev u -stable de N_i . En effet, si V est un sous- K -ev u -stable, on a d'abord $V \cap N_i$ u -stable et $\bigoplus_i (V \cap N_i) \subset V$. Et, pour $x \in V$, $g_i(x) \in N_i$, mais $g_i(x) \in V$ car $g_i \in K[u]$, donc

$$x = \sum g_i(x) \in \sum (V \cap N_i), \quad \text{d'où : } V = \bigoplus_i (V \cap N_i).$$

Endomorphismes diagonalisables

A titre de première conséquence du théorème XV.5.1, il est aisé de caractériser les endomorphismes diagonalisables :

THÉORÈME XV.5.2

|| Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. Pour que u soit **diagonalisable**, il faut et il suffit qu'il existe $p \in \mathbb{N}^*$ et des éléments $\lambda_1, \dots, \lambda_p$ de K tous **distincts** tels que $(X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_p)$ soit annulé par u .

Démonstration :

Si u est diagonalisable, soit $(\lambda_1, \dots, \lambda_p)$ une liste de ses valeurs propres distinctes. Posons $E_i = \text{Ker}(u - \lambda_i \text{Id}_E)$, d'où

$$E = \bigoplus_{i=1}^p E_i \quad \text{et} \quad u|_{E_i} = \lambda_i \text{Id}_{E_i} \quad \text{pour tout } i.$$

Soit $P(X) = (X - \lambda_1) \dots (X - \lambda_p)$ et, pour chaque i ,

$$P_i(X) = \prod_{j \neq i} (X - \lambda_j), \quad \text{d'où} \quad P(u) = P_i(u) \cdot (u - \lambda_i \text{Id}_E).$$

Cette expression montre que $P(u)$ s'annule sur E_i , donc $P(u)$ s'annule sur $E = \bigoplus_i E_i$.

Inversement, supposons satisfaite la condition de l'énoncé : le théorème XV.5.1 entraîne alors $E = \bigoplus_{i=1}^p \text{Ker}(u - \lambda_i \text{Id}_E)$, donc u est diagonalisable, ses sous-espaces propres étant ceux des $\text{Ker}(u - \lambda_i \text{Id}_E)$ qui ne sont pas nuls. ■

Exemple 1 : Supposons $K = \mathbb{C}$. Si $u^q = \text{Id}_E$ pour un certain $q \in \mathbb{N}^*$, alors u est diagonalisable, car $X^q - 1 = \prod_{\zeta \in \mu_q} (X - \zeta)$ a toutes ses racines simples.

Exemple 2 : Si $u \in \text{Hom}_K(E)$ est diagonalisable (E de dimension finie) et si F est un sous- K -ev u -stable de E , le théorème XV.5.2 montre immédiatement que $u|_F$ est diagonalisable, ce qui n'est pas évident à prouver directement.

COROLLAIRE

|| Avec les notations du théorème XV.5.2, pour que u soit **diagonalisable**, il faut et il suffit que son polynôme minimal $Q_u(X)$ soit **dissocié** dans $K[X]$ et à **facteurs simples**.

C'est une conséquence évidente du théorème XV.5.2, compte tenu de la définition du polynôme minimal, car tout facteur d'un polynôme dissocié et à facteurs simples l'est encore.

Espaces caractéristiques

DÉFINITION XV.5.1

{ Soit E un K -ev de dimension finie $n \geq 1$, $u \in \text{Hom}_K(E)$ et λ une valeur propre de u , de multiplicité α . On appelle **espace caractéristique** de u relatif à λ le sous- K -ev N_λ de E égal à $\text{Ker}(u - \lambda \text{Id}_E)^\alpha$.

Avec ces notations et hypothèses, $N_\lambda \neq \{0\}$ (car $\alpha \geq 1$, et donc N_λ contient le sous-espace propre $\text{Ker}(u - \lambda \text{Id}_E)$ associé à λ , qui est déjà $\neq \{0\}$).

De plus, soit $w = (u - \lambda \text{Id}_E)|_{N_\lambda}$. On a : $w^\alpha = 0$, i.e. w est nilpotent, de période $\leq \alpha$. On en déduit (proposition XV.4.2) que

$$\chi_w(X) = (-1)^d X^d, \quad \text{où} \quad d = \dim(N_\lambda),$$

ce qui revient à dire que, avec $v = u|_{N_\lambda}$:

$$(2) \quad \chi_v(X) = (-1)^d (X - \lambda)^d.$$

THÉORÈME XV.5.3

Les notations et hypothèses étant celles de la définition XV.5.1, on a :

$$\dim(N_\lambda) = \alpha.$$

Démonstration :

Soit $d = \dim(N_\lambda)$, $v = u|_{N_\lambda}$. On a :

$$\chi_u(X) = (X - \lambda)^\alpha P(X), \text{ avec } P(X) \in K[X] \text{ et } P(\lambda) \neq 0.$$

Notons $S = \text{Ker}(P(u))$ et $g = u|_S$. Le théorème XV.5.2 donne, en tenant compte que $\chi_u(u) = 0$ (théorème XV.4.2) : $E = N_\lambda \oplus S$. On en déduit (théorème XV.1.3) :

$$\chi_u(X) = \chi_v(X) \chi_g(X)$$

(si $S = \{0\}$, on remplace $\chi_g(X)$ par 1).

Mais on a vu en (2) que $\chi_v(X) = (-1)^d (X - \lambda)^d$. D'autre part $P(g) = 0$ et $P(\lambda) \neq 0$, donc (théorème XV.4.1) $\chi_g(\lambda) \neq 0$. Puisque

$$\chi_u(X) = (-1)^d (X - \lambda)^d \chi_g(X)$$

et que $\chi_g(\lambda) \neq 0$ il s'ensuit $d = \alpha$. ■

THÉORÈME XV.5.4

Soit E un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$ dont le polynôme caractéristique soit **dissocié** dans $K[X]$:

$$\chi_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$$

($p \geq 1$, les λ_i distincts, les $\alpha_i \geq 1$). Alors, si

$$N_i = \text{Ker}(u - \lambda_i \text{Id}_E)^{\alpha_i} \quad (1 \leq i \leq p),$$

on a

$$(3) \quad E = \bigoplus_{i=1}^p N_i \quad \text{et} \quad (\forall i) \quad \dim N_i = \alpha_i.$$

Démonstration :

On sait que $\chi_u(u) = 0$, donc $E = \bigoplus_{i=1}^p N_i$ est conséquence immédiate du théorème XV.5.2. Quant aux relations $\dim(N_i) = \alpha_i$, elles résultent du théorème XV.5.3. ■

Conservons les hypothèses du théorème XV.5.4, et posons, pour tout i , $u_i = u|_{N_i}$, $w_i = u_i - \lambda_i \text{Id}_{N_i}$, $E_i = \text{Ker}(u - \lambda_i \text{Id}_E)$. Alors chaque w_i est nilpotent, et u_i , $\lambda_i \text{Id}_{N_i}$ et w_i sont deux à deux permutables. Le théorème XV.5.4 entraîne la décomposition fondamentale :

$$(4) \quad u = \bigoplus_{i=1}^p (\lambda_i \text{Id}_{N_i} + w_i).$$

Pour chaque i , nous pouvons construire une base ordonnée \mathcal{B}_i de N_i dans laquelle w_i soit représenté par une matrice T_i *trigonale supérieure* (corollaire de la proposition XV.4.2).

En juxtaposant ces bases \mathcal{B}_i , on obtient une base \mathcal{B} de E telle que $M = \text{Mat}_{\mathcal{B}}(u)$ soit *diagonale par blocs* :

$$(5) \quad M = \text{Diag}(M_1, \dots, M_p), \quad \text{avec} \quad (\forall i) \quad M_i = \lambda_i I_{\alpha_i} + T_i \in \mathfrak{M}_{\alpha_i}(K)$$

(M_i est donc trigonale supérieure et ses termes diagonaux sont tous égaux à λ_i).

Ce résultat améliore considérablement (du moins quand $p \geq 2$), le théorème XV.2.2. L'obtention d'une base \mathcal{B} où $\text{Mat}_{\mathcal{B}}(u)$ prend la forme (5) s'appelle une *réduction de u suivant ses espaces caractéristiques*. Même si, pour $p \geq 2$, elle n'est pas facile à obtenir concrètement, son existence a néanmoins une vaste portée théorique (pour $p = 1$, on dispose au moins de l'algorithme décrit dans l'exercice 1 du § XV.4).

En passant aux matrices, on obtient le corollaire : toute matrice carrée $M \in \mathfrak{M}_n(K)$ telle que $\chi_M(X)$ soit *dissocié dans $K[X]$* est semblable à une matrice du type (5).

Somme d'un diagonalisable et d'un nilpotent permutables

Conservons les notations et hypothèses du théorème XV.5.4. Afin d'interpréter la décomposition (4), introduisons les endomorphismes :

$$(6) \quad \mathcal{D} = \bigoplus_{i=1}^p \lambda_i \text{Id}_{N_i}, \quad w = \bigoplus_{i=1}^p w_i.$$

\mathcal{D} est diagonalisable ; son polynôme caractéristique est

$$\chi_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i};$$

ses sous-espaces propres sont les N_i :

$$N_i = \text{Ker} (\mathcal{D} - \lambda_i \text{Id}_E) .$$

w est nilpotent, et puisque $\lambda_i \text{Id}_{N_i}$ et w_i sont permutables pour tout i , il s'ensuit

$$(7) \quad \mathcal{D}w = w\mathcal{D} .$$

De plus, puisque $u|_{N_i} = u_i = \lambda_i \text{Id}_{N_i} + w_i$ pour tout i , on a :

$$(8) \quad u = \mathcal{D} + w .$$

Il est remarquable que (7) et (8) caractérisent le couple (\mathcal{D}, w) de manière unique :

THÉORÈME XV.5.5

Avec les notations et hypothèses du théorème XV.5.4, il existe un et un seul couple (Δ, ν) d'endomorphismes de E tels que : Δ est diagonalisable, ν est nilpotent,

$$\Delta\nu = \nu\Delta, \quad \text{et} \quad u = \Delta + \nu .$$

De plus, pour ce couple, $\Delta \in K[u]$, $\nu \in K[u]$.

Démonstration :

En prenant pour (Δ, ν) le couple (\mathcal{D}, w) qui est défini par (6), on a vu que les conditions : \mathcal{D} est diagonalisable, w est nilpotent, $\mathcal{D}w = w\mathcal{D}$ et $u = \mathcal{D} + w$ sont bien satisfaites.

Désignons, pour tout i , par g_i le projecteur de E , sur N_i parallèlement à $\bigoplus_{j \neq i} N_j$. Le théorème XV.5.1 montre que $g_i \in K[u]$, donc

$$\mathcal{D} = \sum_{i=1}^p \lambda_i g_i \in K[u]$$

et, par différence, $w = u - \mathcal{D} \in K[u]$. Il reste à prouver l'unicité : supposons trouvé un couple (Δ, ν) satisfaisant aux conditions requises. Alors $\Delta u = u\Delta$, donc $\Delta\mathcal{D} = \mathcal{D}\Delta$, car $\mathcal{D} \in K[u]$, et on voit de même que $\nu w = w\nu$. On en déduit d'abord que $\nu - w$ est nilpotent (formule du binôme). Ensuite, on voit que \mathcal{D} et Δ admettent une base commune de vecteurs propres : car chaque N_i est Δ -stable ($x \in E$ et $\mathcal{D}(x) = \lambda_i x$ entraîne $\Delta\mathcal{D}(x) = \mathcal{D}\Delta(x) = \lambda_i \Delta(x)$) ; donc

$$\Delta(x) \in \text{Ker} (\mathcal{D} - \lambda_i \text{Id}_E) = N_i ,$$

et pour tout i , $\Delta|_{N_i}$ est diagonalisable (cf. exemple 2). En conséquence $\mathcal{D} - \Delta$ est diagonalisable. Comme $\mathcal{D} - \Delta = \nu - w$ est aussi nilpotent, c'est qu'il est nul, d'où $(\mathcal{D}, w) = (\Delta, \nu)$. ■

On observera que u est diagonalisable ssi $w = 0$.

Bien entendu, si $p \geq 2$, le calcul effectif du couple (\mathcal{D}, w) nécessite la connaissance des λ_i et des N_i dont l'obtention peut être ardue. En revanche, si $p = 1$, c'est-à-dire si u n'a qu'une seule valeur propre λ , avec $\chi_u(X)$ dissocié, ce qui signifie : si $\chi_u(X) = (\lambda - X)^n$, alors le couple (\mathcal{D}, w) est tout simplement $(\lambda \text{ Id}_E, u - \lambda \text{ Id}_E)$. Pour le déterminer il suffit de savoir calculer λ .

COROLLAIRE

Si, avec les hypothèses du théorème XV.5.4, u est de plus **inversible**, il existe un couple et un seul (Δ, μ) d'endomorphismes de E tels que : $\Delta\mu = \mu\Delta$, $u = \Delta(\text{Id}_E + \mu)$, Δ est diagonalisable et μ est nilpotent.

Démonstration :

\mathcal{D} est sûrement inversible, puisque par hypothèse tous les λ_i sont $\neq 0$. Puisque $\mathcal{D}w = w\mathcal{D}$, $\mu = \mathcal{D}^{-1}w$ est nilpotent, donc le couple $(\Delta = \mathcal{D}, \mu = \mathcal{D}^{-1}w)$ convient.

Réciproquement si un couple (Δ, μ) convient, $\Delta\mu$ est nilpotent (car $\Delta\mu = \mu\Delta$) et $u = \Delta + \Delta\mu$, $\Delta(\Delta\mu) = (\Delta\mu)\Delta$, donc $\Delta = \mathcal{D}$ et $\Delta\mu = w$ d'après le théorème XV.5.5, d'où $\mu = \mathcal{D}^{-1}w$. ■

Nous laissons au lecteur le soin d'énoncer la version matricielle du théorème XV.5.5 et de son corollaire.

Polynôme minimal

Reprenons une fois de plus les notations et hypothèses du théorème XV.5.4. Pour chaque $i \in \llbracket 1, p \rrbracket$, notons β_i la période du nilpotent

$$w_i \in \text{Hom}_K(N_i).$$

On a $1 \leq \beta_i \leq \alpha_i$. β_i peut être appelé **l'indice** de la valeur propre λ_i . Puisque $0 = w_i^{\beta_i} = (u_i - \lambda_i \text{ Id}_{N_i})^{\beta_i}$ pour tout i , le polynôme $P(X) = \prod_{i=1}^p (X - \lambda_i)^{\beta_i}$ est annulé par chaque u_i , donc par u . En fait :

PROPOSITION XV.5.2

Avec les notations ci-dessus, le polynôme $P(X) = \prod_{i=1}^p (X - \lambda_i)^{\beta_i}$ est le polynôme minimal de u , et on a : $N_i = \text{Ker}(u - \lambda_i \text{ Id}_E)^{\beta_i}$ pour tout i .

Démonstration :

On a déjà vu que $P(X)$ est annulé par u , donc $P(X)$ est multiple du polynôme minimal $Q_u(x)$, d'où

$$Q_u(X) = \prod_{i=1}^p (X - \lambda_i)^{\gamma_i} \quad \text{avec} \quad 0 \leq \gamma_i \leq \beta_i \quad \text{pour tout } i.$$

Le théorème XV.5.1 montre donc que $E = \bigoplus_{i=1}^p \text{Ker} (u - \lambda_i \text{Id}_E)^{\gamma_i}$, et puisque

$$\text{Ker} (u - \lambda_i \text{Id}_E)^{\gamma_i} \subset N_i \quad (1 \leq i \leq p),$$

il s'ensuit $\text{Ker} (u - \lambda_i \text{Id}_E)^{\gamma_i} = N_i$ (ce qui prouve au passage que $\gamma_i \geq 1$, puisque $N_i \neq \{0\}$, résultat qui découle aussi directement du théorème XV.4.1). On en déduit, pour chaque i , que $w_i^{\gamma_i} = 0$, donc que $\gamma_i \geq \beta_i$ (par définition de la période), d'où finalement $\gamma_i = \beta_i$. ■

Dire que $\beta_i = 1$ pour un $i \in \llbracket 1, p \rrbracket$ signifie que $w_i = 0$, donc que N_i est égal au sous-espace propre E_i de u associé à la valeur propre λ_i . En particulier, les facteurs de $Q_u(X)$ sont tous simples ssi u est diagonalisable. On retrouve ainsi une partie du corollaire du théorème XV.5.2.

Les applications des résultats de ce paragraphe sont innombrables et ne sauraient donc être recensées ici. Bornons-nous à quelques exemples représentatifs :

Exemple 3 : Puissances d'une matrice.

Soit $n \in \mathbb{N}^*$ et $M = [a_{ij}] \in \mathfrak{M}_n(\mathbb{C})$. Proposons-nous de chercher à quelle condition la matrice M^k est le terme général d'une suite $(M^k)_{k \in \mathbb{N}}$ ayant pour limite 0 quand $n \rightarrow +\infty$. Si nous fixons la matrice $P \in \text{GL}(n, \mathbb{C})$, il revient au même de chercher si la suite

$$(P M^k P^{-1})_{k \in \mathbb{N}} = ((P M P^{-1})^k)_{k \in \mathbb{N}}$$

tend vers zéro.

Choisissons donc $P \in \text{GL}(n, \mathbb{C})$ telle que $P M P^{-1} = N$ soit *diagonale par blocs*, de la forme (5) :

$$N = \text{Diag} (N_1, \dots, N_p) \quad \text{avec} \quad N_i = \lambda_i I_{\alpha_i} + T_i \in \mathfrak{M}_{\alpha_i}(\mathbb{C}),$$

T_i trigonale supérieure à valeurs propres nulles, pour chaque $i \in \llbracket 1, p \rrbracket$, et $\prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ désignant la décomposition du polynôme caractéristique $\chi_M(X)$ ($p \geq 1$, les $\alpha_i \geq 1$, les λ_i distincts).

Pour $k \in \mathbb{N}$, on a $N^k = \text{Diag}(N_1^k, \dots, N_p^k)$, donc (N^k) tend vers 0 ssi chaque (N_i^k) tend vers zéro. Fixons $i \in \llbracket 1, p \rrbracket$. Alors, du fait que $\lambda_i I_{\alpha_i}$ et T_i sont permutables :

$$(9) \quad (\forall k \geq n) \quad N_i^k = (\lambda_i I_{\alpha_i} + T_i)^k = \sum_{j=0}^{\alpha_i-1} \binom{k}{j} \lambda_i^{k-j} T_i^j,$$

car $(T_i)^j = 0$ si $j \geq \alpha_i$. Prenons sur $\mathfrak{M}_{\alpha_i}(\mathbb{C})$ la norme (cf. cours d'Analyse)

$$Z = [z_{q,r}] \mapsto \|Z\| = \max_{(q,r)} |z_{qr}|.$$

De (9), on déduit pour $k \geq n$:

$$(10) \quad \|N_i^k\| \leq |\lambda_i|^{k-n} f_i(k), \text{ avec } f_i(k) = \sum_{j=0}^{\alpha_i-1} \binom{k}{j} \|T_i^j\| \times |\lambda_i|^{n-j}.$$

Si $|\lambda_i| < 1$, (10) montre que $\|N_i^k\| \xrightarrow{k \rightarrow \infty} 0$, car $f_i(k)$ est une fonction *polynomiale* de k .

Si $|\lambda_i| \geq 1$, soit β_i la période de T_i . Pour $k \geq n$, (9) permet d'écrire :

$$(11) \quad N_i^k = \binom{k}{\beta_i-1} \lambda_i^k [(T_i)^{\beta_i-1} + g_i(k)],$$

$$\text{avec } g_i(k) = \sum_{j=0}^{\beta_i-2} \binom{k}{j} \binom{k}{\beta_i-1}^{-1} (\lambda_i)^{-j} (T_i)^j$$

$$(\text{si } \beta_i = 1, g_i(k) = 0 \text{ et } T_i^{\beta_i-1} = I_{\alpha_i}).$$

Comme $\binom{k}{r}$ est une fonction polynomiale de k de degré r (pour r fixé), on voit que $g_i(k) \xrightarrow{k \rightarrow \infty} 0$. Or, $(T_i)^{\beta_i-1} \neq 0$. Donc (11) nous donne *un équivalent* de $\|N_i^k\|$ pour $k \rightarrow +\infty$:

$\|N_i^k\| \underset{k \rightarrow \infty}{\sim} \binom{k}{\beta_i-1} \|T_i^{\beta_i-1}\| |\lambda_i|^k$, et on voit que $\|N_i^k\| \xrightarrow{k \rightarrow \infty} +\infty$ si $\beta_i \geq 2$, ou si $|\lambda_i| > 1$, et que $\|N_i^k\| \xrightarrow{k \rightarrow \infty} 1$ si $|\lambda_i| = 1$ et $\beta_i = 1$. En résumé, on peut conclure que

$$\boxed{M^k \xrightarrow{k \rightarrow \infty} 0 \quad \text{ssi} \quad (\forall i) \quad |\lambda_i| < 1}.$$

Exemple 4 : Exponentielles d'endomorphismes ou de matrices.

Le corps de base sera ici $K = \mathbb{R}$ ou \mathbb{C} . Soit $n \in \mathbb{N}^*$ et $M \in \mathfrak{M}_n(K)$, et supposons $\chi_M(X)$ dissocié sur K (évidemment si $K = \mathbb{C}$, cette

toujours vérifiée). On a donc

$$\chi_M(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i} \quad (p \geq 1, \text{ les } \alpha_i \geq i, \text{ les } \lambda_i \text{ distincts}).$$

Pour étudier commodément $\exp(M)$, raisonnons comme dans l'exemple 1, en remplaçant M par

$$N = PMP^{-1} = \text{Diag} (N_1, \dots, N_p)$$

(par blocs), pour $P \in \text{GL}(n, K)$ convenable, avec, pour tout i , $N_i = \lambda_i I_{\alpha_i} + T_i$, T_i trigonale supérieure et nilpotente dans $\mathfrak{M}_{\alpha_i}(K)$ (l'existence de P est assurée par le théorème XV.5.4, cf. (4) et (5)).

On a alors :

$$\exp(M) = P \exp(N) P^{-1} \quad \text{et} \quad \exp(N) = \text{Diag} (\exp(N_1), \dots, \exp(N_p)).$$

Fixons $i \in \llbracket 1, p \rrbracket$. On a, du fait que $\lambda_i I_{\alpha_i}$ et T_i sont permutables :

$$(12) \quad \exp(N_i) = \exp(\lambda_i I_{\alpha_i}) \exp(T_i) = e^{\lambda_i} \exp(T_i).$$

Mais, notant β_i la période de la matrice nilpotente T_i :

$$\exp(T_i) = \sum_{j=0}^{\beta_i-1} \frac{1}{j!} (T_i)^j.$$

La série se réduit donc à un polynôme en T_i . Finalement :

$$(13) \quad \exp(N_i) = e^{\lambda_i} \sum_{j=0}^{\beta_i-1} \frac{1}{j!} (T_i)^j.$$

Comme application, cherchons toutes les matrices M telles que

$$\exp(M) = I_n,$$

ce qui équivaut à $\exp(N) = I_n$, c'est-à-dire à $\exp(N_i) = I_{\alpha_i}$ pour tout i .

Fixons i . Ecrivons (13) sous la forme :

$$\exp(N_i) = e^{\lambda_i} I_{\alpha_i} + W_i, \quad \text{où} \quad W_i = \sum_{j=1}^{\beta_i-1} \frac{1}{j!} (T_i)^j$$

(si $\beta_i = 1$, $W_i = 0$). Alors $W_i = T_i h_i(T_i)$, où h_i est un polynôme, donc W_i est nilpotente car T_i et $h_i(T_i)$ sont permutables. Et W_i est permutable avec $e^{\lambda_i} I_{\alpha_i}$. Le théorème XV.5.5 montre donc l'équivalence

$$\exp(N_i) = I_{\alpha_i} \Leftrightarrow e^{\lambda_i} I_{\alpha_i} = I_{\alpha_i} \quad \text{et} \quad W_i = 0.$$

Enfin,

$$h_i(T_i) = I_{\alpha_i} + \varphi_i(T_i) T_i ,$$

où φ_i est un nouveau polynôme, donc $h_i(T_i) = I_{\alpha_i} + U_i$, avec U_i nilpotente (si $\beta_i = 1$, on pose $h_i(T_i) = I_{\alpha_i}$), donc $h_i(T_i)$ est *invertible*, d'inverse

$$I_{\alpha_i} - U_i + U_i^2 - \cdots + (-1)^{n-1} U_i^{n-1} ,$$

et par suite $W_i = 0$ ssi $T_i = 0$. On en déduit donc : $\exp(N_i) = I_{\alpha_i}$ ssi $T_i = 0$ et $e^{\lambda_i} = 1$. Par suite, $\exp(M) = I_n \Leftrightarrow M$ est diagonalisable et toutes ses valeurs propres vérifient $e^{\lambda_i} = 1$. En particulier, pour $K = \mathbb{C}$, cela signifie que toutes ses valeurs propres sont des multiples entiers de $2i\pi$.

Pour $K = \mathbb{R}$, si l'on se souvient que nous avons imposé au début la condition « $\chi_M(X)$ dissocié sur K », on ne trouve comme seule solution de $\exp(M) = I_n$ que la matrice $M = 0$. En revanche, si l'on ne s'impose plus cette condition, il ne faudra pas s'étonner de trouver beaucoup plus de solutions : on prouvera par exemple facilement que

$$M = \begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix} \in \mathfrak{M}_2(\mathbb{R})$$

est telle que $\exp(M) = I_2$.

Remarque 3 : Sous les hypothèses et notations de l'exemple 4, supposons $p = 1$, i.e. $\chi_M(X)$ de la forme $(\lambda - X)^n$. Alors le calcul de $\exp(M)$ est *immédiat*, sans même avoir besoin d'un changement de base. En effet, en posant $W = M - \lambda I_n$, on est sûr que W est nilpotente et comme $M = \lambda I_n + W$, on a directement, par la formule du binôme :

$$(14) \quad \exp(M) = \exp(\lambda I_n + W) = \exp(\lambda I_n) \sum_{k=0}^{n-1} \frac{1}{k!} W^k = e^\lambda \sum_{k=0}^{n-1} \frac{1}{k!} W^k .$$

C'est évidemment la circonstance la plus favorable qui puisse se présenter pour calculer une exponentielle de matrice, plus favorable encore que le cas où M a toutes ses valeurs propres dans K , et de multiplicité 1.

Exemple 5 : Soit $M \in \text{GL}(n, \mathbb{C})$ ($n \geq 1$), et $p \in \mathbb{N}^*$ ($p \geq 2$). Montrer que, si M est diagonalisable, toute matrice $N \in \mathfrak{M}_n(\mathbb{C})$ telle que $N^p = M$ est encore diagonalisable.

Solution : Nous allons ici utiliser directement le théorème XV.5.5, sous sa forme matricielle. Supposons $N^p = M$, et soit D et W les matrices, éléments de $\mathfrak{M}_n(\mathbb{C})$, telles que $N = D + W$, $DW = WD$, D diagonalisable et W nilpotente. Alors la formule du binôme s'applique et donne :

$$N^p = D^p + WA , \quad \text{avec} \quad A = \sum_{k=1}^p \binom{p}{k} D^{p-k} W^{k-1} .$$

Mais, puisque $DW = WD$, on a aussi $WA = AW$; donc WA est nilpotente et $(WA) D^p = D^p(WA)$. Puisque $M = D^p + WA$ est diagonalisable, une nouvelle application du théorème XV.5.5 montre que $D^p = M$ et $WA = 0$.

Or,
$$A = pD^{p-1}(I_n + WB), \quad \text{avec} \quad B = \frac{1}{p} (D^{-1})^{p-1} \sum_{k=2}^p D^{p-k} W^{k-2},$$

formule qui a un sens car D est inversible ($D^p = M \in \text{GL}(n, \mathbb{C})$). W et B sont permutables, donc WB est nilpotente, donc $I_n + WB$ est inversible (son inverse est $I_n - (WB) + (WB)^2 - \dots$). Donc A est aussi inversible et $WA = 0$ entraîne $W = 0$. Par suite $N = D$ est bien diagonalisable.

Exercice 1 : Soit $n \in \mathbb{N}^*$ et $M \in \mathfrak{M}_n(\mathbb{R})$ telle que le polynôme $\chi_M(X)$ soit *dissocié* sur \mathbb{R} . Montrer que, pour que M soit diagonalisable dans $\mathfrak{M}_n(\mathbb{R})$, il faut et il suffit qu'elle le soit dans $\mathfrak{M}_n(\mathbb{C})$. Généraliser en remplaçant \mathbb{R} et \mathbb{C} par un corps commutatif K et une extension L de K .

Exercice 2 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$ dont le polynôme caractéristique est *dissocié* dans $K[X]$. Montrer que u est diagonalisable ssi $(\forall \lambda \in \mathbb{C}) \text{rg}(u - \lambda \text{Id}_E) = \text{rg}(u - \lambda \text{Id}_E)^2$.

Exercice 3 : Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$, supposé *diagonalisable*. On donne un sous- K -ev F de E qui est u -stable. Montrer que F admet dans E un supplémentaire u -stable.

Exercice 4 : Soit E un K -ev de dimension finie $n \geq 1$. On donne une partie non vide \mathcal{H} de $\text{Hom}_K(E)$ telle que $(\forall u \in \mathcal{H}) (\forall v \in \mathcal{H}) uv = vu$, et $(\forall u \in \mathcal{H}) u$ est diagonalisable. Montrer qu'il existe une base \mathcal{B} de E telle que, pour tout $u \in \mathcal{H}$, \mathcal{B} est une base de vecteurs propres de u (raisonner par récurrence sur n).

Exercice 5 : Soit E un K -ev non nul, $u \in \text{Hom}_K(E)$ et S_1, S_2 deux éléments de $K[X]$, D leur pgcd. Montrer :

$$\text{Ker}(S_1(u)) \cap \text{Ker}(S_2(u)) = \text{Ker}(D(u)).$$

Exercice 6 : Soit E un \mathbb{C} -ev de dimension finie $n \geq 1$ et $f \in \text{Hom}_{\mathbb{C}}(E)$. On note T_f l'endomorphisme de $\text{Hom}_{\mathbb{C}}(E)$ tel que

$$T_f(u) = f \circ u - u \circ f, \quad \text{pour } u \in \text{Hom}_{\mathbb{C}}(E).$$

Montrer : $(T_f \text{ est diagonalisable}) \Leftrightarrow (f \text{ diagonalisable})$.

Exercice 7 : a) Soit $M \in \mathfrak{M}_2(\mathbb{Z})$. On suppose qu'il existe un $k \in \mathbb{N}^*$ tel que $M^k = I_2$. Montrer que $M^{12} = I_2$ (remarquer que M est diagonalisable, étudier ses valeurs propres).

b) Plus généralement, soit $p \in \mathbb{N}$, $p \geq 3$. Montrer par la même méthode qu'il existe un entier $f(p)$ ne dépendant que de p tel que, pour toute matrice $M \in \mathfrak{M}_p(\mathbb{Z})$ vérifiant $M^k = I_p$ pour au moins un $k \in \mathbb{N}^*$, on ait : $M^{f(p)} = I_p$.

Exercice 8 : Soit p et n deux entiers ≥ 2 et $A_0, A_1, \dots, A_{p-1} \in \mathfrak{M}_n(\mathbb{C})$ des matrices deux à deux *permutables* et *diagonalisables*. On considère la matrice circulante par blocs $M \in \mathfrak{M}_{np}(\mathbb{C})$:

$$M = \begin{bmatrix} A_0 & A_1 & \dots & A_{p-1} \\ A_{p-1} & A_0 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \dots & \dots & A_0 \end{bmatrix}.$$

Montrer que M est diagonalisable.

Indication : Penser à l'exercice 4 ci-dessus.

Exercice 9 : Soit $n \in \mathbb{N}$ ($n \geq 2$) et E un \mathbb{C} -ev de dimension n . Montrer qu'il existe un voisinage V de Id_E dans $\text{Hom}_{\mathbb{C}}(E)$ tel que $\{\text{Id}_E\}$ soit le seul sous-groupe de $\text{GL}_{\mathbb{C}}(E)$ inclus dans V .

Exercice 10 : Soit $n \in \mathbb{N}^*$, $a = (a_1, \dots, a_n) \in \mathbb{C}^n$ et $M_a \in \mathfrak{M}_{n+1}(\mathbb{C})$ définie par

$$M_a = \begin{bmatrix} 0 & a_1 & a_2 & \dots & \dots & a_n \\ a_1 & 0 & a_2 & \dots & \dots & a_n \\ a_1 & a_2 & 0 & \dots & \dots & a_n \\ \vdots & \vdots & & \ddots & & \vdots \\ a_1 & a_2 & a_3 & \dots & a_n & 0 \end{bmatrix}.$$

On note f_a l'élément de $\text{Hom}_{\mathbb{C}}(\mathbb{C}^{n+1})$ tel que, si $\mathcal{B} = (e_1, \dots, e_{n+1})$ est la base canonique de \mathbb{C}^{n+1} , $\text{Mat}_{\mathcal{B}}(f_a) = M_a$.

a) Valeurs propres de f_a ?

b) Soit $s = \sum_{i=1}^n a_i$. Montrer : si $s \neq -a_j$ pour tout $j \in \llbracket 1, n \rrbracket$, alors f_a est diagonalisable, et

donner dans ce cas une base de vecteurs propres. Exprimer une telle base en fonction des vecteurs ε_i :

$$\varepsilon_0 = \sum_{i=1}^{n+1} e_i \quad \text{et} \quad \varepsilon_i = \sum_{j=1}^i e_j \quad \text{pour } i \in \llbracket 1, n \rrbracket.$$

c) Calculer $\text{Mat}_{(\varepsilon_0, \dots, \varepsilon_n)}(f_a)$. En déduire tous les a pour lesquels f_a est diagonalisable, et lorsqu'il l'est, donner une base de vecteurs propres.

d) Polynôme minimal de f_a ?

Exercice 11 : Soit $n \in \mathbb{N}$ ($n \geq 2$) et $\mathcal{N}(n, \mathbb{C})$ l'ensemble des matrices nilpotentes dans $\mathfrak{M}_n(\mathbb{C})$. On donne une série formelle $S \in \mathbb{C}[[X]]$ de valuation 1. Démontrer que l'application

$$\hat{S} : \mathcal{N}(n, \mathbb{C}) \rightarrow \mathcal{N}(n, \mathbb{C}), \quad M \mapsto S(M)$$

est bijective, en montrant d'abord, que pour M nilpotente fixée, $S \mapsto \hat{S}(M), \mathbb{C}[[X]] \rightarrow \mathfrak{M}_n(\mathbb{C})$ est un homomorphisme de \mathbb{C} -algèbres.

Exercice 12 : Soit $n \in \mathbb{N}^*$. Une matrice $S = [p_{ij}]_{(i,j) \in \llbracket 1, n \rrbracket^2} \in \mathfrak{M}_n(\mathbb{R})$ est dite *stochastique* ssi

$$\forall (i, j) \quad p_{ij} \in \mathbb{R}_+ \quad \text{et} \quad (\forall i) \quad \sum_{j=1}^n p_{ij} = 1.$$

Elle est dite *stochastique stricte* ssi, de plus, on a : $\forall (i, j) \quad p_{ij} > 0$. On notera \mathcal{S} (resp. \mathcal{S}^*) l'ensemble des matrices stochastiques (resp. stochastiques strictes) de $\mathfrak{M}_n(\mathbb{R})$. Ces ensembles sont stables par le produit.

a) Soit $S = [p_{ij}] \in \mathcal{S}^*$, S fixée. On note U l'élément de $\text{Hom}_{\mathbb{C}}(\mathbb{C}^n)$ dont la matrice, dans la base canonique $\mathcal{C} = (e_1, \dots, e_n)$ de $E = \mathbb{C}^n$, est S .

1) Montrer que 1 est valeur propre de U et que $\dim_{\mathbb{R}}(\text{Ker}(U - \text{Id}_E)) = 1$.

2) Soit λ une valeur propre de U autre que 1. Montrer que $|\lambda| < 1$.

Indication : Considérer un vecteur propre $x = (x_1, \dots, x_n)$ associé à λ et choisir $i \in \llbracket 1, n \rrbracket$ tel que $|x_i| = \text{Max}_{j \in \llbracket 1, n \rrbracket} (|x_j|)$.

3) On munit E de la norme

$$x = (x_1, \dots, x_n) \mapsto \text{Max}_{i \in \llbracket 1, n \rrbracket} (|x_i|) = \|x\|.$$

• Si $x \in E$, comparer $\|U(x)\|$ et $\|x\|$.

• Soit $N_1 = \text{Ker}(U - \text{Id}_E)^{\alpha_1}$, où α_1 est la multiplicité de la valeur propre 1 de U . Soit U_1 induit par U sur N_1 et $\nu = U_1 - \text{Id}_{N_1}$. Montrer : $\nu \neq 0 \Leftrightarrow \alpha_1 \geq 2$. On suppose $\alpha_1 \geq 2$ et on note β la période de ν et y un vecteur de N_1 tel que $\nu^{\beta-1}(y) \neq 0$. Trouver $\lim_{k \rightarrow \infty} \|U^k(x)\|$.

- En déduire finalement que $\alpha_1 = 1$.
- 4) Soit P la somme vectorielle des espaces caractéristiques de U autres que N_1 , et soit $\varphi : E \rightarrow N_1$ la projection de E sur N_1 parallèlement à P .
 - Si $x \in E$, on a $\lim_{m \rightarrow \infty} U^m(x) = \varphi(x)$.
- En déduire qu'il existe $c_1, c_2, \dots, c_n \in \mathbb{R}_+$ tels que $\sum c_i = 1$ et que

$$\lim_{m \rightarrow \infty} S^m = \begin{bmatrix} c_1 & c_2 & \dots & c_n \\ c_1 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ c_1 & c_2 & \dots & c_n \end{bmatrix}.$$

b) Etudier de même le cas où S fixée est une matrice de \mathcal{S} . Montrer que les valeurs propres d'une telle S dont le module est égal à 1 sont en réalité des racines N -ièmes de 1 pour N convenable.

c) Application numérique :

$$S = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \quad \text{et} \quad S = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercice 13 : Soit $A \in \mathfrak{M}_n(\mathbb{C})$ ($n \geq 2$). Trouver une condition nécessaire et suffisante pour que l'ensemble $\{A^k\}_{k \in \mathbb{N}}$ soit borné dans $\mathfrak{M}_n(\mathbb{C})$.

Exercice 14 : Soit $n \in \mathbb{N}^*$ ($n \geq 2$) et $a \in \mathbb{C}$

a) Pour $M = \begin{bmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a^2 \\ \vdots & \vdots & \vdots & \ddots & a \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}$, donner une expression de $\exp(M)$.

b) Pour $p \in \mathbb{N}^*$ donné, trouver $N \in \mathfrak{M}_n(\mathbb{C})$ telle que $N^p = M$.

Exercice 15 : On donne $n \in \mathbb{N}^*$ ($n \geq 2$), et $M \in \text{GL}(n, \mathbb{C})$. A l'aide de l'exercice 11, trouver toutes les matrices $N \in \mathfrak{M}_n(\mathbb{C})$ telles que $\exp(N) = M$.

Exercice 16 : On reprend les notations de l'exercice 11.

a) Soit $\mathcal{E} : \mathcal{N}(n, \mathbb{C}) \rightarrow \mathfrak{M}_n(\mathbb{C})$, $N \mapsto \sum_{k=0}^{\infty} \frac{1}{k!} N^k$. Montrer que \mathcal{E} définit une bijection de $\mathcal{N}(n, \mathbb{C})$ sur l'ensemble Γ des $M \in \mathfrak{M}_n(\mathbb{C})$ telles que $M - I_n \in \mathcal{N}(n, \mathbb{C})$, dont la réciproque est \mathcal{L} :

$$M \mapsto \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (M - I_n)^k.$$

b) On donne $\alpha \in \mathbb{C}$. Si $M \in \Gamma$, soit $M^\alpha = \mathcal{E}(\alpha \mathcal{L}(M))$. Montrer que si on pose $S_\alpha = (1 + X)^\alpha \in \mathbb{C}[[X]]$, on a $M^\alpha = \hat{S}_\alpha(M - I_n)$; que si $\alpha \neq 0$ l'application $\Gamma \rightarrow \Gamma$, $M \mapsto M^\alpha$ est surjective; et que, pour $M \in \Gamma$ fixée, on a :

$$\forall (\alpha, \beta) \in \mathbb{C}^2 \quad M^{\alpha+\beta} = M^\alpha M^\beta.$$

c) Soit $M \in \text{GL}(n, \mathbb{C})$ et $q \in \mathbb{N}^*$. Montrer qu'il existe $N \in \text{GL}(n, \mathbb{C})$ telle que $N^q = M$ et $N \in \mathbb{C}[M]$. Trouver l'ensemble de tous ces N ; cas particulier où M est diagonalisable.

Application numérique : Résoudre $N^2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ 8 & 2 & 4 \end{bmatrix}$; $N^2 = \begin{bmatrix} 1 & -5 & 5 \\ 0 & 3 & -3 \\ 0 & -3 & 3 \end{bmatrix}$.

Exercice 17 : On donne $x \in \mathbb{N}^*$ ($n \geq 2$) et $f \in \mathbb{C}_n[X]$, f non constant.

a) Si M et N sont deux éléments permutables dans $\mathfrak{M}_n(\mathbb{C})$, prouver :

$$f(M + N) = f(M) + \frac{1}{1!} f'(M) N + \dots + \frac{1}{k!} f^{(k)}(M) N^k + \dots$$

- b) Chercher les couples $(\lambda, M) \in \mathbb{C} \times \mathfrak{M}_n(\mathbb{C})$ tels que : M est nilpotent et $f(\lambda I_n + M) = 0$.
 c) Décrire toutes les solutions dans $\mathfrak{M}_n(\mathbb{C})$ de l'équation à l'inconnue M : $f(M) = 0$. Vérifier que le nombre des polynômes caractéristiques possibles pour M est fini.
 d) Application numérique : Résoudre dans $\mathfrak{M}_2(\mathbb{C})$ l'équation $M^2 + M + I_2 = 0$ et dans $\mathfrak{M}_3(\mathbb{C})$ les équations

$$M^3 - 3M + 2I_3 = 0 \quad \text{et} \quad M^3 + M = 0.$$

Résoudre également l'équation $M^3 + M = 0$ dans $\mathfrak{M}_3(\mathbb{R})$.

Montrer que si $M^2 + I = 0$ ($M \in \mathfrak{M}_2(\mathbb{R})$), M est semblable à $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Résoudre cette équation dans $\mathfrak{M}_4(\mathbb{R})$.

Exercice 18 : Le corps de base est \mathbb{C} . Soit $n \in \mathbb{N}$ ($n \geq 2$) et E un \mathbb{C} -ev de dimension n , muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. A $\sigma \in \mathfrak{S}_n$, on associe $f_\sigma \in \text{Hom}_{\mathbb{C}}(E)$ tel que $f_\sigma(e_i) = e_{\sigma(i)}$ pour tout i .

a) Si σ est le cycle $\langle 1, 2, \dots, n \rangle$, calculer $\chi_{f_\sigma}(X)$. Montrer que f_σ est diagonalisable et donner une base de vecteurs propres. Que vaut $Q_{f_\sigma}(X)$?

b) σ étant quelconque, décomposer σ en cycles, montrer que f_σ est diagonalisable et donner une base de vecteurs propres ainsi que $\chi_{f_\sigma}(X)$.

Exprimer $Q_{f_\sigma}(X)$ en utilisant les polynômes cyclotomiques (cf. exercice 16 du § VII.4 et exercice 12 du § IX.7).

Exercice 19 : a) Calculer l'exponentielle de

$$M = \begin{bmatrix} 1 & -4 & 2 & -1 \\ 0 & -3 & 2 & -1 \\ -1 & -15 & 9 & -4 \\ -2 & -14 & 8 & -3 \end{bmatrix} \quad (M \in \mathfrak{M}_4(\mathbb{C})).$$

b) Pour cette même matrice, calculer une matrice N telle que $N^p = M$ (cf. exercice 16 ci-dessus), où $p \in \mathbb{N}^*$ est donné.

Exercice 20 : On donne une extension L du corps K et $M \in \mathfrak{M}_n(K)$ ($n \in \mathbb{N}^*$).

a) Soit $Q_M^{[K]}(X)$ (resp. $Q_M^{[L]}(X)$) le polynôme minimal de M dans $\mathfrak{M}_n(K)$ (resp. $\mathfrak{M}_n(L)$). Prouver : $Q_M^{[K]}(X) = Q_M^{[L]}(X)$.

b) En déduire : M diagonalisable dans $\mathfrak{M}_n(K) \Leftrightarrow M$ diagonalisable dans $\mathfrak{M}_n(L)$ sous réserve que $Q_M(X)$ soit dissocié dans $K[X]$.

Exercice 21 : Algorithme de trigonalisation.

On donne l'entier $n \geq 2$, le K -ev E de dimension n et $u \in \text{Hom}_K(E)$. On note, comme au § XV.4, $\chi_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ le polynôme caractéristique de u dissocié dans $K[X]$, et

$Q_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\beta_i}$ son polynôme minimal.

a) Prouver que pour tout i et pour tout k , et pour tout sous- K -ev H de E tel que

$$\text{Ker}(\varphi_i)^k \subset H \subset \text{Ker}(\varphi_i)^{k+1},$$

H est u -stable, φ_i désignant l'endomorphisme $u - \lambda_i \text{Id}_E$ ($1 \leq i \leq p$).

b) En déduire que si on complète la chaîne

$$\{0\} \subsetneq \text{Ker}(\varphi_i) \subsetneq \text{Ker}(\varphi_i)^2 \subsetneq \dots \subsetneq \text{Ker}(\varphi_i)^{\beta_i} = N_i$$

de façon quelconque en un drapeau de E , ce drapeau est u -stable.

c) Déduire de ce qui précède une méthode pour trigonaliser u .

Application numérique : trigonaliser les matrices suivantes :

$$A = \begin{bmatrix} 67 & 82 & 24 & -14 \\ -51 & -63 & -19 & 8 \\ -3 & -4 & -2 & -1 \\ 12 & 15 & 5 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 61 & 61 & 83 & 71 \\ -47 & -51 & 9 & 44 \\ -3 & -3 & 1 & 4 \\ 11 & 12 & -2 & -10 \end{bmatrix}.$$

Exercice 22 : Soit L une extension du corps K . On donne $x \in L$ algébrique de degré d sur K . On note $\varphi_x(X)$ le polynôme minimal de x , au sens de la définition VII.4.2.

a) Soit E le K -ev $K[x]$ (qui est un corps). On considère l'endomorphisme $f_x \in \text{Hom}_K(E)$ tel que $f_x(t) = tx$ pour tout $t \in E$. Prouver :

$$\varphi_x(X) = (-1)^d \chi_{f_x}(X) = Q_{f_x}(X).$$

b) Si $y \in K[x]$, trouver $\varphi_y(X)$ et $\chi_{f_y}(X)$ ($f_y \in \text{Hom}_K(E)$, $f_y : t \mapsto ty$).

Exercice 23 : Soit E un \mathbb{C} -ev de dimension finie $p \geq 2$ et n un entier $\geq p$. Un élément $f \in \text{GL}_{\mathbb{C}}(E)$ est dit *cyclique de rang n* ssi il existe une partie finie Γ de E , de cardinal n , qui engendre E comme \mathbb{C} -ev, qui soit f -stable, et telle que l'application $\Gamma \rightarrow \Gamma$, $x \mapsto f(x)$ soit un cycle de longueur n . Lorsqu'il en est ainsi, toute partie Γ vérifiant ces conditions sera appelée un *n -cycle propre* de f . Dans les questions 1), 2) a)-b)-c) et 3), f désigne un élément fixé de $\text{GL}_{\mathbb{C}}(E)$ cyclique de rang n .

1) Soit Γ un n -cycle propre de f et $a \in \Gamma$.

a) Prouver que la suite $(a, f(a), \dots, f^{p-1}(a))$ est une base du \mathbb{C} -ev E .

b) Montrer que f est un élément d'ordre n du groupe $\text{GL}_{\mathbb{C}}(E)$. En déduire qu'il existe des racines n -ièmes de 1 : $\lambda_1, \dots, \lambda_q$ ($q \geq 1$), distinctes, telles que pour tout i λ_i soit valeur propre de f et qu'on ait (1) $Q(f) = 0$, où $Q \in \mathbb{C}[X]$ est défini par $Q(X) = \prod_{i=1}^q (\lambda_i - X)$. Prouver que f est diagonalisable.

c) En utilisant a), prouver que dans (1) on a forcément $q = p$. En déduire $\chi_f(X)$.

2) On pose $\omega = \exp\left(\frac{2i\pi}{n}\right)$; a désigne toujours un point donné du n -cycle propre Γ de f .

On définit les vecteurs $(W_k)_{0 \leq k \leq n-1}$ de E par :

$$(2) \quad (\forall k \in \llbracket 0, n-1 \rrbracket) \quad W_k = \sum_{j=0}^{n-1} \omega^{kj} f^j(a).$$

a) Calculer $f(W_k)$.

b) Quel est le rang du système de vecteurs $(W_k)_{0 \leq k \leq n-1}$? En déduire qu'il existe $J \subset \llbracket 0, n-1 \rrbracket$ tel que $(W_k)_{k \in J}$ soit une base de vecteurs propres de f . Étudier les vecteurs $(W_k)_{k \in J}$. Calculer les $f^j(a)$ en fonction des $(W_k)_{k \in J}$.

c) Étudier les valeurs propres $(\omega^{-k})_{k \in J}$ de f .

d) Réciproquement, soit $J \subset \llbracket 0, n-1 \rrbracket$ une partie de cardinal p et soit $(W_k)_{k \in J}$ une base de E . On définit $g \in \text{GL}_{\mathbb{C}}(E)$ en posant, pour $k \in J$, $g(W_k) = \omega^k W_k$. Donner une CNS, portant sur les entiers (p.g.c.d. $(k, n)_{k \in J}$), pour que g soit cyclique de rang n . En déduire qu'il existe des automorphismes de E cycliques de rang n .

e) Soit $g \in \text{GL}_{\mathbb{C}}(E)$ un élément d'ordre n du groupe $\text{GL}_{\mathbb{C}}(E)$. A quelle CNS, portant sur les valeurs propres de g , g est-il cyclique de rang n ?

3) On conserve les notations a , Γ , W_k et ω du 2) et on suppose $f \in \text{GL}_{\mathbb{C}}(E)$ cyclique de rang n .

a) Comment choisir $b \in E$ pour que l'ensemble $\Omega(b) = \{b, f(b), \dots, f^{n-1}(b)\}$ soit un n -cycle propre de f ?

b) Lorsque $\Omega(b)$ n'est pas un n -cycle propre de f , discuter, suivant b , l'entier $\text{card}(\Omega(b))$.

4) On donne l'entier $n \geq p$. Soit $\mathcal{B} = (V_0, V_1, \dots, V_{p-1})$ une base de E . Montrer qu'il existe $f \in \text{GL}_{\mathbb{C}}(E)$ cyclique de rang n et des vecteurs $V_p, V_{p+1}, \dots, V_{n-1}$, tels que $f(V_i) = V_{i+1}$ pour $i \leq n-2$, $f(V_{n-1}) = V_0$ et que $\Gamma = \{V_0, V_1, \dots, V_{n-1}\}$ soit un n -cycle propre de f .

La base \mathcal{B} étant fixée, montrer que l'ensemble des $f \in \text{GL}_{\mathbb{C}}(E)$ qui conviennent est fini.

Indication : examiner $\text{Mat}_{\mathcal{B}}(f)$. Calculer $\chi_f(X)$ et conclure par 2) c).

§ XV.6 SUITES DÉFINIES PAR UNE RELATION DE RÉCURRENCE LINÉAIRE

Nous allons compléter ci-dessous l'étude, amorcée dans l'exemple 4 du § XV.3, du K -ev \mathcal{E} des suites $(u_n)_{n \in \mathbb{N}}$ à valeurs dans K vérifiant la relation

$$(1) \quad (\forall n \in \mathbb{N}) \quad u_{n+k} - \zeta_{k-1} u_{n+k-1} - \cdots - \zeta_0 u_n = 0$$

où $k \in \mathbb{N}^*$ et $(\zeta_0, \zeta_1, \dots, \zeta_{k-1}) \in K^n$ sont donnés, avec $\zeta_0 \neq 0$. Nous avons associé à ces données, le *polynôme caractéristique* de la relation (1) :

$$P(X) = X^k - \zeta_{k-1} X^{k-1} - \cdots - \zeta_0.$$

Rappelons que \mathcal{E} est de dimension k , l'application $\Phi : K^k \rightarrow \mathcal{E}$ associant à $(a_1, \dots, a_k) \in K^k$ l'unique suite $(u_n) \in \mathcal{E}$ pour laquelle

$$u_0 = a_1, \dots, u_{k-1} = a_k$$

étant un isomorphisme de K -ev. Enfin, on a un remarquable endomorphisme τ de \mathcal{E} qui associe, à la suite (u_n) , la suite (v_n) définie par $v_n = u_{n+1}$ pour tout $n \in \mathbb{N}$. Nous avons vu que $(-1)^k P(X)$ est précisément le *polynôme caractéristique* $\chi_\tau(X)$.

La définition même de \mathcal{E} nous montre que : $\mathcal{E} = \text{Ker}(P(T))$, où $T \in \text{Hom}_K(K^{\mathbb{N}})$ envoie chaque suite (u_n) sur la suite (v_n) avec $(\forall n) v_n = u_{n+1}$. On a donc : $\tau = T|_{\mathcal{E}}$, et la relation $P(\tau) = 0$ est ici évidente, sans qu'il soit besoin d'utiliser le théorème XV.4.2.

Etude de \mathcal{E} lorsque $P(X)$ est dissocié dans $K[X]$

Supposons $P(X)$ dissocié sur K :

$$(2) \quad P(X) = \prod_{i=1}^p (X - \lambda_i)^{\alpha_i}$$

(les $\alpha_i \geq 1$, $p \geq 1$ et les λ_i distincts).

Le lemme des noyaux (théorème XV.5.1), appliqué successivement à T et à τ , montre que :

$$(3) \quad \mathcal{E} = \bigoplus_{i=1}^p \text{Ker}(T - \lambda_i \text{Id}_{K^{\mathbb{N}}})^{\alpha_i} = \bigoplus_{i=1}^p \text{Ker}(\tau - \lambda_i \text{Id}_{\mathcal{E}})^{\alpha_i}.$$

Les inclusions

$$\text{Ker}(\tau - \lambda_i \text{Id}_{\mathcal{E}})^{\alpha_i} \subset \text{Ker}(T - \lambda_i \text{Id}_{K^{\mathbb{N}}})^{\alpha_i} \quad (1 \leq i \leq p)$$

sont immédiates, donc on déduit de (3) que, pour tout i

$$\text{Ker } (T - \lambda_i \text{Id}_{K^{\mathbb{N}}})^{\alpha_i} = \text{Ker } (\tau - \lambda_i \text{Id}_{\mathcal{E}})^{\alpha_i}.$$

Notons \mathcal{N}_i ces sous-espaces. Puisque $P(X) = (-1)^k \chi_\tau(X)$, ce sont les *sous-espaces caractéristiques* de τ , donc le théorème XV.5.4 montre

$$(4) \quad (\forall i) \quad \dim(\mathcal{N}_i) = \alpha_i.$$

Notons Δ l'opérateur de différence de $K[X] : Q(X) \mapsto Q(X+1) - Q(X)$. Pour tout $n \in \mathbb{N}^*$, on a $\Delta(K_n[X]) \subset K_{n-1}[X]$; et $\Delta(K) = \{0\}$. Soit

$$i \in \llbracket 1, p \rrbracket \quad \text{et} \quad Q \in K_{\alpha_i-1}[X].$$

La suite $u_{(Q)} : n \mapsto u_n = \lambda_i^n Q(n)$ se transforme, par l'endomorphisme $f_i^q = (T - \lambda_i \text{Id}_{K^{\mathbb{N}}})^q$ de $K^{\mathbb{N}}$, en la suite $n \mapsto \lambda_i^{n+q} [\Delta^q(Q)](n)$ pour tout $q \in \mathbb{N}^*$, et en particulier $f_i^{\alpha_i}(u_{(Q)}) = 0$. On a donc ce résultat :

Pour tout polynôme $Q \in K_{\alpha_i-1}[X]$, la suite $u_{(Q)} : n \mapsto \lambda_i^n Q(n)$ ⁽¹⁾ vérifie la relation de récurrence (1). Cette suite appartient à \mathcal{N}_i .

Supposons maintenant le corps K de *caractéristique nulle*. Fixons $i \in \llbracket 1, p \rrbracket$: l'application $Q \mapsto u_{(Q)}$, $K_{\alpha_i-1}[X] \rightarrow \mathcal{N}_i$ est linéaire. Son noyau est formé des $Q \in K_{\alpha_i-1}[X]$ tels que $Q(n) = 0$ pour tout $n \in \mathbb{N}$ (car $\lambda_i \neq 0$), donc ce noyau est nul (cf. théorème VII.4.2). Et puisque $\dim(\mathcal{N}_i) = \dim K_{\alpha_i-1}[X] = \alpha_i$, il en résulte que $Q \mapsto u_{(Q)}$ est bijective de $K_{\alpha_i-1}[X]$ sur \mathcal{N}_i . Nous avons donc obtenu dans ce cas *toutes* les solutions de (1). Résumons l'ensemble de ces résultats :

THÉORÈME XV.6.1

*Si le corps K est de **caractéristique nulle**, et si le polynôme caractéristique de la relation (1) est **dissocié** sous la forme (2), le K -ev des suites solutions de (1) est exactement*

$$\mathcal{E} = \bigoplus_{i=1}^p \mathcal{N}_i$$

où, pour chaque $i \in \llbracket 1, p \rrbracket$, \mathcal{N}_i est le K -ev formé des suites $u_{(Q)} : n \mapsto \lambda_i^n Q(n)$, avec $Q \in K_{\alpha_i-1}[X]$. Chaque \mathcal{N}_i est de dimension α_i , l'application $Q \mapsto u_{(Q)}$ étant une bijection linéaire de $K_{\alpha_i-1}[X]$ sur \mathcal{N}_i .

⁽¹⁾ $Q(n)$ est une abréviation pour $Q(n.1_K)$.

Donnons-nous $(a_1, \dots, a_k) = \mathbf{a} \in K^k$ en conservant les notations et hypothèses du théorème XV.6.1, et cherchons la suite $u = (u_n) \in \mathcal{E}$ telle que

$$(5) \quad u_0 = a_1, \dots, u_{k-1} = a_k.$$

Pour cela, prenons pour inconnues les coefficients (au nombre total de k) des polynômes

$$Q_1 \in K_{\alpha_1-1}[X], \dots, Q_p \in K_{\alpha_p-1}[X],$$

et écrivons que $u = u_{(Q_1)} + u_{(Q_2)} + \dots + u_{(Q_p)}$.

Les conditions (5) constituent, relativement à ces inconnues, un système linéaire de k équations dont on sait à l'avance qu'il a une solution unique : c'est donc un système de Cramer et sa résolution donne Q_1, \dots, Q_p , d'où la suite u cherchée.

Exemple 1 : Le corps de base est \mathbb{C} . On donne $\lambda \in \mathbb{C}$, $\lambda \notin \{0, 1\}$, et un triplet $(a_1, a_2, a_3) \in \mathbb{C}^3$. Trouver la suite (u_n) telle que

$$(6) \quad \begin{aligned} u_0 &= a_1, u_1 = a_2, u_2 = a_3 \quad \text{et} \quad (\forall n \geq 3) \\ u_n &= (2\lambda + 1)u_{n-1} - (\lambda^2 + 2\lambda)u_{n-2} + \lambda^2 u_{n-3}. \end{aligned}$$

Solution : le polynôme caractéristique est $P(X) = (X - \lambda)^2 (X - 1)$. La suite cherchée est donc, d'après le théorème XV.6.1, de la forme

$$u_n = (x_1 + x_2 n) \lambda^n + x_3 \quad (n \in \mathbb{N}).$$

Les conditions (6) équivalent au système linéaire des 3 équations aux inconnues x_1, x_2, x_3 :

$$(7) \quad \begin{cases} x_1 + x_3 = a_1 \\ \lambda x_1 + \lambda x_2 + x_3 = a_2 \\ \lambda^2 x_1 + 2\lambda^2 x_2 + x_3 = a_3 \end{cases} \quad \text{dont la résolution est immédiate.}$$

On trouve, en utilisant les formules de Cramer,

$$\begin{aligned} x_1 &= \frac{(1 - 2\lambda)a_1 + 2\lambda a_2 - a_3}{(\lambda - 1)^2}, \quad x_2 = \frac{\lambda a_1 - (\lambda + 1)a_2 + a_3}{\lambda(\lambda - 1)}, \\ x_3 &= \frac{\lambda^2 a_1 - 2\lambda a_2 + a_3}{(\lambda - 1)^2}. \end{aligned}$$

Pour $\lambda = 1$, la relation (6) devient $u_n = 3u_{n-1} - 3u_{n-2} + u_{n-3}$ et le polynôme caractéristique $(X - 1)^3$. La solution doit donc être cherchée sous la forme $x_1 + x_2 n + x_3 n^2$ et le lecteur achèvera de la déter

Remarque 1 : Le théorème XV.6.1 tombe en défaut si K n'est pas de caractéristique 0.

Utilisation d'une fraction rationnelle

Le corps K étant à nouveau quelconque, cherchons la suite $u = (u_n)$ qui vérifie (1) et

$$u_0 = a_1, u_1 = a_2, \dots, u_{k-1} = a_k \quad ((a_1, \dots, a_k) \in K^k \text{ donné}).$$

Posons $c_i = \zeta_{k-i}$ ($1 \leq i \leq k$), et soit

$$S = \sum_{n \geq 0} u_n X^n \in K[[X]], \quad \text{et} \quad Q(X) = 1 - c_1 X - c_2 X^2 - \dots - c_k X^k.$$

En calculant le produit $QS = \sum_{n \geq 0} v_n X^n$, on voit que la relation (1) équivaut à la propriété que $v_n = 0$ pour tout $n \geq k$ (d'où $QS \in K[X]$).

D'ailleurs $\text{val}(Q) = 1$, donc $Q(X)$ est inversible dans $K[[X]]$, et finalement (1) signifie que S est le développement en série formelle de la fraction rationnelle

$$(8) \quad F = \frac{QS}{Q} = \frac{A_0 + A_1 X + \dots + A_{k-1} X^{k-1}}{Q(X)},$$

$F \in K(X)$ (cf. théorème VIII.5.2), avec

$$A_0 = a_1, A_1 = a_2 - c_1 a_1, \dots, A_{k-1} = a_k - c_1 a_{k-1} - \dots - c_{k-1} a_1.$$

On peut (mais le calcul devient lourd dès que $k \geq 3$) écrire directement F sous la forme

$$(A_0 + A_1 X + \dots + A_{k-1} X^{k-1}) \sum_{n \geq 0} (c_1 X + \dots + c_k X^k)^n,$$

en regroupant par puissances égales de X après développement de chaque $(c_1 X + \dots + c_k X^k)^n$ par la formule du multinôme. Mais il est évident que si $P(X)$ est dissocié dans $K[X]$ sous la forme (2), d'où

$$Q(X) = \prod_{i=1}^p (1 - \lambda_i X)^{\alpha_i},$$

il est plus rapide de décomposer F en éléments simples sur K :

$$F = \sum_{i=1}^p \left(\sum_{j=1}^{\alpha_i} B_{i,j} (1 - \lambda_i X)^{-j} \right),$$

et développer chaque $(1 - \lambda_i X)^{-j}$ (cf. théorème VIII.5.3), d'où :

$$(1 - \lambda_i X)^{-j} = 1 + \sum_{q \geq 1} \binom{q+j-1}{q} \lambda_i^q X^q$$

et

$$F(X) = \sum_{q \geq 0} X^q \left[\sum_{i=1}^p \lambda_i^q \left\{ \sum_{j=1}^{\alpha_i} B_{i,j} \binom{q+j-1}{q} \right\} \right]$$

et donc

$$(9) \quad (\forall n) \quad u_n = \sum_{i=1}^p \lambda_i^n \left\{ \sum_{j=1}^{\alpha_i} \binom{n+j-1}{n} B_{ij} \right\}.$$

Si K est de caractéristique 0, chaque fonction $n \mapsto \binom{n+j-1}{n}$ au second membre de (9) est *polynomiale* de degré $j-1$, et à partir de là on peut donner une nouvelle preuve du théorème XV.6.1.

Remarque 2 : Dans des cas concrets (sauf si $k \leq 2$ comme dans l'exemple 1 du § VIII.5) cette méthode est rarement aussi avantageuse que la précédente. Cependant elle présente un grand intérêt théorique : par exemple, lorsque K est de caractéristique > 0 , le second membre de (9) nous donne (lorsque P est dissocié sur K) la forme générale des solutions de (1), qui restait à étudier.

Exercice 1 : Le corps de base est \mathbb{C} . Expliciter u_n sachant que $(u_0, u_1, u_2) = (a, b, c) \in \mathbb{C}^3$ donné et

$$(\forall n \geq 3) \quad u_n = 3u_{n-1} - 2u_{n-3}.$$

Comparer, sur cet exemple, la méthode du système de Cramer et celle des fractions rationnelles.

Exercice 2 : Soit $p \in \mathbb{N}$ ($p \geq 2$) et (u_n) une suite de \mathbb{R}_+^* définie par u_0, u_1, \dots, u_{p-1} donnés > 0 et $(\forall n \geq p) \quad u_n = (u_{n-1} u_{n-2} \dots u_{n-p})^{1/p}$. Étudier la suite (u_n) pour $n \rightarrow \infty$.

Exercice 3 : Le corps de base est \mathbb{C} . On donne la relation de récurrence

$$(*) \quad (\forall n \geq 4) \quad u_n = 4u_{n-1} - 4u_{n-2} + u_{n-4}.$$

Trouver tous les systèmes de valeurs de $(u_0, u_1, u_2, u_3) \in \mathbb{C}^4$.

- a) pour que la suite (u_n) définie par ces conditions initiales et la relation $(*)$ soit *bornée* ;
- b) pour que cette même suite tende vers 0 quand $n \rightarrow \infty$ (il est inutile d'expliciter u_n en fonction de (u_0, u_1, u_2, u_3)).

Exercice 4 : Trouver toutes les suites de nombres complexes telles que

$$(\forall n \geq 0) \quad u_{n+5} = u_{n+4} + 5u_{n+3} - u_{n+2} - 8u_{n+1} - 4u_n.$$

Exercice 5 : On donne u_1 et v_1 . Étudier les suites définies pour $n \geq 1$ par

$$u_{n+1} = \frac{u_1 + \dots + u_n}{n} \quad \text{et par} \quad v_{n+1} = \frac{2}{n} (v_1 + \dots + v_n).$$

Exercice 6 : On se place dans \mathbb{R}_+^* et on donne $u_0 > 0$, $u_1 > 0$. Calculer u_n défini, pour tout $n \geq 2$ par $\frac{2}{u_n} = \frac{1}{u_{n-1}} + \frac{1}{u_{n-2}}$. Quelle est la limite de u_n ?

Exercice 7 : On donne $u_0 = 1$ et $u_1 = 2$. Etudier la suite réelle définie par la relation

$$(\forall n \geq 1) \quad u_{n+1} - 2u_n + u_{n-1} = n.$$

Exercice 8 : On donne $u_0 = 0$ et $v_0 = 1$. Etudier le couple de suites réelles définies par

$$(\forall n \in \mathbb{N}^*) \quad u_n = \frac{2u_{n-1} + 3v_{n-1}}{5}, \quad v_n = \frac{3u_{n-1} + 2v_{n-1}}{5}.$$

Montrer qu'elles ont une limite commune.

Exercice 9 : On désigne par D_n le nombre de *dérangements* de \mathfrak{S}_n (i.e. le nombre de permutations σ de $\llbracket 1, n \rrbracket$ telles que $(\forall i) \sigma(i) \neq i$). Montrer que $D_1 = 0$, $D_2 = 1$, et que $D_{n+1} = n(D_n + D_{n-1})$ pour tout $n \geq 2$. Si on pose $u_n = \frac{D_n}{n!}$ montrer que la suite (u_n) vérifie la relation de récurrence $nu_n = (n-1)u_{n-1} + u_{n-2}$.

En calculant $v_n = u_n - u_{n-1}$, en déduire l'expression exacte de u_n , puis de D_n .

Exercice 10 : Trouver toutes les applications (u, v, w, t) de \mathbb{N} dans \mathbb{C}^4 vérifiant les relations suivantes :

$$\begin{cases} u_{n+1} = -5u_n - 3v_n - 2w_n + 4t_n \\ v_{n+1} = 2u_n + w_n - t_n \\ w_{n+1} = 10u_n + 7v_n + 4w_n - 9t_n \\ t_{n+1} = 2u_n + w_n \end{cases}$$

Exercice 11 : On suppose $a \in \mathbb{C}$, $a \neq 1$. Exprimer $\sum_{p=1}^n p^2 a^p$, où $n \in \mathbb{N}^*$, par une formule d'échelon 0.

Exercice 12 : Le corps K est supposé de caractéristique $p > 0$. Dédurre de la formule (9), et de l'étude qui précède le théorème XV.6.1, que les fonctions

$$\theta_j : \mathbb{N} \rightarrow K, \quad n \mapsto \binom{n+j-1}{n} \cdot 1_K$$

sont linéairement indépendantes (on pourra aussi essayer de prouver directement cette propriété). Quelle est la dimension du K -ev engendré par les fonctions $\mathbb{N} \rightarrow K$, $n \mapsto Q(n) \lambda^n$ où $\lambda \in K^*$ est donné, lorsque Q parcourt $K_{\alpha-1}[X]$, avec $\alpha \in \mathbb{N}^*$? Conclusion si $\alpha > p$? (N.B. les fonctions θ_j sont parfois appelées *pseudopolynômes* par les combinatoriens).

Chapitre XVI

COMPLÉMENT : RÉDUCTION DE JORDAN

Dans tout ce chapitre, K désigne un corps commutatif fixé, et tous les K -ev considérés sont de dimension finie.

§ XVI.1 ÉTUDE DES ENDOMORPHISMES NILPOTENTS

Soit E un K -ev de dimension $n \geq 1$, et $u \in \text{Hom}_K(E)$ un endomorphisme nilpotent, de période r . Au § XV.4, nous avons vu que $\chi_u(X) = (-1)^n X^n$ et $Q_u(X) = X^r$. Nous allons caractériser la classe de similitude de u . Nous utiliserons pour cela la notion suivante :

DÉFINITION XVI.1.1

Soit $n \in \mathbb{N}^*$ et $\lambda \in K$. On appelle **matrice de Jordan** ⁽¹⁾ $J_\lambda(n)$ la (n, n) -matrice à coefficients dans K de terme général $a_{i,j}$ $((i, j) \in \llbracket 1, n \rrbracket^2)$ telle que $a_{i,i} = \lambda$ pour tout i ; $a_{i,i-1} = 1$ pour $i \geq 2$ et $a_{i,j} = 0$ pour tous les autres couples (i, j) . La matrice $J_0(n)$ sera simplement notée $J(n)$ et appelée **matrice nilpotente de Jordan** d'ordre n .

On a donc $J(1) = [0]$, $J_\lambda(1) = [\lambda]$, et pour $n \geq 2$:

$$J_\lambda(n) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix} = \lambda I_n + J(n).$$

Cas où $r = n$

Si la période r du nilpotent u est égale à la dimension de E , on a :

THÉORÈME XVI.1.1

Soit E un K -ev de dimension $n \geq 1$ et $u \in \text{Hom}_K(E)$ un endomorphisme nilpotent de période r . Les propriétés suivantes sont équivalentes :

- (I) $r = n$.
- (II) Il existe une base ordonnée $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $\text{Mat}_{\mathcal{B}}(u) = J(n)$.
- (III) Il existe $a \in E$ tel que $(u^{n-1}(a), \dots, u(a), a)$ soit une base de E .

⁽¹⁾ Marie Ennemond Camille Jordan (1838-1922), ingénieur des Mines et mathématicien français spécialiste de la théorie des groupes, auteur du *Traité des substitutions et des équations algébriques* (1870).

Démonstration :

Si $r = n$, prenons $a \in E$ tel que $u^{n-1}(a) \neq 0$. Alors les vecteurs $(u^{n-1}(a), u^{n-2}(a), \dots, u(a), a)$ sont linéairement indépendants (cf. proposition XV.4.1). Donc ils forment, dans cet ordre, une base \mathcal{B} de E , et il est clair que $\text{Mat}_{\mathcal{B}}(u) = J(n)$, d'où (II) et (III).

Réciproquement, si (III) est vraie, on a $u^{n-1}(a) \neq 0$, d'où (I) ; et si (II) est vraie, pour un choix déterminé de base $\mathcal{B} = (e_1, \dots, e_n)$ vérifiant $\text{Mat}_{\mathcal{B}}(u) = J(n)$, en posant $a = e_n$, on constate que $u(a) = e_{n-1}, \dots, u^{n-1}(a) = e_1$, d'où (III). ■

Remarque 1 : Si u vérifie les conditions équivalentes du théorème XVI.1.1, en choisissant une base qui vérifie (III), on voit que $\text{Ker}(u) = Ke_1 = Ku^{n-1}(a)$, donc $\text{rg}(u) = n - 1$.

Un endomorphisme nilpotent de période r induit évidemment, sur tout sous-espace u -stable, un endomorphisme nilpotent, de période $\leq r$.

DÉFINITION XVI.1.2

Soit u un endomorphisme nilpotent du K -ev E de dimension $n \geq 1$. Un sous- K -ev non nul F de E est dit **u -monogène** ssi : il est u -stable, et l'endomorphisme $u|_F$ est de période $\dim(F)$ (i.e. vérifie les conditions équivalentes du théorème XVI.1.1).

Avec cette définition, dire que u est de période n revient à dire que E est u -monogène.

Soit alors $x \in E \setminus \{0\}$. Le sous-espace engendré par les $(u^k(x))_{0 \leq k \leq r-1}$ est u -monogène. On le note $K[u] \cdot x$, en conformité avec le § XV.4. Il est clair qu'on obtient ainsi tous les sous-espaces u -monogènes de E .

Décomposition en somme directe d'espaces u -monogènes

THÉORÈME XVI.1.2

Soit u un endomorphisme nilpotent de période r dans le K -ev E ($\dim(E) = n \geq 1$). Alors E est somme directe interne de sous-espaces u -monogènes.

Démonstration :

Raisonnons par récurrence sur n . La propriété est évidente pour $n = 1$ ($u(E) = \{0\} \subset E$ et u est de période $1 = \dim(E)$). Supposons-la vraie en toute dimension $< n$, avec $n \geq 2$. Soit $v = {}^t u \in \text{Hom}_K(E^*)$; v est nilpotent, de période r comme u (car $(\forall k \in \mathbb{N}) v^k = {}^t(u^k)$), d'où $v^{r-1} \neq 0$. Prenons $\varphi \in E^*$ telle que $v^{r-1}(\varphi) \neq 0$, i.e. $\varphi \circ u^{r-1} \neq 0$, puis, $x \in E$ tel que $\varphi(u^{r-1}(x)) \neq 0$. Posons $F = K[u] \cdot x$, et $G = {}^0(K[v] \cdot \varphi)$. Le sous-espace F est u -monogène non nul, car $x \neq 0$. Si $F = E$, alors E est u -monogène et il n'y a plus rien à prouver. Plaçons-nous donc dans le cas où $F \neq E$.

Montrons d'abord que G est u -stable :

Puisque $u^{r-1}(x) \neq 0$, les vecteurs $(x, u(x), \dots, u^{r-1}(x))$ sont linéairement indépendants dans E (cf. la preuve de la proposition XV.4.1), et forment donc une base de F , d'où $\dim(F) = r$. Soit $z \in E$: la définition de G montre que

$$z \in G \text{ ssi } (\forall k \in \mathbb{N}) \varphi(u^k(z)) = 0.$$

Donc, si $z \in G$, il s'ensuit $\varphi(u^k(u(z))) = \varphi(u^{k+1}(z)) = 0$, ce qui entraîne $u(z) \in G$ et prouve bien que G est u -stable.

Montrons ensuite que $F \cap G = \{0\}$:

Soit $y \in F \cap G$. y s'écrit de manière unique sous la forme

$$y = \lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{r-1} u^{r-1}(x),$$

avec $(\lambda_0, \dots, \lambda_{r-1}) \in K^r$, car $y \in F$. Si l'on avait $y \neq 0$, on pourrait définir l'entier

$$k = \text{Min } \{i \in \llbracket 0, r-1 \rrbracket \mid \lambda_i \neq 0\}$$

et puisque $y \in G$, la forme linéaire $v^{r-1-k}(\varphi) = \varphi \circ u^{r-1-k}$ s'annulerait sur y , d'où $\varphi \circ u^{r-1}(x) = 0$, ce qui contredirait le choix de x .

Donc $y = 0$ et $F \cap G = \{0\}$.

Montrons enfin que $F \oplus G = E$:

On a vu que $\dim(F) = r$. On verrait de la même manière que $\dim(K[v] \cdot \varphi) = r$, car les vecteurs $(\varphi, v(\varphi), \dots, v^{r-1}(\varphi))$ forment une base du K -ev $K[v] \cdot \varphi$ (à cause de $v^{r-1}(\varphi) \neq 0$). Par dualité (théorème XII.2.5), il s'ensuit que

$$\dim(G) = n - r = n - \dim(F),$$

donc, puisque $F \cap G = \{0\}$, cela entraîne bien $F \oplus G = E$.

Fin de la démonstration :

Puisque $\dim(G) = n - r < n$, l'hypothèse de récurrence s'applique à l'endomorphisme nilpotent $u|_G$ de G . En écrivant G comme somme directe interne de sous-espaces $u|_G$ -monogènes (donc u -monogènes), on voit que $E = F \oplus G$ est aussi somme directe interne de tels sous-espaces. ■

Facteurs invariants

Conservons les hypothèses et notations du théorème XVI.1.2. Nous allons maintenant analyser de plus près une décomposition de E sous la forme

$$(1) \quad E = E_1 \oplus E_2 \oplus \cdots \oplus E_k,$$

où les E_i sont des sous-espaces monogènes non nuls, pour lesquels nous poserons $\dim(E_i) = r_i$, la numérotation étant choisie pour que $r_1 \geq r_2 \geq \cdots \geq r_k \geq 1$.

Introduisons d'abord quelques notations : $u_i = u|_{E_i}$ ($1 \leq i \leq k$) ; $N_j = \text{Ker}(u^j)$ ($j \in \mathbb{N}$) ; $d_j = \dim(N_j)$.

La période de u est r , celle de u_i est r_i pour $1 \leq i \leq k$, donc $r = \text{Max}_{i=1}^k (r_i)$, d'où $r_1 = r$. La propriété $N_j = N_{j+1}$ est héréditaire, donc on a :

$$N_0 = \{0\} \subsetneq \text{Ker}(u) = N_1 \subsetneq N_2 \subsetneq \cdots \subsetneq N_r = E \quad (\text{car } u^r = 0 \text{ et } u^{r-1} \neq 0).$$

Il est facile de voir que $k = d_1$: en effet, le rang de chaque u_i est $r_i - 1$ (cf. remarque 1) ; étant donné que $u = \bigoplus_{i=1}^r u_i$, on en déduit $\text{rg}(u) = \sum_{i=1}^k (r_i - 1) = n - k$, d'où $k = \dim(\text{Ker}(u)) = d_1$.

La fonction décroissante $\rho : \llbracket 1, k \rrbracket \rightarrow \llbracket 1, r \rrbracket$, $i \mapsto r_i$ est définie de manière unique par la suite (ν_1, \dots, ν_r) , où $\nu_\alpha = \text{card}(\rho^{-1}(\alpha))$ pour tout $\alpha \in \llbracket 1, r \rrbracket$. A l'aide de ρ , nous définissons ainsi la fonction $\psi : \llbracket 0, r+1 \rrbracket \rightarrow \mathbb{N} : \psi(0) = k$, $\psi(r+1) = 0$, $\psi(j) = \nu_j + \nu_{j+1} + \dots + \nu_r$, pour $j \in \llbracket 1, r \rrbracket$; ψ est décroissante. La connaissance de ρ équivaut à celle de ψ , car on voit que la valeur de ρ sur chaque intervalle $\llbracket \psi(j+1), \psi(j) \rrbracket$ est j , pour $j \in \llbracket 1, r \rrbracket$.

Calcul de ψ en fonction des d_j .

Pour chaque $i \in \llbracket 1, k \rrbracket$, considérons une base \mathcal{B}_i de E_i de la forme

$$\mathcal{B}_i = (u^{r_i-1}(x_i), u^{r_i-2}(x_i), \dots, u(x_i), x_i),$$

avec un x_i convenable. En juxtaposant ces bases dans l'ordre $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$, on obtient une base \mathcal{B} de E . Etudions l'action de u^j sur les vecteurs de cette base pour $j \in \llbracket 1, r \rrbracket$.

Si $j = 1$, les premiers vecteurs des $\mathcal{B}_i : u^{r_1-1}(x_1), u^{r_2-1}(x_2), \dots, u^{r_k-1}(x_k)$ appartiennent à $\text{Ker}(u)$. Quant aux autres vecteurs de \mathcal{B} , leurs images par u sont des vecteurs indépendants : on retrouve ainsi le fait que $d_1 = \dim(\text{Ker}(u)) = k$.

Si $j \geq 2$, l'image par u^j des vecteurs de \mathcal{B}_l pour $l > \psi(j)$, et des vecteurs $u^{r_l-1}(x_l), \dots, u^{r_l-j}(x_l)$ pour $1 \leq l \leq \psi(j)$, est 0. Le nombre total de ces vecteurs est

$$A_j = j\psi(j) + \sum_{l=1+\psi(j)}^k r_l. \text{ Quant aux images par } u^j \text{ des vecteurs } u^{r_l-j-1}(x_l), \dots,$$

$u(x_l), x_l$ pour $1 \leq l \leq \psi(j)$, ce sont $u^{r_l-1}(x_l), \dots, u^{j+1}(x_l), u^j(x_l)$ qui forment une famille libre, ce qui fait en tout $B_j = \sum_{l=1}^{\psi(j)} (r_l - j) = r_1 + \dots + r_{\psi(j)} - j\psi(j)$ vecteurs

indépendants dans $\text{Im}(u^j)$.

Or, $A_j + B_j = \sum_{l=1}^k r_l = n = \dim(E)$. En exploitant la formule du rang pour u^j , on en déduit que $\text{Ker}(u^j)$ admet pour base la réunion des vecteurs des bases \mathcal{B}_l , $l > \psi(j)$, et des vecteurs $(u^{r_l-1}(x_l), \dots, u^{r_l-j}(x_l))$, pour l décrivant $\llbracket 1, \psi(j) \rrbracket$, donc $\dim(\text{Ker}(u^j)) = d_j = A_j$; quant à $\text{Im}(u^j)$, il admet pour base la réunion de tous les vecteurs $(u^{r_l-1}(x_l), \dots, u^j(x_l))$ pour $l \in \llbracket 1, \psi(j) \rrbracket$, et $\dim(\text{Im}(u^j)) = B_j$.

Remarque 2 : On verrait de même que les vecteurs (x_l) , $l > \psi(j+1)$ forment une base d'un supplémentaire de $u(N_{j+1})$ dans N_j .

De tout ce qui précède résulte la relation :

$$(\forall j \in \llbracket 1, r \rrbracket) \quad d_j = A_j = j\psi(j) + \sum_{\psi(j) < l \leq k} r_l$$

(pour $j = 1$ on fait la convention que $\sum_{\psi(j) < l \leq k} r_l = 0$).

Si l'on tient compte du fait, vu plus haut, que $r_l = j-1$ pour $\psi(j) < l \leq \psi(j-1)$, on obtient, pour tout $j \in \llbracket 1, r \rrbracket$:

$$\begin{aligned} d_j - d_{j-1} &= j\psi(j) - (j-1)\psi(j-1) + \sum_{\psi(j) < l \leq \psi(j-1)} r_l = \\ &= j\psi(j) - (j-1)\psi(j-1) + (j-1)[\psi(j-1) - \psi(j)] \end{aligned}$$

Nous avons donc exprimé la fonction ψ à l'aide de la suite (d_0, d_1, \dots, d_r) par la formule

$$(2) \quad \psi(j) = d_j - d_{j-1} \quad (1 \leq j \leq r)$$

et, comme nous l'avons signalé en introduisant ψ , cela permet d'exprimer la fonction ρ à l'aide de la seule suite (d_0, d_1, \dots, d_r) . Donc cette fonction ρ ne dépend que de u et non du choix des E_i dans (1).

Résumons le résultat obtenu :

THÉOREME XVI.1.3

Soit u un endomorphisme nilpotent de période r du K -ev E de dimension $n \geq 1$. Soit $d_j = \dim(\text{Ker}(u^j))$ pour $j \in \mathbb{N}$, et $\psi : \llbracket 0, r+1 \rrbracket \rightarrow \mathbb{N}$ la fonction telle que $\psi(0) = d_1$, $\psi(r+1) = 0$ et $\psi(j) = d_j - d_{j-1}$ pour $j \in \llbracket 1, r \rrbracket$. Alors ψ est décroissante.
Soit $\rho : \llbracket 1, d_1 \rrbracket \rightarrow \llbracket 1, r \rrbracket$ la fonction qui vaut j sur $\llbracket \psi(j+1), \psi(j) \rrbracket$ pour tout $j \in \llbracket 1, r \rrbracket$. Pour toute décomposition de E sous la forme $E = E_1 \oplus E_2 \oplus \dots \oplus E_k$, où les E_i sont non nuls, u -monogènes, et tels que $\dim(E_1) \geq \dim(E_2) \geq \dots \geq \dim(E_k)$, on a : $k = d_1$ et $\dim(E_i) = \rho(i)$ pour $i \in \llbracket 1, k \rrbracket$.

DÉFINITION XVI.1.3

Avec les notations du théorème XVI.1.3, la suite $X^{r_1}, X^{r_2}, \dots, X^{r_k}$, où $r_i = \rho(i)$ pour tout i , s'appelle suite des **facteurs invariants** de u .

Dans toute décomposition du type (1) de E , en notant $u_i = u|_{E_i}$ les facteurs invariants X^{r_1}, \dots, X^{r_k} sont les *polynômes minimaux* respectifs de u_1, u_2, \dots, u_k et $X^{r_1} = X^r$ est le *polynôme minimal* de u .

Pour toute base \mathcal{B} , construite comme indiqué dans l'étude précédente, à partir d'une décomposition (1) de E , on a : $\text{Mat}_{\mathcal{B}}(u_i) = J(r_i)$ ($1 \leq i \leq k$), et $\text{Mat}_{\mathcal{B}}(u)$ est diagonale par blocs de dimensions r_1, r_2, \dots, r_k :

$$\text{Mat}_{\mathcal{B}}(u) = \text{Diag}(J(r_1), J(r_2), \dots, J(r_k)).$$

Comme d'habitude, on définit les **facteurs invariants d'une matrice nilpotente** $M \in \mathfrak{M}_n(K)$ comme étant les facteurs invariants de l'endomorphisme u_M de K^n que définit M dans la base canonique de K^n .

Exemple 1 : Si l'on se donne à l'avance des entiers k, r_1, r_2, \dots, r_k , avec $k \geq 1$, $r_1 \geq r_2 \geq \dots \geq r_k \geq 1$ et $r_1 + r_2 + \dots + r_k = n$, les facteurs invariants de la matrice

$$(2) \quad M = \text{Diag}(J(r_1), \dots, J(r_k)) \in \mathfrak{M}_n(K)$$

sont $X^{r_1}, X^{r_2}, \dots, X^{r_k}$. En effet, notons (e_1, \dots, e_n) la base canonique \mathcal{B} de K^n . Soit $x_1 = e_{r_1}, x_2 = e_{r_1+r_2}, \dots, x_k = e_{r_1+r_2+\dots+r_k}$. On vérifie immédiatement que \mathcal{B} s'obtient en juxtaposant les suites $\mathcal{B}_1, \dots, \mathcal{B}_k$, où $\mathcal{B}_i = (u_M^{r_i-1}(x_i), \dots, u_M(x_i), x_i)$ pour tout i .

Il n'y a aucune difficulté, à partir de là, à montrer que la fonction ρ construite à partir des entiers $\dim (\text{Ker } (u_M^i))_{0 \leq i \leq r_1}$ est précisément la suite $i \mapsto r_i$ ($1 \leq i \leq k$). (On pose $\xi(i) = \text{Max } \{j \mid r_j \geq i\}$ pour $i \in \llbracket 1, r \rrbracket$, et on contrôle que

$$\dim (\text{Ker } (u_M^j)) = j\xi(j) + \sum_{\psi(j) < l \leq k} r_l$$

pour tout $j \in \llbracket 1, r_1 \rrbracket$, d'où

$$\xi(j) = \dim (\text{Ker } (u_M^j)) - \dim (\text{Ker } (u_M^{j-1})),$$

donc ξ est la fonction ψ associée au nilpotent u_M , donc ρ est la fonction $i \mapsto r_i$.

Classes de similitude de nilpotents

THÉORÈME XVI.1.4

Soit E un K -ev de dimension $n \geq 1$ et u, v deux endomorphismes **nilpotents** de E . Pour que u et v soient **semblables**, il faut et il suffit qu'ils aient les **mêmes facteurs invariants**.
De même, pour que deux matrices M et N de $\mathfrak{M}_n(K)$ **nilpotentes** soient **semblables**, il faut et il suffit qu'elles aient les **mêmes facteurs invariants**.

Démonstration :

Si u et v ont mêmes facteurs invariants : X^{r_1}, \dots, X^{r_k} , on peut trouver des bases ordonnées \mathcal{B} et \mathcal{C} de E telles que

$$\text{Mat}_{\mathcal{B}}(u) = \text{Mat}_{\mathcal{C}}(v) = \text{Diag } (J(r_1), \dots, J(r_k)).$$

Donc u et v sont semblables (si φ est l'automorphisme qui envoie \mathcal{B} sur \mathcal{C} , on a $v = \varphi u \varphi^{-1}$).

Réciproquement, si u et v sont semblables, soit $d_j(u) = \dim (\text{Ker } (u^j))$, $d_j(v) = \dim (\text{Ker } (v^j))$ pour $j \in \mathbb{N}$, et soit $\varphi \in \text{GL}_K(E)$ tel que $v = \varphi u \varphi^{-1}$. On a $v^j = \varphi u^j \varphi^{-1}$ pour tout j , d'où $\text{Ker } (v^j) = \varphi (\text{Ker } (u^j))$, et par suite $d_j(u) = d_j(v)$. Le théorème XVI.1.3 et la définition XVI.1.3 montrent alors que u et v ont mêmes facteurs invariants.

La seconde assertion de l'énoncé est la forme matricielle de la première et lui est donc équivalente. ■

En vertu de l'exemple 1 et du théorème XVI.1.4, si $k \in \mathbb{N}^*$ et si $(r_1, \dots, r_k) \in \mathbb{N}^k$ avec $r_1 + r_2 + \dots + r_k = n$, $r_1 \geq r_2 \geq \dots \geq r_k \geq 1$, pour qu'un endomorphisme nilpotent $u \in \text{Hom}_K(E)$, où $\dim(E) = n$ (resp. pour qu'une matrice nilpotente $M \in \mathfrak{M}_n(K)$) ait pour facteurs invariants $X^{r_1}, X^{r_2}, \dots, X^{r_k}$, il faut et il suffit qu'il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $\text{Mat}_{\mathcal{B}}(u) = \text{Diag } (J(r_1), \dots, J(r_k))$ (resp. que M soit semblable à $\text{Diag } (J(r_1), \dots, J(r_k))$).

En fin de compte, on a des bijections naturelles entre les trois ensembles suivants (E étant un K -ev donné de dimension $n \geq 1$) :

1° l'ensemble des suites finies (r_1, \dots, r_k) d'entiers ≥ 1 telles que $r_1 \geq r_2 \geq \dots \geq r_k$ et $r_1 + r_2 + \dots + r_k = n$, où k est arbitraire ($1 \leq k \leq n$),

2° l'ensemble des classes de similitude des matrices nilpotentes dans $\mathfrak{M}_n(K)$,

3° l'ensemble des classes de similitude d'endomorphismes nilpotents de E .

Ces trois ensembles sont finis, puisque le premier l'est de toute é

Exercice 1 : Avec les notations utilisées dans l'étude des facteurs invariants, montrer directement que la suite $(d_j - d_{j-1})$ est décroissante, en construisant une injection de N_j/N_{j-1} dans N_{j-1}/N_{j-2} pour $j \geq 2$.

Exercice 2 : Dans un K -ev E de dimension $n \geq 1$, soit u un endomorphisme nilpotent de période n . Montrer que les seuls sous- K -ev u -stables de E sont $\{0\}$, $\text{Ker}(u)$, $\text{Ker}(u^2)$, ..., $\text{Ker}(u^{n-1})$, E .

Exercice 3 : Soit $n \in \mathbb{N}^*$. Montrer que les matrices $J(n)$ et $\sum_{k \geq 1} (1+k)(J(n))^k$ sont semblables, le corps de base K étant supposé de caractéristique nulle.

Exercice 4 : Pour $n \in \llbracket 2, 6 \rrbracket$, dénombrer les classes de similitude de matrices nilpotentes de $\mathfrak{M}_n(K)$.

Exercice 5 : Soit u un endomorphisme nilpotent d'un K -ev E de dimension $n \geq 2$. Prouver que la période de u est n ssi $\text{rg}(u) = n - 1$.

Exercice 6 : On prend pour K le corps $\mathbb{Z}/p\mathbb{Z}$ (p premier impair). Dans $\mathfrak{M}_p(K)$ on considère la matrice

$$M = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 0 & 1 \\ +1 & 0 & \dots & 0 & -1 \end{bmatrix}.$$

Montrer que M est nilpotente et préciser sa période.

Exercice 7 : $K = \mathbb{C}$. On considère la matrice

$$M = \begin{bmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{bmatrix}.$$

Montrer qu'elle est nilpotente et chercher dans quelle base de \mathbb{C}^4 on peut la réduire à la forme de Jordan $J(4)$.

§ XVI.2 RÉDUCTION DE JORDAN QUAND $\chi_u(X)$ EST DISSOCIÉ

Considérons un K -ev E de dimension $n \geq 1$, et un endomorphisme u de E dont le polynôme caractéristique $\chi_u(X)$ est dissocié sur K :

$$(1) \quad \chi_u(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$$

(les λ_i distincts, les $\alpha_i \geq 1$ et $p \geq 1$).

Nous noterons $(F_j)_{1 \leq j \leq p}$ les sous-espaces caractéristiques de u , et pour chaque j , $u_j = u|_{F_j}$, $w_j = u_j - \lambda_j \text{Id}_{F_j}$. Nous savons que w_j est nilpotent de période β_j , avec $1 \leq \beta_j \leq \alpha_j$, le polynôme minimal $Q_u(X)$ étant $\prod_{i=1}^p (X - \lambda_i)^{\beta_i}$.

Notons $(X^{r_{i,1}}, X^{r_{i,2}}, \dots, X^{r_{i,k_i}})$ les facteurs invariants de w_i ($1 \leq i \leq p$). On a donc $r_{i,1} = \beta_i$; $k_i = \dim(\text{Ker}(w_i))$ est la dimension du sous-espace propre de u associé à la valeur propre λ_i .

DÉFINITION XVI.2.1

Avec les notations ci-dessus, on appelle **facteurs invariants** de u la suite des polynômes

$$(2) \quad \Phi_1(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,1}},$$

$$\Phi_2(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,2}}, \dots, \Phi_k(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,k_i}},$$

où $k = \max_{i=1}^p (k_i)$. Il résulte de cette définition que

$$\Phi_k(X) \mid \Phi_{k-1}(X) \mid \dots \mid \Phi_1(X), \quad \text{et} \quad \deg(\Phi_k) \geq 1.$$

On définit de manière analogue les **facteurs invariants d'une matrice** $M \in \mathfrak{M}_n(K)$ dont le polynôme caractéristique est donné par (1). En appliquant les résultats du § XVI.1 aux endomorphismes w_i , on obtient :

THÉORÈME XVI.2.1

Soit un K -ev E de dimension $n \geq 1$ et $u \in \text{Hom}_K(E)$ un endomorphisme dont le polynôme caractéristique est défini par (1) et dont les facteurs invariants sont définis par (2). Alors il existe au moins une base \mathcal{C} de E telle que $\text{Mat}_{\mathcal{C}}(u)$ soit diagonale par blocs de dimensions $\alpha_1, \dots, \alpha_p$:

$$(3) \quad \text{Mat}_{\mathcal{C}}(u) = \text{Diag}(S_1, \dots, S_p)$$

($S_i \in \mathfrak{M}_{\alpha_i}(K)$ pour tout i). Chaque bloc S_i étant lui-même une matrice diagonale par blocs de dimensions $r_{i,1}, \dots, r_{i,k_i}$:

$$S_i = \lambda_i I_{\alpha_i} + \text{Diag}(J(r_{i,1}), J(r_{i,2}), \dots, J(r_{i,k_i}))$$

$$= \text{Diag}(J_{\lambda_i}(r_{i,1}), J_{\lambda_i}(r_{i,2}), \dots, J_{\lambda_i}(r_{i,k_i})).$$

De manière équivalente, on peut énoncer ce théorème sous sa forme matricielle

Toute matrice $M \in \mathfrak{M}_n(K)$ dont le polynôme caractéristique est donné par (1) et dont les facteurs invariants sont donnés par (2) est semblable à une matrice du type (3).

Déterminer une base \mathcal{C} satisfaisant aux conditions de cet énoncé, c'est par définition, *réduire u à la forme de Jordan* dans la base \mathcal{C} .

Il s'agit maintenant de caractériser la classe de similitude de u . Il est d'abord clair que si deux endomorphismes à polynômes caractéristiques dissociés sur K ont **mêmes facteurs invariants**, ils sont **semblables**, en vertu du théorème XVI.2.1.

Pour établir la réciproque, nous allons montrer que les facteurs invariants de u (dont le polynôme caractéristique est donné par (1) et les facteurs inv

peuvent être déterminés à partir de certains sous-espaces remarquables. Pour cela, reprenons toutes les notations du début de ce paragraphe.

Pour chaque $i \in \llbracket 1, p \rrbracket$, nous avons la chaîne de sous-espaces de E :

$$\{0\} \subsetneq \text{Ker} (u_i - \lambda_i \text{Id}_{F_i}) \subsetneq \dots \subsetneq \text{Ker} (u_i - \lambda_i \text{Id}_{F_i})^{\beta_i} = F_i.$$

Si nous posons : $d_{i,q} = \dim (\text{Ker} (u_i - \lambda_i \text{Id}_{F_i})^q)$ pour $0 \leq q \leq \beta_i$, nous en déduisons, d'après les résultats du § XVI.1, la fonction décroissante $\psi_i : \llbracket 0, \beta_i + 1 \rrbracket \rightarrow \mathbb{N}$ telle que $\psi_i(0) = d_{i,1}$, $\psi_i(q) = d_{i,q} - d_{i,q-1}$ pour $q \in \llbracket 1, \beta_i \rrbracket$ et $\psi_i(\beta_i + 1) = 0$. Nous avons vu au § 1 que cette fonction ψ_i détermine la suite des facteurs invariants $X^{r_{i,1}}, \dots, X^{r_{i,k_i}}$ de w_i : le nombre k_i est $d_{i,1}$, et la fonction $q \mapsto r_{i,q}$ prend la valeur j pour $\psi_i(j+1) < q \leq \psi_i(j)$ ($1 \leq j \leq \beta_i$).

Cela dit, considérons $u' \in \text{Hom}_K(E)$ semblable à u , et soit $\varphi \in \text{GL}_K(E)$ tel que $u' = \varphi u \varphi^{-1}$. On sait déjà que $\chi_{u'}(X) = \chi_u(X)$. Notons $F'_i = \text{Ker} (u' - \lambda_i \text{Id}_E)^{\alpha_i}$ ($1 \leq i \leq p$) les sous-espaces caractéristiques de u' , et pour chaque i , $u'_i = u'|_{F'_i}$, $w'_i = u'_i - \lambda_i \text{Id}_{F'_i}$. Quel que soit $q \in \mathbb{N}$, on a :

$$\text{Ker} (u' - \lambda_i \text{Id}_E)^q = \varphi (\text{Ker} (u - \lambda_i \text{Id}_E)^q). \quad \text{Notons } \varphi_i = \varphi|_{F'_i}.$$

On voit que $F'_i = \varphi(F_i)$, et $u'_i = \varphi_i u_i \varphi_i^{-1}$, $w'_i = \varphi_i w_i \varphi_i^{-1}$. Cela entraîne que les nilpotents w_i et w'_i ont même période β_i (car $(w'_i)^q = \varphi_i w_i^q \varphi_i^{-1}$ pour tout $q \in \mathbb{N}$), puis, que $\text{Ker} ((w'_i)^q) = \varphi(\text{Ker} (w_i^q))$ pour tout $q \in \mathbb{N}$. Posons

$$d'_{i,q} = \dim (\text{Ker} (u'_i - \lambda_i \text{Id}_{F'_i})^q) \quad (q \in \mathbb{N}),$$

et soit ψ'_i la fonction construite à partir des $d'_{i,q}$ comme ψ_i l'est à partir des $d_{i,q}$. Il résulte de ce qui précède que $d'_{i,q} = d_{i,q}$ pour tout q , donc que $\psi'_i = \psi_i$. Donc la suite des facteurs invariants de w'_i est la même que celle de w_i , et cela est vrai pour tout $i \in \llbracket 1, p \rrbracket$. Il s'ensuit, par construction même des facteurs invariants, que les facteurs invariants de u et u' sont les mêmes, d'où :

THÉORÈME XVI.2.2

Soit E un K -ev de dimension $n \geq 1$ et u, u' deux endomorphismes dont le polynôme caractéristique est **dissocié** dans $K[X]$. Pour que u et u' soient **semblables**, il faut et il suffit qu'ils aient **mêmes facteurs invariants**.

Traduit en langage matriciel, cet énoncé devient :

Pour que deux matrices $M \in \mathfrak{M}_n(K)$, $M' \in \mathfrak{M}_n(K)$ à polynôme caractéristique **dissocié** sur K soient **semblables**, il faut et il suffit qu'elles aient **mêmes facteurs invariants**.

Exemple 1 : Donnons-nous à l'avance un polynôme $\chi(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i}$ dissocié dans $K[X]$ ($p \geq 1$, les $\alpha_i \geq 1$, les λ_i distincts) et, pour chaque $i \in \llbracket 1, p \rrbracket$, des entiers $k_i \geq 1$, $r_{i,1}, \dots, r_{i,k_i}$ tels que

$$r_{i,1} + r_{i,2} + \dots + r_{i,k_i} = \alpha_i \quad \text{et} \quad r_{i,1} \geq r_{i,2} \geq \dots \geq r_{i,k_i} \geq$$

Définissons, pour $i \in \llbracket 1, p \rrbracket$ la matrice

$$(4) \quad S_i = \lambda_i I_{\alpha_i} + \text{Diag} (J(r_{i,1}), \dots, J(r_{i,k_i})) = \\ = \text{Diag} (J_{\lambda_i}(r_{i,1}), \dots, J_{\lambda_i}(r_{i,k_i})).$$

Soit enfin M la matrice

$$(5) \quad M = \text{Diag} (S_1, S_2, \dots, S_p).$$

Alors les facteurs invariants de M sont nécessairement

$$(6) \quad \Phi_1(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,1}}, \Phi_2(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,2}}, \dots, \\ \Phi_k(X) = \prod_{i=1}^p (X - \lambda_i)^{r_{i,k}}, \quad \text{où} \quad k = \max_{i=1}^p (k_i).$$

C'est une conséquence immédiate du fait, expliqué dans l'exemple 1 du § XVI.1 que, pour chaque i , les facteurs invariants de S_i sont $X^{r_{i,1}}, \dots, X^{r_{i,k_i}}$.

Compte tenu des théorèmes XVI.2.1 et XVI.2.2, cet exemple 1 nous permet d'affirmer, en conservant ces notations :

THÉOREME XVI.2.3

Soit E un K -ev de dimension $n \geq 1$ et $u \in \text{Hom}_K(E)$ un endomorphisme à polynôme caractéristique **dissocié** sur K . Pour que u admette pour facteurs invariants $\Phi_1(X), \dots, \Phi_k(X)$ donnés par (6), il faut et il suffit qu'il existe une base ordonnée \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(u)$ soit la matrice M donnée par (4) et (5).

Sous sa forme purement matricielle, cet énoncé devient :

Pour qu'une matrice $N \in \mathfrak{M}_n(K)$, à polynôme caractéristique **dissocié** sur K , admette $\Phi_1(X), \dots, \Phi_k(X)$ donnés par (6) pour facteurs invariants, il faut et il suffit que N soit semblable à la matrice M donnée par (4) et (5).

Finalement, on a des bijections naturelles entre les trois ensembles suivants (E étant un K -ev de dimension $n \geq 1$) :

1° l'ensemble de toutes les matrices données par les formules (4) et (5) (où l'on fait varier les λ_i , les entiers k_i et les suites $r_{i,1}, \dots, r_{i,k_i}$),

2° l'ensemble des classes de similitude d'endomorphismes à polynôme caractéristique dissocié du K -ev donné E ,

3° l'ensemble des classes de similitude de matrices $N \in \mathfrak{M}_n(K)$ à polynôme caractéristique dissocié sur K .

En particulier, lorsque K est algébriquement clos, les théorèmes XVI.2.2 et XVI.2.3 résolvent complètement le problème de la détermination des classes de similitude dans $\text{Hom}_K(E)$ (resp. dans $\mathfrak{M}_n(K)$).

Exemple 2 : Le corps de base est \mathbb{C} . On considère $M \in \mathfrak{M}_5(\mathbb{C})$ donnée par :

$$M = \begin{bmatrix} 1 & 1 & -1 & 2 & -1 \\ 2 & 0 & 1 & -4 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & 3 & -1 \end{bmatrix}.$$

Trouver la forme réduite de Jordan de M et trouver une matrice de passage $P \in \text{GL}(5, \mathbb{C})$ telle que $P^{-1}MP$ soit cette forme réduite.

Solution. On note $\mathcal{B} = (e_i)_{1 \leq i \leq 5}$ la base canonique du \mathbb{C} -ev \mathbb{C}^5 , et u l'endomorphisme de \mathbb{C}^5 tel que $\text{Mat}_{\mathcal{B}}(u) = M$.

Le calcul de $\chi_M(X)$ est relativement aisé et donne :

$$\chi_M(X) = (1 - X)^3(1 + X)^2.$$

Notons $F_1 = \text{Ker}(u - \text{Id}_{\mathbb{C}^5})^3$ et $F_{-1} = \text{Ker}(u + \text{Id}_{\mathbb{C}^5})^2$ les sous-espaces caractéristiques de u , et $M_1 = M - I_5$, $M_{-1} = M + I_5$, $u_1 = u|_{F_1}$, $u_{-1} = u|_{F_{-1}}$.

On a successivement :

$$M_1^2 = \begin{bmatrix} 2 & 0 & 8 & -10 & 2 \\ -2 & 0 & -8 & 10 & -2 \\ 2 & 0 & 0 & -2 & -2 \\ 2 & 0 & -4 & 2 & -2 \\ 0 & 0 & 12 & -12 & 4 \end{bmatrix},$$

$$M_1^3 = \begin{bmatrix} 0 & 0 & -28 & 28 & -8 \\ 0 & 0 & 28 & -28 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & -8 & 0 \\ 0 & 0 & -36 & 36 & -8 \end{bmatrix},$$

$$M_{-1}^2 = \begin{bmatrix} 6 & 4 & 4 & -2 & -2 \\ 6 & 0 & -4 & -6 & -6 \\ 2 & 4 & 4 & 2 & 2 \\ 2 & 4 & 4 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

On sait à l'avance que $\text{rg}(M_1^3) = 2$ et $\text{rg}((M_{-1})^2) = 3$ car $\dim(F_1) = 3$ et $\dim(F_{-1}) = 2$.

On constate que $\text{rg}(M_1^2) = 3$ et $\text{rg}(M_{-1}) = 4$; donc $u_1^2 \neq 0$ et $u_{-1} \neq 0$, ce qui indique que les périodes respectives de M_1 et M_{-1} sont 3 et 2 : leurs facteurs invariants respectifs sont : X^3 pour u_1 et X^2 pour u_2 . Donc la forme réduite de Jordan de M est

$$N = \text{Diag}(I_3 + J(3), -I_2 + J(2)) = \text{Diag}(J_1(3), J_{-1}(2)),$$

$$\text{soit } N = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \text{ Il reste à trouver } P.$$

Pour cela, construisons des vecteurs x_1 et x_{-1} tels que $x_1 \in F_1 \setminus \text{Ker}(u_1^2)$ et $x_{-1} \in F_{-1} \setminus \text{Ker}(u_{-1})$. Une matrice P convenable s'en déduira en écrivant la matrice de passage de \mathcal{B} à \mathcal{C} , où

$$\mathcal{C} = ((u - \text{Id}_{\mathbb{C}^5})^2(x_1), (u - \text{Id}_{\mathbb{C}^5})(x_1), x_1; (u + \text{Id}_{\mathbb{C}^5})(x_{-1}), x_{-1}).$$

Or par exemple si $x_1 = e_3 + e_4$, on a $x_1 \in F_1$, mais $x_1 \notin \text{Ker}(u_1^2)$. De même $x_{-1} = e_4 - e_5 \in F_{-1}$, mais $x_{-1} \notin \text{Ker}(u_{-1})$. On trouve ainsi que la matrice suivante convient :

$$P = \begin{bmatrix} -2 & 1 & 0 & 3 & 0 \\ 2 & -3 & 0 & -3 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ -2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 & -1 \end{bmatrix}.$$

Le lecteur vérifiera sans peine que l'on a bien $MP = PN$:

$$\begin{aligned} & \begin{bmatrix} 1 & 1 & -1 & 2 & -1 \\ 2 & 0 & 1 & -4 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & 3 & -1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 & 3 & 0 \\ 2 & -3 & 0 & -3 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ -2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 & -1 \end{bmatrix} = \\ & = \begin{bmatrix} -2 & 1 & 0 & 3 & 0 \\ 2 & -3 & 0 & -3 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ -2 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \\ & = \begin{bmatrix} -2 & -1 & 1 & -3 & 3 \\ 2 & -1 & -3 & 3 & -3 \\ -2 & -1 & 2 & 0 & 0 \\ -2 & -1 & 2 & 0 & -1 \\ 0 & 0 & 0 & -3 & 4 \end{bmatrix} \end{aligned}$$

Exercice 1 : Si le corps de base K est algébriquement clos, et si $n \in \mathbb{N}^*$, prouver que toute matrice $M \in \mathfrak{M}_n(K)$ est semblable à sa transposée, en utilisant une réduction de Jordan, ce qui permet de se ramener au cas où M est une matrice de Jordan.

Etendre ce résultat au cas où M n'est pas algébriquement clos en utilisant le théorème VII.7.3 et l'exercice 11 du § XIV.3.

Exercice 2 : Le corps de base est \mathbb{C} . Trouver la forme réduite de Jordan des matrices suivantes :

$$a) M = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad b) \begin{bmatrix} 3 & -5 & 2 & -6 \\ 0 & 5 & 0 & 4 \\ -2 & 7 & -1 & 11 \\ 0 & -4 & 0 & -3 \end{bmatrix}; \quad c) \begin{bmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -1 & 7 \end{bmatrix};$$

$$d) \begin{bmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{bmatrix}; \quad e) M = \begin{bmatrix} 0 & 0 & \dots & 0 & n \\ \vdots & & & & \\ 0 & \dots & \dots & \dots & n-1 \\ n & n-1 & \dots & 2 & 1 \end{bmatrix}.$$

Pour chacune de ces matrices on précisera le changement de base à effectuer pour cette réduction et on calculera également M^k pour $k \in \mathbb{N}$ ainsi que $\exp(M)$.

Exercice 3 : On donne $P(X) = \prod_{i=1}^p (\lambda_i - X)^{\alpha_i} \in K[X]$,

avec des λ_i distincts, des $\alpha_i \geq 1$ et $p \geq 1$ et

$$\alpha_1 + \alpha_2 + \dots + \alpha_p = n \geq 1$$

et on donne des entiers $\beta_i \in \llbracket 1, \alpha_i \rrbracket$ pour $i \in \llbracket 1, p \rrbracket$. Trouver une matrice $M \in \mathfrak{M}_n(K)$ telle que $\chi_M(X) = P(X)$ et $\mathcal{Q}_M(X) = \prod_{i=1}^p (X - \lambda_i)^{\beta_i}$.

Exercice 4 : Le corps de base est \mathbb{C} . On donne un entier $n \geq 2$.

a) Pour $(z_1, z_2, \dots, z_{n-1}) \in \mathbb{C}^{n-1}$, donner la classe de similitude de la matrice

$$M(z_1, \dots, z_{n-1}) = [a_{i,j}]$$

telle que $a_{i,i+1} = z_i$ pour $i \in \llbracket 1, n-1 \rrbracket$ et $a_{i,j} = 0$ pour tous les autres couples (i, j) .

b) Soit \mathcal{S} l'ensemble des suites (r_1, r_2, \dots, r_k) d'entiers ≥ 1 , où k est variable ($k \geq 1$) telles que

$$r_1 \geq r_2 \geq \dots \geq r_k \geq 1 \quad \text{et} \quad r_1 + r_2 + \dots + r_k = n.$$

A chaque élément $\rho = (r_1, r_2, \dots, r_k) \in \mathcal{S}$ on associe la suite $s(\rho) = (s_1, s_2, \dots, s_k)$

où $s_j = \sum_{i=1}^j r_i$ pour $j \in \llbracket 1, k \rrbracket$. Soit $\rho = (r_1, \dots, r_k)$ et $\rho' = (r'_1, \dots, r'_k)$ deux éléments de \mathcal{S} .

On écrit que $\rho \leq \rho'$ ssi $s(\rho)$ est une sous-suite de $s(\rho')$. Vérifier que \leq est une relation d'ordre sur \mathcal{S} .

c) Soit $\rho = (r_1, \dots, r_k) \in \mathcal{S}$. On note S_ρ la classe de similitude des matrices nilpotentes de $\mathfrak{M}_n(\mathbb{C})$ de facteurs invariants $(X^{r_1}, \dots, X^{r_k})$. Dédurre du a) que l'adhérence de S_ρ dans $\mathfrak{M}_n(\mathbb{C})$ contient l'ensemble \mathcal{S}_ρ des $S_{\rho'}$ pour $\rho' \geq \rho$.

d) Avec les notations du c), soit $M' \in \mathfrak{M}_n(\mathbb{C})$ une matrice adhérente à S_ρ dans $\mathfrak{M}_n(\mathbb{C})$. Montrer : $\text{rg}(M'^k) \leq \text{rg}(M^k)$ pour tout $k \in \mathbb{N}$, et en déduire que l'adhérence de S_ρ dans $\mathfrak{M}_n(\mathbb{C})$ est \mathcal{S}_ρ .

e) Soit S une classe de similitude *quelconque* dans $\mathfrak{M}_n(\mathbb{C})$. A l'aide des résultats ci-dessus, déterminer l'adhérence de S dans $\mathfrak{M}_n(\mathbb{C})$.

Exercice 5 : Utiliser les invariants de similitude pour reconnaître que les matrices

$$\begin{bmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 6 & 20 & -34 \\ 6 & 32 & -51 \\ 4 & 20 & -32 \end{bmatrix}$$

sont semblables. Même question pour

$$\begin{bmatrix} 4 & 10 & -19 & 4 \\ 1 & 6 & -8 & 3 \\ 1 & 4 & -6 & 2 \\ 0 & -1 & 1 & 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 41 & -4 & -26 & -7 \\ 14 & -13 & -91 & -18 \\ 40 & -4 & -25 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercice 6 : On donne $n \in \mathbb{N}^*$ ($n \geq 2$). On suppose ici le corps K algébriquement clos et de caractéristique $p > 0$. La base canonique (e_1, \dots, e_n) du K -ev $K^n = E$ est notée \mathcal{E} . Pour $\sigma \in \mathfrak{S}_n$, on note f_σ l'élément de $\text{Hom}_K(E)$ tel que $f_\sigma(e_j) = e_{\sigma(j)}$ pour $1 \leq j \leq n$, et P_σ la matrice $\text{Mat}_{\mathcal{E}}(f_\sigma)$. On rappelle que $\sigma \mapsto f_\sigma$ définit un isomorphisme du groupe \mathfrak{S}_n sur un sous-groupe, qu'on notera S_n , de $\text{GL}_K(E)$.

a) On suppose que σ est un cycle de longueur n . On écrit $n = p^s m$, avec $s \in \mathbb{N}$ et $m \in \mathbb{N}^*$, m non divisible par p . Déterminer $\chi_{f_\sigma}(X)$ et $Q_{f_\sigma}(X)$. Prouver que f_σ est diagonalisable ssi $s = 0$. Donner alors une base de vecteurs propres de f_σ .

b) Ici σ est quelconque ($\sigma \neq \text{Id}_{\mathbb{E}_n}$). On pourra décomposer σ en cycles disjoints : $\sigma = c_1 c_2 \dots c_r$ ($r \geq 1$, longueur de $c_i = l_i \geq 2$).

Déterminer $\chi_{f_\sigma}(X)$ et $Q_{f_\sigma}(X)$. Donner une CNS sur les l_i pour que f_σ soit diagonalisable. Donner alors une base de vecteurs propres de f_σ .

c) On suppose à nouveau que σ est un cycle de longueur n , avec $n = p^s m$. Soit \mathcal{G}_m le sous-groupe multiplicatif $\{x \in K \mid x^m = 1_K\}$ de K^* . Vérifier que $\text{card}(\mathcal{G}_m) = m$. Posons $g = (f_\sigma)^{p^s}$. Montrer que g est diagonalisable, trouver $\chi_g(X)$, $Q_g(X)$ et les sous-espaces propres de g .

On écrit $\mathcal{G}_m = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$. Démontrer que la matrice P_σ est semblable à la matrice diagonale par blocs

$$G = \text{Diag} (J_{\lambda_1}(p^s), J_{\lambda_2}(p^s), \dots, J_{\lambda_m}(p^s))$$

et donner une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(f_\sigma) = G$.

Indication : on pourra s'aider de la décomposition en cycles de σ^{p^s} .

§ XVI.3 SOUS-ESPACES MONOGÈNES

Dans ce qui suit, E désigne un K -ev de dimension $n \geq 1$, et u un endomorphisme fixé de E . A chaque vecteur $x \in E$, on associe l'ensemble $\{P(u) \cdot x\}_{P \in K[X]}$ noté en abrégé $K[u] \cdot x$. Cet ensemble est un sous- K -ev u -stable de E qui contient x . Il est d'ailleurs clair que tout sous-espace u -stable de E qui contient x contient $K[u] \cdot x$, donc $K[u] \cdot x$ est, au sens de l'inclusion, le plus petit sous-espace u -stable de E contenant x .

DÉFINITION XVI.3.1

~ Pour $x \in E$, on appelle **sous-espace u -monogène engendré par x** le sous- K -ev $K[u] \cdot x$. Un sous-espace F de E est dit **u -monogène** ssi il existe $x \in E$ pour lequel $F = K[u] \cdot x$, et s'il en est ainsi, tout $x \in E$ vérifiant cette condition est appelé un **u -générateur** de F .

Lorsque u est *nilpotent*, cette notion coïncide avec la notion de sous-espace u -monogène déjà définie au § XVI.1 (définition XVI.1.2).

Idéal annulateur

Considérons une partie non vide A de E . L'ensemble des $P \in K[X]$ tels que

$$(\forall a \in A) \quad P(u) \cdot a = 0$$

est un idéal de $K[X]$ appelé **idéal u -annulateur de A** (ou **idéal annulateur de A** , si aucune confusion n'est à craindre sur u). Nous noterons $\text{Ann}_u(A)$ ou $\text{Ann}(A)$ cet idéal. On a toujours $\text{Ann}(A) = \text{Ann}(\text{Vect}(A))$. L'idéal $\text{Ann}(A)$ est toujours non nul, car il contient évidemment $\chi_u(X)$, et même le polynôme minimal de u : $Q_u(X)$.

Plus généralement, soit

$$K[u] \cdot A = \{P(u) \cdot a\}_{P \in K[X], a \in A}.$$

Alors $\text{Ann}(A) = \text{Ann}(K[u] \cdot A)$ (en particulier, pour $x \in E$, $\text{Ann}(x) = \text{Ann}(K[u] \cdot x)$). Finalement

$$\text{Ann}(A) = \text{Ann}(K[u] \cdot \text{Vect}(A)).$$

$K[u] \cdot \text{Vect}(A)$ est, au sens de l'inclusion, le plus petit sous-espace u -stable contenant A . Si $A \subset B \subset E$, on a bien sûr :

$$(1) \quad \text{Ann}(B) \subset \text{Ann}(A).$$

C'est encore une évidence que, pour toute partie A non vide de E :

$$(2) \quad \text{Ann}(A) = \bigcap_{a \in A} \text{Ann}(a).$$

Soit maintenant F un sous- K -ev u -stable de E . Puisque $\text{Ann}(F)$ est un idéal non nul de $K[X]$, on a un unique polynôme normalisé $Q_{u,F} \in K[X]$ tel que $Q_{u,F} K[X] = \text{Ann}(F)$. Il est clair que $Q_{u,E}(X) = Q_u(X)$ et que $Q_{u,F} = Q_{u \parallel F}$ lorsque $F \neq \{0\}$, par définition même du polynôme minimal, enfin que $Q_{u,\{0\}} = 1$.

Aux relations (1) et (2) correspondent les propriétés suivantes :

(3) si F et G sont deux sous-espaces u -stables de E tels que $F \subset G$, alors $Q_{u,F}$ divise $Q_{u,G}$. En particulier $Q_{u,F}$ divise toujours $Q_{u,E} = Q_u$.

(4) Pour tout sous-espace u -stable F de E ,

$$Q_{u,F}(X) = \text{ppcm}_{x \in F} (Q_{u,\{x\}}), \text{ et si } F \neq \{0\} : Q_{u,F}(X) = \text{ppcm}_{x \in F \setminus \{0\}} (Q_{u,\{x\}})$$

(le ppcm étant toujours défini du fait que tous les $Q_{u,K[u] \cdot x}$ divisent $Q_u(X)$). Pour abréger, nous noterons, lorsque $x \in E$, $Q_{u,x}$ à la place de $Q_{u,\{x\}}$.

THÉORÈME XVI.3.1

$$\left\| \begin{array}{l} \text{Soit } x \in E. \text{ Alors } \dim(K[u] \cdot x) = \deg(Q_{u,x}). \\ \text{Donc, si } x \neq 0_E, \text{ on a, en notant } v = u \parallel_{K[u] \cdot x} \text{ et } d = \dim(K[u] \cdot x) \\ \chi_v(X) = (-1)^d Q_{u,x}(X) = (-1)^d Q_v(X). \end{array} \right.$$

Démonstration :

Si $x = 0_E$, $K[u] \cdot x = \{0_E\}$ et $Q_{u,x} = 1$, d'où le résultat. Si $x \neq 0_E$, c'est-à-dire $K[u] \cdot x \neq \{0_E\}$, on a déjà vu que $Q_v(X) = Q_{u,x}(X)$. Soit $a_0 + a_1 X + \dots + a_{d-1} X^{d-1} + X^d$ ce polynôme. Alors

$$K[u] \cdot x = \text{Vect}(x, v(x), \dots, v^{d-1}(x)).$$

Les vecteurs $x, v(x), \dots, v^{d-1}(x)$ sont linéairement indépendants, car si

$$(\lambda_0, \lambda_1, \dots, \lambda_{d-1}) \in K^d$$

vérifie

$$\lambda_0 x + \lambda_1 v(x) + \dots + \lambda_{d-1} v^{d-1}(x) = 0,$$

le polynôme $\lambda_0 + \dots + \lambda_{d-1} X^{d-1}$ appartient à $\text{Ann}(x)$, donc est multiple de $Q_v(X)$, ce qui n'est possible que si

$$\lambda_0 = \lambda_1 = \dots = \lambda_{d-1} = 0.$$

Par suite $(x, v(x), \dots, v^{d-1}(x))$ est une base de $K[u] \cdot x$, d'où

$$\dim(K[u] \cdot x) = d = \deg(Q_v(X)).$$

Puisque $Q_v(X)$ divise $\chi_v(X)$ qui est aussi de degré d , il s'ensuit bien que

$$\chi_v(X) = (-1)^d Q_v(X). \quad \blacksquare$$

Remarque 1 : La matrice de v dans la base $(v^{d-1}(x), \dots, v(x), x)$ est

$$M = \begin{bmatrix} -a_{d-1} & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ -a_1 & 0 & \vdots & \vdots & 1 \\ -a_0 & 0 & \dots & \dots & 0 \end{bmatrix}.$$

Il ne serait pas difficile par un calcul direct de vérifier que

$$\chi_M(X) = (-1)^d Q_v(X),$$

ce qui est une autre façon d'achever la démonstration précédente.

Nous allons maintenant établir une sorte de réciproque du théorème XVI.3.1. De façon précise, si F est un sous-espace u -stable $\neq \{0\}$, si on pose $v = u \parallel_F$, nous allons voir que la relation $\chi_v(X) = (-1)^d Q_v(X)$ entraîne que F est u -monogène. Il suffit évidemment de le prouver avec $F = E$.

THÉORÈME XVI.3.2

|| Si $\chi_u(X) = (-1)^n Q_u(X)$, alors E est u -monogène. La condition $\deg Q_u(X) = n$ est donc nécessaire et suffisante pour que E soit u -monogène.

Démonstration :

Décomposons $Q_u(X)$ en facteurs irréductibles dans $K[X]$:

$$Q_u(X) = \prod_{i=1}^p P_i^{e_i}(X)$$

(avec $p \geq 1$, des P_i irréductibles *normalisés* deux à deux distincts, et des $e_i \geq 1$). Rappelons que

$$Q_u(X) = \text{ppcm}_{y \in E \setminus \{0\}} (Q_{u,y}(X)).$$

Nous pouvons donc choisir, et nous le faisons, pour chaque $i \in \llbracket 1, p \rrbracket$, un $y_i \in E \setminus \{0\}$ tel que

$$Q_{u,y_i}(X) = P_i^{e_i}(X) S_i(X),$$

avec $S_i(X) \in K[X] \setminus \{0\}$. Posons $z_i = S_i(u) \cdot y_i$ ($1 \leq i \leq p$).

Alors $P_i^{e_i}(u) \cdot z_i = 0$, et il est immédiat que

$$Q_{u,z_i}(X) = P_i^{e_i}(X)$$

car, si $f \in K[X]$, on a $S_i \neq 0$, et :

$$(f(u) \cdot z_i = 0) \Leftrightarrow (f(X) S_i(X) \equiv 0 \text{ mod } (P_i^{e_i}(X) S_i(X))).$$

Soit alors $z = z_1 + z_2 + \dots + z_p$. Nous allons montrer que $Q_{u,z}(X) = Q_u(X)$, et alors par application du théorème XVI.3.1, on aura :

$$\dim (K[u] \cdot z) = \deg (Q_u(X)) = n = \dim (E),$$

d'où $K[u] \cdot z = E$, ce qui achèvera la démonstration. Pour établir cela, posons $F_i = K[u] \cdot z_i$ ($1 \leq i \leq p$).

Prouvons d'abord que les sous-K-ev F_i sont indépendants :

Soit $g_1, \dots, g_p \in K[X]$ tels que $g_1(u) \cdot z_1 + \dots + g_p(u) \cdot z_p = 0$.

Posons $T_k(X) = Q_u(X)/P_k^{e_k}(X)$ pour $1 \leq k \leq p$, et fixons $i \in \llbracket 1, p \rrbracket$. On a :

$$0 = T_i(u)[g_1(u) \cdot z_1 + \dots + g_p(u) \cdot z_p] = [T_i(u) \circ g_i(u)] \cdot z_i,$$

car $[T_i(u) \circ g_j(u)] \cdot z_j = 0$ pour $i \neq j$. Donc

$$T_i g_i \equiv 0 \text{ mod } (P_i^{e_i}(X)),$$

et comme T_i et $P_i^{e_i}$ sont premiers entre eux, il s'ensuit que :

$$g_i(X) \equiv 0 \text{ mod } (P_i^{e_i}(X)),$$

d'où $g_i(u) \cdot z_i = 0$, et c'est vrai pour tout i , d'où l'assertion.

Montrons que $Q_{u,z}(X) = Q_u(X)$:

Il est déjà certain que $Q_{u,z}(X)$ divise $Q_u(X)$; soit $f \in \text{Ann}_u(z)$, alors

$$f(u) \cdot z = 0 = \sum_{i=1}^p f(u) \cdot z_i,$$

et $f(u) \cdot z_i \in F_i$ pour tout i , donc, d'après ce qui précède, $f(u) \cdot z_i = 0$ pour tout i d'où $f \equiv 0 \pmod{(P_i^{e_i})}$ pour tout i ; il s'ensuit :

$$f \equiv 0 \pmod{\left(\prod_{i=1}^p P_i^{e_i}\right)},$$

donc en particulier $Q_{u,z}(X)$ est multiple de $Q_u(X)$, et finalement $Q_{u,z}(X) = Q_u(X)$; on achève la preuve comme indiqué plus haut. ■

Remarque 2 : avec les notations de cette démonstration, on a pour tout i , $\deg(P_i^{e_i}) = \dim(F_i)$, d'où

$$\dim\left(\bigoplus_{i=1}^p F_i\right) = \sum_{i=1}^p \deg(P_i^{e_i}) = n = \dim(E),$$

donc $\bigoplus_{i=1}^p F_i = E$, et $n = \deg(Q_{u,z}(X) = \dim(K[u] \cdot z))$, donc

$$K[u] \cdot z = E = \bigoplus_{i=1}^p F_i,$$

alors qu'*a priori* seule l'inclusion $K[u] \cdot z \subset \bigoplus_{i=1}^p F_i$ était évidente.

Retour sur la réduction de Jordan

Commençons par un résultat préliminaire intéressant en soi :

LEMME 1

Soit F_1, F_2, \dots, F_p ($p \in \mathbb{N}^*$) des sous- K -ev u -stables dans E , et soit $F = \sum_{i=1}^p F_i$. Alors

$$Q_{u,F}(X) = \text{ppcm}_{i=1}^p Q_{u,F_i}(X).$$

Démonstration :

Soit P le polynôme $\text{ppcm}_{i=1}^p Q_{u,F_i}(X)$; il est clair que $P(u)$ s'annule sur chaque F_i , donc sur $F = \sum_{i=1}^p F_i$, donc $P(X)$ est multiple de $Q_{u,F}(X)$. D'autre part, pour chaque i , on a : $F_i \subset F$, donc $Q_{u,F_i}(X)$ divise $Q_{u,F}(X)$, et par suite $P(X)$ divise $Q_{u,F}(X)$; finalement $Q_{u,F}(X) = P(X)$. ■

Nous supposons maintenant $\chi_u(X)$ dissocié dans $K[X]$, et nous reprenons toutes les notations en vigueur au début du § XVI.2. Nous allons interpréter les facteurs invariants $\Phi_1(X), \dots, \Phi_k(X)$ de u .

Pour chaque $i \in \llbracket 1, p \rrbracket$, choisissons des sous-espaces w_j -monogènes $(G_{i,j})_{1 \leq j \leq k_i}$ tels que

$$\bigoplus_{j=1}^{k_i} G_{i,j} = F_i \quad \text{et} \quad \dim(G_{i,j}) = r_{i,j} \quad \text{pour tout } j.$$

Cela est possible en vertu des résultats du § XVI.1. Chaque $G_{i,j}$ est u_i -monogène, donc u -monogène, et l'on a :

$$(\forall i) \quad (\forall j) \quad Q_{u_i, G_{i,j}}(X) = (X - \lambda_i)^{r_{i,j}} = Q_{u, G_{i,j}}(X).$$

Définissons les sous-espaces $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ par :

$$\Gamma_j = \bigoplus_{i=1}^p G_{i,j} \quad (1 \leq j \leq k).$$

Les sous-espaces Γ_j sont non nuls, u -stables, et l'on a : $E = \bigoplus_{j=1}^k \Gamma_j$. D'après le lemme 1,

$$Q_{u, \Gamma_j}(X) = \text{ppcm}_{i=1}^p (X - \lambda_i)^{r_{i,j}} = \prod_{i=1}^p (X - \lambda_i)^{r_{i,j}} = \Phi_j(X),$$

de sorte que

$$\deg(Q_{u, \Gamma_j}(X)) = \sum_{i=1}^p r_{i,j} = \sum_{i=1}^p \dim(G_{i,j}) = \dim(\Gamma_j).$$

En vertu du théorème XVI.3.2, on en déduit que Γ_j est u -monogène, et que

$$Q_{u, \Gamma_j}(X) = \Phi_j(X).$$

On a donc prouvé le résultat suivant :

THÉORÈME XVI.3.3

Soit u un endomorphisme du K -ev E de dimension $n \geq 1$, dont le polynôme caractéristique est **dissocié** sur K , et dont les facteurs invariants sont $\Phi_1(X), \dots, \Phi_k(X)$. Il existe une décomposition de E de la forme :
 $E = \bigoplus_{i=1}^p \Gamma_j$, où les Γ_j sont des sous-espaces **u -monogènes** tels que
 $(\forall j) \quad Q_{u, \Gamma_j}(X) = \Phi_j(X).$

Réciproquement, on peut montrer que, si $E = \bigoplus_{j=1}^k \mathcal{F}_j$, où les \mathcal{F}_j sont des sous-espaces u -monogènes tels que

$$Q_{u, \mathcal{F}_k}(X) \mid Q_{u, \mathcal{F}_{k-1}}(X) \mid \dots \mid Q_{u, \mathcal{F}_1}(X),$$

alors nécessairement la suite $(Q_{u, \mathcal{F}_1}(X), \dots, Q_{u, \mathcal{F}_k}(X))$ est la suite

$$(\Phi_1(X), \dots, \Phi_k(X))$$

des facteurs invariants de u .

Exercice 1 : On suppose le K -ev de dimension $n \geq 1$, et u -monogène, pour $u \in \text{Hom}_K(E)$. Trouver l'ensemble \mathcal{C}_u (dont on vérifiera que c'est une sous- K -algèbre de $\text{Hom}_K(E)$) :

$$\mathcal{C}_u = \{v \in \text{Hom}_K(E) \mid v \circ u = u \circ v\}$$

(\mathcal{C}_u est appelé le commutant de u). Préciser $\dim_K(\mathcal{C}_u)$.

Exercice 2 : Dans le K -ev E de dimension $n \geq 1$, on considère $u \in \text{Hom}_K(E)$ et le commutant \mathcal{C}_u de u défini dans l'exercice 1.

a) On suppose u nilpotent, de facteurs invariants X^{r_1}, \dots, X^{r_k} . Calculer $\dim_K(\mathcal{C}_u)$ à l'aide des r_i . Trouver la CNS pour que $\dim_K(\mathcal{C}_u) = n$.

b) On suppose $\chi_u(X)$ dissocié dans $K[X]$. Trouver une CNS pour que $\dim_K(\mathcal{C}_u) = n$.

Exercice 3 : Soit E un K -ev de dimension finie $n \geq 1$, et $u \in \text{Hom}_K(E)$.

a) On suppose $Q_u(X) = P^e$, où $e \geq 1$ et où $P \in K[X]$ est irréductible et normalisé dans $K[X]$. Soit $v = 'u$.

1) Le polynôme $Q_v(X)$ est P^e .

2) En déduire qu'il existe $\varphi \in E^*$ telle que $\varphi \circ P^{e-1}(u) \neq 0$. On choisit φ ainsi, et on choisit $x \in E$ tel que $(\varphi \circ P^{e-1}(u)) \cdot x \neq 0$. On pose : $F = K[u] \cdot x$, $G = {}^0(K[v] \cdot \varphi)$. Montrer que F est u -stable et que $\dim(F) = de$, où $d = \deg(P)$.

3) Prouver que G est u -stable et de dimension $n - de$.

4) Prouver que $F \cap G = \{0\}$ (on remarquera que

$$(1, X, \dots, X^{d-1}, Q, XQ, \dots, X^{d-1}Q, \dots, Q^{e-1}, XQ^{e-1}, \dots, X^{d-1}Q^{e-1})$$

est une base du K -ev $K_{de-1}[X]$), et en déduire que $F \oplus G = E$.

5) En déduire par récurrence sur n que E est somme directe interne de sous-espaces u -monogènes.

b) On suppose encore $Q_u(X) = P^e$. Montrer que $\chi_u(X)$ est une puissance de P .

c) On suppose $\chi_u(X) = \prod_{i=1}^r P_i^{e_i}(X)$, où les P_i sont irréductibles, deux à deux premiers entre eux, les $e_i \geq 1$, et $r \geq 1$. Montrer que :

$$E = \bigoplus_{i=1}^r \text{Ker}(P_i^{e_i}(u)),$$

et qu'on a : $(*) \quad (\forall i) \quad \dim(\text{Ker}(P_i^{e_i}(u))) = e_i \deg(P_i)$.

Montrer aussi que E est somme directe interne de sous-espaces u -monogènes et obtenir aussi une nouvelle démonstration du théorème XVI.3.2.

Montrer enfin que $Q_u(X) \equiv 0 \pmod{(P_1 P_2 \dots P_r)}$. Obtenir directement ce dernier résultat en raisonnant sur des matrices et en étendant le corps de base K par utilisation du théorème VII.7.3. Montrer que cette dernière méthode permet aussi de retrouver les relations (*).

Exercice 4 (endomorphismes semi-simples) :

Soit E un K -ev de dimension finie $n \geq 1$ et $u \in \text{Hom}_K(E)$. On dit que u est *semi-simple* ssi tout sous-espace u -stable de E admet au moins un supplémentaire u -stable.

a) Montrer que si u est diagonalisable, u est semi-simple.

b) On suppose $Q_u(X)$ irréductible dans $K[X]$. Montrer que u est semi-simple.

Indication : Soit V un sous- K -ev u -stable de E . Montrer que $f \in K[X]$, $x \in E$ et $f(u) \cdot x \in V$ entraînent : $f \equiv 0 \pmod{(Q_u(X))}$ ou $x \in V$.

En déduire que si $y \in E \setminus V$, le sous- K -ev $K[u] \cdot y = V'$ est u -stable et vérifie $V \cap V' = \{0\}$. Acheter en considérant un sous- K -ev u -stable de dimension maximum de V . On peut aussi procéder de façon plus rapide mais plus abstraite en considérant le corps $L = K[X]/Q_u(X)$ (cf. théorème VII.7.5) et en munissant E d'une structure naturelle de L -ev.

c) Montrer l'équivalence entre les deux conditions :

(I) u est semi-simple.

(II) Les exposants des facteurs irréductibles de $Q_u(X)$ sont tous égaux à 1 dans la décomposition de $Q_u(X)$.

Qu'en déduit-on si K est algébriquement clos ? Pouvait-on alors prouver directement ce résultat ?

BIBLIOGRAPHIE

Outre les ouvrages cités dans le texte, signalés par un astérisque, nous nous sommes efforcés de réunir une liste de références accessibles en rapport avec le présent traité.

- [1]* BOREVITCH Z. I., CHAFAREVITCH I. R., *Théorie des nombres*, Gauthier-Villars, 1967.
- [2]* BOURBAKI N., *Algèbre*, chap. I à III ; Hermann, 1982.
- [3]* BOURBAKI N., *Groupes et Algèbres de Lie*, Hermann, 1982.
- [4] BOUTELOUP, *L'algèbre linéaire*, P.U.F., coll. « Que Sais-je ? ».
- [5] CALAIS J., *Éléments de Théorie des groupes*, P.U.F., coll. « Mathématiques ».
- [6] CARREGA J. C., *Théorie des corps, la règle et le compas*, Hermann, 1981.
- [7]* CHAMBADAL L., OVAERT J.-L., *Algèbre linéaire et tensorielle*, Dunod, 1968.
- [8]* COHEN Paul, *Set Theory and the continuum hypothesis*, Benjamin, 1966.
- [9] COMTET L., *Analyse combinatoire*, T 1 et 2, P.U.F.
- [10]* DELLACHERIE C., *Nombres au hasard*, Public. Univ. Louis Pasteur, Strasbourg, 1978.
- [11] DICKSON L. E., *Introduction to the theory of numbers*, Dover.
- [12]* DUBREIL P., *Leçons d'Algèbre moderne*, Dunod, 1964.
- [13]* GAAL L., *Classical Galois Theory, with examples*, Chelsea, 3^e éd., 1979.
- [14]* GALOIS E., *Œuvres*, Gauthier-Villars, 1962.
- [15] GAUSS C. F., *Recherches Arithmétiques*, Blanchard, Paris.
- [16] GODEMENT R., *Cours d'Algèbre*, Hermann, 1969.
- [17] HARDY G. H., WRIGHT E. M., *An Introduction to the theory of numbers*, Oxford Un. Press, 1959.

- [18] ITARD J., *Les nombres premiers*, P.U.F., coll. « Que sais-je ? ».
- [19] ITARD J., *Arithmétique et théorie des nombres*, P.U.F., coll. « Que sais-je ? ».
- [20]* JACOBSON N., *Lectures in abstract algebra*, Tome 3, Springer.
- [21]* JORDAN C., *Traité des Substitutions*, Gauthier-Villars, 1927, Blanchard, 1957.
- [22]* KREISEL G., KRIVINE J. L., *Eléments de Logique Mathématique*, Dunod, 1967.
- [23]* KRIVINE J. L., *Théorie axiomatique des ensembles*, P.U.F., coll. « Sup ».
- [24] LEDERMANN W., *Introduction to the theory of finite groups*, Wiley, 1961.
- [25] MOISOTTE L., *Exercices de Mathématiques*, Dunod, 1982.
- [26] RYSER H. J., *Mathématiques combinatoires*, Dunod, 1969.
- [27]* SAMUEL P., ZARISKI O., *Commutative Algebra*, Tome 1, Van Nostrand.
- [28]* SAMUEL P., *Théorie algébrique des nombres*, Hermann, 1967.
- [29] SCHWERDTFEGGER, *Geometry of complex numbers*, Toronto.
- [30] SIERPINSKI W., *250 Problèmes de théorie des nombres*, Hachette.
- [31] SIERPINSKI W., *Elementary theory of numbers*, Warszawa, 1964.
- [32] WARUSFEL A., *Structures algébriques finies*, Hachette.
- [33]* WEBER H., *Lehrbuch der Algebra*, T. 1 à 3, Chelsea.

INDEX ALPHABÉTIQUE

- Abélien* (monoïde —, 65 ; groupe —, 66)
Absurde (Démonstration par l'—, 4)
Action (D'un groupe sur un ensemble, 184)
Addition (Dans \mathbb{N} , 44 ; dans $\mathbb{Z}/n\mathbb{Z}$, 122)
Affine (Application —, 561 ; droite —, 560 ; bijection —, 563 ; fonction —, 564 ; plan —, 560 ; sous-var. lin. —, 559)
d'Alembert (Théorème de —, 285 ; démonstration du théorème de —, 452)
Algèbre (— à division, 80 ; Structure d'—, 215 ; — de polynômes, 257 ; — d'endomorphismes, 209 ; — de matrices carrées, 462 ; — de type fini, 218 ; — d'un monoïde, 216 ; — de polynômes, 256, 407)
Algébrique (Nombre —, 62 ; élément —, 279, 397)
Alterné (groupe —, 173 ; polynôme —, 420 ; forme multilinéaire — *e*, 525)
Angles orientés, 242
Anneau (72 ; — commutatif, 73 ; — de Boole, 77 ; — euclidien, 263, 264 ; — factoriel, 275 ; — intègre, 75 ; — principal 116, 269 ; — quotient $\mathbb{Z}/n\mathbb{Z}$, 122 ; — quotient de $K[X]$, 308)
Antisymétrie (propriété d'—, 33)
Antisymétrique (forme —, 522 ; matrice —, 463)
Application (15 ; — composée, 16 ; — bijective, injective, surjective, 17 ; — croissante, décroissante, 34 ; — identique, 16 ; — involutive, 20 ; — linéaire, 206 ; — multilinéaire, 518 ; — affine, 561 ; — réciproque d'une bijection, 19 ; — transposée, 494)
Archimède (Propriété d'—, 53)
Argand-Cauchy (Plan d'—, 222)
Argument (— d'un nombre complexe, 232)
Arrangement (Nombre d'—s, 97)
Associative (Loi —, 63, 87 ; algèbre —, 215)
Associés, (Eléments — dans un anneau intègre, 116 ; éléments — dans $K[X]$, 260)
Automorphisme (— de monoïdes, 65 ; — de groupes, 67 ; — d'anneaux, 76 ; — intérieur, 177 ; groupe des —s, 169)
Axiome (— de fondation, 10 ; — du choix, 21 ; — de récurrence, 41)
Base (— canonique, 213 ; — duale, 498 ; — d'une K-e.v., 212 ; théorème de la — incomplète, 384)
Bell (Nombres de —, 364)
Bergers (principe des —, 95)
Bernoulli (Nombres et polynômes de —, 361)
Bernstein (Théorème de

- Bezout* (théorème de —, 129, 267, relation de —, 130)
Biadditivité, (457)
Bidual (497)
Bijection (18 ; — réciproque, 19)
Binôme (formule du —, 107)
Blocs (— d'une matrice, 459)
Boole (anneaux de — finis, 125)
Borne (— inférieure, supérieure, 37)
- Canonique* (Projection —, 30 ; injection —, 16)
Cantor (60)
Caractères (— d'un groupe, 239)
Caractéristique (— d'un corps, 143 ; déterminant —, 582 ; polynôme —, 604 ; sous-espace —, 640)
Cardan (formule de —, 437)
Cardinal (— d'un ensemble fini, 56)
Catalan (Nombre de —, 104, 360)
Cayley (théorème de — -Hamilton, 633 ; théorème de —, 170)
Centre (— d'un groupe, 186 ; — d'une algèbre, 219 ; — d'un p -groupe, 189 ; — de $GL_K(E)$, 395)
Cercles (— de \mathbb{C} , 253)
Chaîne (105 ; — simple, 49)
Chinois (Théorème —, 134)
Choix (axiome du —, 21 ; p - —, 99)
Clan (126)
Classes — à droite, à gauche, 186 ; — d'équivalence, 28 ; — de conjugaison, 177 ; — modulo n , 120 ; équation aux —, 188 ; — de similitude, 491, 492 ; — de similitude, 491, 492 ; — de similitude de nilpotents, 665)
Clos (corps algébriquement —, 285)
Codimension (510, 511)
Coefficients (— binômiaux, 107 ; — multinômiaux, 108 ; — d'une combinaison linéaire, 211 ; — d'une matrice, 454 ; — d'un polynôme, 257)
Cofacteur (537)
Collectivisante (relation —, 9)
Colonne (— d'une matrice, 453)
Comatrice (537)
Combinaison (Nombre de —s, 97 ; — avec répétitions, 99 ; — linéaire, 211 ; — A -linéaire, 115)
Commutant (— d'un endomorphisme, 628, 679)
Commutateurs (groupe des —, 175)
Commutative (Loi —, 63, 88)
Compatible (Système —, 567)
Complémentaire (— d'un ensemble, 9 ; matrice —, 537)
Composé (application —e, 16 ; —s itérés, 86 ; — de familles à support fini, 93)
Composition (Loi de —, 62)
Congruences (— dans \mathbb{Z} , 119 ; — dans $K[X]$, 306 ; — dans un anneau commutatif, 311)
Conjecture (— de Goldbach, 2 ; — de Waring, 2)
Conjonction (3)
Conjugaison (— dans un groupe, 177 ; classes de —, 177)
Conjugué (Nombre complexe —, 220, sous-groupe —, 177 ; éléments — dans un groupe, 177)
Connecteurs (— logiques, 2)
Continu (hypothèse du —, 60 ; puissance du —, 60)
Contradictoire (théorie —, 2)
Coordonnées (— relativement à une base, 213 ; fonctions —, 213)
Corestriction (17)
Corps (80 ; — des nombres complexes, 221 ; — des fractions d'un anneau intègre, 82 ; — des fractions rationnelles, 315 ; — des rationnels, 78 ; — fini, 400 ; — de nombres algébriques, 399 ; — de dissociation, 311)
Correspondance, (14)
Couple (11)
Cramer (formules de —, 573 ; systèmes de —, 570)
Crible (Formule du —, 104)
Crochet (216)
Croissante (application —, 34)
Cycles (décomposition d'une permutation en —, 180 ; — conjugués, 181)
Cyclique (Groupe —, 160 ; endomorphisme —, 653)
Cyclotomique (polynôme —, 289, 401)

- Décomposition* (— d'un entier en facteurs premiers, 146 ; — d'une permutation en cycles, 180 ; — d'une fraction rationnelle en éléments simples, 322 ; — d'un polynôme en facteurs irréductibles, 273 ; — dans $\mathbb{R}[X]$, 303 ; — canonique d'une application, 30 ; — canonique d'une application linéaire, 509)
- Décroissante* (application —, 34)
- Degré* (— d'un polynôme, 258, 410 ; d'une fraction rationnelle, 319 ; — partiel, total, 415 ; — d'un nombre algébrique, 279)
- De Moivre* (Formule de —, 227)
- Dénombrable* (Ensemble —, 59)
- Dénombrement* (95)
- Dérangements* (Nombre de —, 357)
- Dérivée* (— d'un polynôme, 296 ; — d'une fraction rationnelle, 338 ; — d'une série formelle, 349 ; — partielle, 414 ; — d'une forme multilinéaire, 520)
- Descente infinie* (Principe de — de Fermat, 51)
- Déterminant* (— caractéristique, 582 ; — de matrices carrées, 533 ; — d'endomorphismes, 529 ; — de n vecteurs dans une base, 528 ; — de Vandermonde, 546 ; — principal, 578 ; — de Sylvester, 585)
- Développement* (— d'un déterminant, 537 ; — en série formelle d'une fraction rationnelle, 351)
- Diagonale* (— d'un produit cartésien, 12 ; matrice —, 461 ; matrice — par blocs, 461)
- Diagonalisable* (endomorphisme —, matrice —, 618, 639)
- Diagramme* (— commutatif, 31)
- Diédral* (groupe —, 161, 201)
- Différence* (— d'ensembles, 13 ; — symétrique, 66)
- Dimension* (— d'un espace vectoriel, 381 ; espace de — finie, 378 ; théorème de la — finie, 380)
- Directe* (Somme interne —, 373)
- Direction* (— d'une variété affine, 560)
- Dirichlet* (Théorème de —, 141)
- Discriminant* (— d'un polynôme, 435 ; — général, 420)
- Disjonction* (3 ; — des cas, 4)
- Dissocié* (polynôme —, 286)
- Distingué* (sous-groupe —, 192)
- Distributive* (Loi —, 73)
- Diviseur de zéro* (74)
- Divisibilité* (119)
- Division* (algèbre à —, 80 ; — euclidienne des entiers, 53 ; — euclidienne dans $K[X]$, 261 ; — en les puissances croissantes, 326)
- Domaine* (— d'opérateurs, 202)
- Dominant* (terme —, coefficient —, 258)
- Drapeau* (611)
- Droite* (— vectorielle, 382 ; — principale d'une similitude indirecte, 246)
- Dual* (— algébrique d'un K -e.v., 387, 497 ; propriété —e, 26 ; — duale, 498)
- Echange* (théorème d'—, 382)
- Eisenstein* (critère d'—, 277)
- Élément* (— inversible, 75 ; — inversible de $\mathbb{Z}/n\mathbb{Z}$, 135 ; — irréductible, 139 ; — maximal, 36 ; — maximum, 36 ; — neutre, 45 ; — nilpotent, 111, 389 ; — unité, 73 ; — régulier, 74 ; —s simples d'une fraction rationnelle, 322)
- Endomorphisme* (— d'un monoïde, 65 ; — diagonalisable, 619 ; — nilpotent, 632 ; — transposé, 494 ; — trigonalisable, 612)
- Engendré* (groupe —, 154 ; idéal —, 115 ; sous- K -ev —, 211)
- Ensemble* (7 ; — d'arrivée, de départ, de définition, 14 ; — des parties, 11 ; — dénombrable, 59 ; — fini, 54 ; — infini, 54 ; — ordonné, 33 ; — quotient, 132)
- Entiers* (— naturels, 41 ; — relatifs (ou rationnels), 70 ; — algébriques, 270 ; — de Gauss, 223)
- Equation* (— algébrique, 434 ; — aux classes, 188 ; —s linéaires, 559)
- Equipotents* (ensembles —, 58)
- Equivalence* (relation d'—, 28 ; classe d'—, 28)
- Equivalentes* (assertions —, 3 ; matrices —, 484)

Espace (— produit, 369 ; — projectif, 29 ; — quotient, 507 ; — vectoriel, 202)

Euclide (5 ; algorithme d'—, 128 ; théorème d'—, 140)

Euclidien (anneau — $K[X]$, 261)

Euclidien (anneau —, 264)

Euler (indicateur d'—, 136, 148 ; formules d'—, 227 ; identité d'—, 415 ; théorème d'—, 137 ; critère d'—, 295 ; polynômes d'—, 302)

Exponentielles (— d'endomorphismes, 614 ; — de matrices, 623, 646)

Exponentiation (— dans \mathbb{N} , 48)

Exposant (— d'un groupe abélien fini, 166)

Extension (— d'un corps commutatif, 81 ; — algébrique simple, 399)

Extrémal (Elément —, 36)

Facteurs invariants (— d'une matrice carrée, 662)

Factoriel (anneau —, 275)

Factorisation (voir aussi *décomposition*, 303)

Famille (22 ; — à support fini, 93 ; — indexée, 22 ; intersection et union de —s, 24 ; — génératrice, libre liée, 212 ; produit d'une — d'ensembles, 24 ; — sommable dans $K[X]$, 345)

Fermat (Nombres de —, 149, 290 ; petit théorème de —, 141 ; grand théorème de, 2, 432 ; descente infinie de —, 51)

Ferrari (méthode de —, 446)

Fibonacci (suite de —, 138, 359)

Fidèlement (groupe opérant —, 185)

Fini (ensemble —, 54)

Fonction (15 ; — affine, 564 ; — coordonnée, 213 ; — polynôme, 281 ; rationnelle, 336 ; — monômiale, 421 ; — symétrique élémentaire, 291 ; — polynômiale, 405)

Fonctionnel (graphe —, 15)

Forme (— trigonométrique d'un nombre complexe, 233 ; — multilinéaire, 522 ; — linéaire, 387 ; — bilinéaire canonique, 494 ; — réduite de Jordan, 666)

Fractions (corps des — d'un anneau intè-

gre, 83 ; — rationnelles, 315 ; décomposition des — rationnelles, 317)

Frobenius (méthode de —, 587)

Fubini (principe de —, 89)

Gauss (entiers de —, 223 ; méthode de —, 594 ; sommes de —, 637 ; théorème de —, 130, 268)

Générateurs (— d'un idéal, 115)

Génératrice (famille —, partie —, 212 ; partie — d'un sous-groupe, 154)

Glisseurs (30)

Gödel (2)

Goldbach (conjecture de —, 2)

Graphe (14 ; — fonctionnel, 15)

Groupe (66 ; — abélien, 66 ; — alterné, 173 ; — circulaire, 255 ; — cyclique, 113 ; diédral, 201 ; — monogène, 113 ; — de permutations, 168 ; — d'isotropie, 186 ; — des unités, 75 ; — de type fini, 155 ; — linéaire, 209, 394 ; — produit, 156 ; — quotient, 194 ; — simple, 192 ; — symétrique, 170 ; — spécial linéaire, 531)

Hadamard (Théorème de —, 141)

Hamilton (Théorème de Cayley —, 633)

Héréditaire (Propriété —, 45)

Hilbert (polynômes de —, 362)

Homogène (polynôme —, 410 ; système linéaire —, 568)

Homographique (fraction rationnelle —, 341)

Homomorphisme (— de groupes, 67 ; — d'anneaux, 75 ; — de K -algèbres, 217 ; — exponentiel, 226 ; — de substitution, 278, 407)

Homothétie (203)

Hyperplan (— affine, 565 ; — vectoriel, 391)

Idéal (— bilatère, 114 ; — engendré, 115 ; — maximal, 117 ; — principal, 115 ; — à droite, à gauche, 218 ; — primaire, 118 ; — premier, 118 ; — de type fini, 115).

Identique (application —, 16)

- Identité* (— d'Euler, 415 ; principe de prolongement des —s algébriques, 412)
- Image* (— d'une application, 16 ; — d'une application linéaire, 207 ; — directe, réciproque d'une partic, 16)
- Implication* (3)
- Inclusion* (8)
- Incompatible* (système linéaire —, 567)
- Inconnue* (— principale, 578)
- Indécidable* (assertion —, 3)
- Indépendance linéaire* (— de vecteurs, 212 ; — de sous-espaces vectoriels, 373)
- Indéterminée* (258)
- Indicateur* (— d'Euler, 136, 148 ; — de torsion, 312)
- Indice* (ensemble d'—s, 22 ; — d'un sous-groupe, 163 ; — d'une valeur propre, 644)
- Indicielle* (notation — d'une suite, 23)
- Induite* (application —, 17 ; loi de groupe —, 67 ; relation d'équivalence —, 31 ; relation d'ordre —, 33)
- Infini* (ensemble —, 54)
- Injection* (— canonique, 16)
- Injective* (application —, 17)
- Intègre* (anneau — 175 ; algèbre —, 216)
- Intersection* (— de deux ensembles, 13 ; — d'une famille d'ensembles, 24)
- Intervalles* (— dans \mathbb{N} , 52 ; — dans \mathbb{Z} , 71)
- Invariant* (sous-groupe, —, 192 ; facteurs —s, 662)
- Inverse* (— d'un élément dans un anneau, 75 ; — d'un élément dans un monoïde, 64 ; fonction — (= bijection réciproque), 19)
- Inversible* (élément —, 74 ; matrice carrée —, 462)
- Inversions* (nombre d'— d'une permutation, 170)
- Involution* (20 ; nombre d'—s dans \mathfrak{S}_n , 359)
- Irréductible* (élément — d'un anneau intègre, 139 ; polynôme —, 271 ; représentant — d'un rationnel, 133)
- Isométrie* (cf. *Tomes de Géométrie et d'Analyse* ; ici, 245)
- Isomorphisme* (— de monoïdes, 65 ; — de groupes, 67 ; — d'anneaux, 76 ; — d'espaces vectoriels, 206 ; — de K -algèbres, 217 ; — de corps, 81)
- Isotropie* (groupe d'—, 186)
- Itérées* (— d'une application, 46 ; — d'une loi de composition, 86)
- Jacobi* (formules de —, 549)
- Jordan* (*Camille*) (matrice de —, 660 ; réduction de —, 660)
- Jordan* (méthode de —, 596 ; *N.B. : il ne s'agit pas de Camille Jordan, mais d'un autre mathématicien*)
- Kronecker* (symbole de —, 213)
- Lagrange* (théorème de —, 164 ; interpolation de —, 283)
- Lambert* (Série de —, 361)
- Laplace* (formule de —, 541)
- Legendre* (symbole de —, 301)
- Leibniz* (formule de —, 296)
- Lexicographique* (ordre —, 39)
- Libre* (famille —, 212 ; partie —, 212 ; éléments algébriquement —s, 279, 406)
- Lié* (famille —e, 212 ; partie —e, 212 ; éléments algébriquement —s, 276)
- Linéaire* (application —, 206 ; combinaison —, 211 ; forme —, 387 ; indépendance —, 212, 373 ; partie — d'une application affine, 562)
- Loi de composition* (62 ; — externe, 202 ; — induite, 67 ; — interne, 62 ; — quotient, 123)
- Longueur* (— d'un cycle, 178)
- Majorant* (37)
- Matrice* (453 ; — antisymétrique, 463 ; — bistochastique, 468 ; — bordante, 483 ; — carrée, 453 ; déterminant de — circulante, 548 ; — colonne, ligne, 453 ; — complémentaire, 537 ; — de dilatation, 487 ; — de Jordan, 660 ; — de passage, 476 ; — de permutation, 469 ; — diagonale, 461 ; — diagonalisable, 619 ; — d'une application linéaire, 471 ; — élémentaire, 455 ; —s équivalentes, 484 ; — inversible, 475 ; — magique, 506 ; —

nilpotente, 469 ; — produit, 456 ; — scalaire, 462 ; —s semblables, 491 ; — stochastique, 650 ; — symétrique, 463 ; — transposée, 458 ; — de transvection, 487 ; — trigonale, 463 ; — trigonale par blocs, 539 ; — trigonalisable, 610 ; — unipotente, 463)

Maximal (élément —, 36 ; idéal —, 117)

Maximum (élément —, 35)

Mersenne (nombres de —, 150)

Mineurs (— d'une matrice, 536 ; — centrés, 537, — principaux, 537, 591)

Minimal (élément —, 36 ; polynôme —, 279)

Minimum (élément —, 35)

Minorant (37)

Möbius (fonction de —, 361)

Module (— d'une congruence, 120 ; — d'un nombre complexe, 220)

Monogène (groupe —, 160 ; sous-espace —, 673)

Monoïde (64)

Moivre (formule de De —, 227)

Monoïde (application —, 34)

Morgan (Lois de De —, 13)

Morley (théorème de —, 244)

Morphisme (65)

Multilinéaire (application —, 518 ; forme —, 518)

Multinôme (formule du —, 107)

Multiplication (— dans \mathbb{N} , 47 ; — dans $\mathbb{Z}/n\mathbb{Z}$, 122)

Multiplicité (— d'une racine, 286 ; — d'une valeur propre, 605 ; — d'un pôle, 320)

Naturel (entier —, 41 ; ordre —, 50)

Négation (2)

Neutre (élément —, 45, 48, 63)

Newton (formules de —, 428)

Nilpotent (élément —, 111 ; endomorphisme —, 632 ; matrice —*e*, 469)

Nilradical (117)

Normal (sous-groupe —, 192)

Normalisateur (— d'un sous-groupe, 187)

Normalisé (polynôme —, 258)

Norme (271)

Noyau (— d'un homomorphisme de groupes, 69 ; — d'une application linéaire, 207 ; lemme des —*x*, 638)

Numération (— des entiers, 94, 151 ; — anglaise, 94)

Opérateur (synonyme d'*endomorphisme*)

Opérations élémentaires (486)

Opération (— d'un groupe sur un ensemble, 184)

Opposé (— d'un élément, dans un groupe abélien, 64)

Orbite (185)

Ordonné (ensemble —, 33 ; ensemble bien —, 35, 50)

Ordre (relation d'—, 33 ; — produit, 38 ; — lexicographique, 39 ; — d'un élément dans un groupe, 112 ; — naturel, 49, 71)

Orthogonal (495)

Parallèles (variétés affines —, 560)

Parfait (nombre —, 150)

Partage (27)

Partie (— d'un ensemble, 8 ; — entière, *P*-fractionnaire d'une fraction rationnelle, 324 ; — génératrice, libre d'un espace vectoriel, 213 ; — polaire, 325 ; — vide, 10 ; — réelle, imaginaire d'un nombre complexe, 221)

Partiel (relation d'ordre —, 33)

Partielles (dérivées —, 414)

Partition (29)

Pascal (relation de —, 98 ; triangle de —, 99)

Peano (4 ; axiomes de, 41)

Période (— de *a modulo n*, 138)

Permutables (éléments —, 92)

Permutation (18 ; — circulaire, 178 ; groupe de —s, 170 ; — paire, impaire, 173 ; nombre de —s, 97)

PGCD (— d'entiers, 127 ; — de polynômes, 265)

Pivot (méthode du — partiel, 597)

Pôle (— d'une fraction rationnelle, 320)

Polynôme (256 ; — alterné, 420 ; — en *n* lettres, 405 ; — caractéristique, 604 ; —

- cyclotomique, 289 ; — d'endomorphisme, 613 ; — dissocié, 286 ; fonction —, 281 ; — homogène, 410 ; — irréductible, 271 ; — minimal, 630 ; — normalisé, 258 ; — primitif, 276 ; — symétrique, 420)
- PPCM** (— d'entiers, 126 ; — de polynômes, 264)
- Prédécesseur** (41)
- Premier** (— élément d'un ensemble bien ordonné, 35 ; nombre —, 139 ; idéal —, 118 ; sous-corps —, 142)
- Premiers entre eux** (entiers —, 129 ; polynômes —, 267)
- Préordre** (119)
- Primaire** (idéal —, 118)
- Primitive** (— d'une fraction rationnelle, 340 ; racine — de l'unité, 235)
- Principal** (anneau —, 116 ; déterminant —, 578 ; équations —es, 578 ; idéal —, 115 ; inconnues —es, 578)
- Produit** (— cartésien d'ensembles, 11 ; — d'espaces vectoriels, 369 ; — d'une famille d'ensembles, 24 ; — de groupes, 156 ; — de matrices, 456 ; — par blocs (matrices), 459 ; — de polynômes, 256, 408 ; — lexicographique, 39)
- Projecteur** (367)
- Projection** (première et deuxième — d'un graphe, 14 ; — canonique, 30 ; — sur un sous-espace vectoriel, 367)
- Prolongement** — d'une application, 17 ; principe de — des identités algébriques, 412)
- Propre** (Valeur —, 603 ; vecteur —, 603 ; sous-espace —, 616)
- Puissance** (— du dénombrable, 59 ; — du continu, 60 ; —s d'une matrice, 645 ; somme de —s, 101)
- Quadratfrei** (401)
- Quadratique** (entier algébrique —, 270 ; loi de réciprocité —, 300)
- Quantificateur** (— existentiel, 5 ; — universel, 6)
- Quaternions** (558 ; groupe —ique, 200)
- Quotient** (ensemble —, 13, 28 ; — euclidien, 53 ; — exact, 48 ; groupe —, 194 ; espace vectoriel —, 508 ; anneau — par un idéal, 311)
- Racines** (— carrées d'un nombre complexe, 224 ; — n -ièmes, 234 ; — d'un polynôme, 282 ; — d'une équation, 434 ; — primitives de l'unité, 235 ; — simples, multiples, 286 ; *racine* d'un idéal, 117)
- Radicaux** (équation résoluble par —, 434)
- Rang** (— d'un système de vecteurs, 388 ; — d'une application linéaire, 386 ; — d'une matrice, 481 ; — d'un système linéaire, 569 ; formule du —, 385)
- Rationnelle** (fraction —, 315 ; fonction —, 335)
- Réciproque** (bijection —, 19 ; équation —, 447)
- Réciprocité** (loi de — quadratique, 300)
- Recouvrement** (27)
- Récurrence** (démonstration par —, 42 ; — transfinie, 51 ; relation de — linéaire à coefficients constants, 624, 654)
- Réduction** (— de Jordan, 660)
- Régulier** (élément — 63, 74)
- Relation** (— d'équivalence, 28 ; — d'ordre, 33 ; — collectivisante, 9 ; — entre coefficients et racines, 292)
- Répartition** (— des nombres premiers, 140)
- Repère cartésien** (239)
- Représentant** (— irréductible d'un rationnel, 153 ; — irréductible d'une fraction rationnelle, 316)
- Résidu** (— d'une fraction rationnelle en un pôle, 325 ; — quadratique mod (p) , 295)
- Résolvante** (445)
- Reste** (— dans la division euclidienne d'entiers, 53 ; — dans la division euclidienne de polynômes, 262 ; — dans la division en puissances croissantes, 326 ; — mod (n) , 121)
- Restriction** (— d'une application, 17 ; — des scalaires, 216)
- Résultant** (585)
- Réunion** (— de deux ensembles, 12 ; — d'une famille d'ensembles, 24)

Réversion (— d'une série formelle, 350)
Rouché-Fontené (théorème de —, 581)

Sarrus (règle de —, 545)

Scalaire (202)

Semi-simple (endomorphisme —, 679)

Séquence (153)

Série formelle (344 ; —(s) usuelles, 355)

Signature (— d'une permutation, 170)

Similitude (— vectorielle, 245 ; — de matrices, 492)

Simple (groupe —, 192)

Simplifiable (élément —, 63)

Simson (droite de —, 255)

Singleton (8)

Sommable (famille — de séries formelles, 345)

Somme (— d'idéaux, 116 ; — de sous-espaces vectoriels, 372 ; — directe interne de sous-espaces vectoriels, 515 ; ensemble —, 26 ; — de deux carrés, 302 ; — directe externe d'espaces vectoriels, 371 ; — directe externe de groupes abéliens, 156)

Sous (— algèbre, 217 ; — anneau, 76 ; — corps, 80 ; — ensemble ordonné, 33 ; — espace vectoriel, 204 ; — espace propre, 603 ; — espace caractéristique, 640 ; — espaces supplémentaires, 365 ; — famille, 22 ; — groupe, 67 ; — groupe conjugué, 177 ; — groupe distingué, 192 ; — groupe engendré, 154 ; structure — jacente, 215 ; — matrice, 455 ; — variété linéaire affine, 559)

Soustraction (— dans \mathbb{N} , 48)

Spectrale (valeur —, 603)

Spectre (603)

Spernérienne (famille —, 104)

Stabilisateur (186)

Stable (partie —, 67 ; sous-espace —, 637)

Stathme euclidien (263, 264)

Stationnaire (suite —, 23)

Steinitz (théorème de —, 286)

Stirling (nombres de —, 363)

Structure (— quotient, 194 ; voir *groupe*, *anneau*, *corps*, *espace vectoriel*, *algèbre*, *ensemble ordonné*)

Substitution (278, 407)

Successeur (application —, 41)

Suite (23 ; — extraite, 23 ; — stationnaire, 23)

Suite finie (23)

Superposée (— de séries formelles, 346)

Supplémentaires (Sous-espaces —, 365)

Support (— premier, 145 ; — fini, 93 ; — d'un cycle, 178)

Syllogisme (3)

Sylvester (déterminant et matrice de —, 585)

Symétrique (— d'un élément, 64 ; fonction — élémentaire, 292 ; groupe —, 170 ; fonctions —s, 418 ; forme n -linéaire —, 522)

Symétrisable (élément —, 64)

Système (— d'équations linéaires, 568 ; — de Cramer, 570 ; — homogène, 568 ; — mal conditionné, 600 ; — complet mod (n) , 121 ; — compatible, incompatible, 567)

Taylor (formule de —, 296, 416)

Tchebychev (polynômes de —, 287)

Torsion (élément de, ou sans —, 112 ; indicateur de —, 312)

Trace (— d'une matrice, 466 ; — d'un endomorphisme, 479)

Transcendant (élément —, 279 ; nombre —, 279)

transformée (— d'une équation, 450)

Transitivement (groupe opérant —, 185)

Transitivité (— d'une relation, 28 ; — des indices, 164)

Translation (— à gauche ou à droite, 162 ; — du plan affine \mathbb{C} , 240 ; — d'un espace vectoriel, 564)

Transposé (endomorphisme —, 495 ; — d'une matrice, 458)

Transposition (172)

Treillis (40)

Trigonalisation (610 ; algorithme de —, 652)

Trisection (— d'un angle, 439)

Type fini (idéal de —, 115 ; groupe de —, 155 ; K -algèbre de —, 218)

- u-monogène* (espaces —, 661)
u-stable (sous-espace vectoriel —, 606)
Unifère (algèbre —, 215)
Unitaire (polynôme —, 258)
Unité (élément —, 73 ; groupe des —, 75)
p-uple (11)
- Valeur* (— absolue dans \mathbb{Z} , 71 ; — propre d'endomorphisme, 603 ; — propre de matrice carrée, 608 ; — spectrale, 603)
Valuation (*p*- — d'un entier, 145 ; — d'un polynôme, 258, 410 ; — d'une série formelle, 344 ; *P*- — d'un polynôme, 272 ; *P*- — d'une fraction rationnelle, 319)
Vandermonde (matrice de —, 504 ; déter-
- minant de —, 546 ; polynôme de —, 420)
Vecteur (— -ligne, -colonne, 453 ; — propre, 603 ; — glissant, 30)
Vectoriel (espace —, 202 ; sous-espace —, 204 ; plan —, 382 ; droite —le, 382)
Vide (ensemble —, 9 ; application —, 16)
- Waring* (formule de —, 430 ; conjecture de —, 2)
Wedderburn (théorème de —, 390)
Wilson (théorème de —, 142)
- Zermelo* (théorème de —, 40)
Zéro (— d'un polynôme, 282 ; — d'une fraction rationnelle, 320)
Zorn (théorème de —, 40)

MATHÉMATIQUES

Cours de mathématiques, par J. LELONG-FERRAND et J.-M. ARNAUDIES.

Tome 1 : *Algèbre*. Tome 2 : *Analyse*. Tome 3 : *Géométrie et cinématique*.

Tome 4 : *Équations différentielles, intégrales multiples*.

Exercices résolus d'analyse, par J. LELONG-FERRAND.

Exercices et problèmes résolus d'algèbre, par L. CHAMBADAL.

Formulaire de mathématiques, par L. CHAMBADAL.

Mathématiques, rappels de cours et exercices résolus, par F. DELMER.

Cours de mathématiques 1. *Algèbre*. 2. *Analyse*. 3. *Compléments d'analyse*. 4. *Algèbre bilinéaire et géométrie*, par J.-M. ARNAUDIES. et H. FRAYSSE.

Analyse, exercices corrigés, T. 1 et 2, par J.-M. MONIER.

Algèbre, exercices corrigés, T. 1 et 2, par J.-M. MONIER.

PHYSIQUE

Formulaire de physique, par J. RENAULT.

Turbo Pascal élémentaire pour la physique, par J. RENAULT

Cours de physique, par M. BERTIN, J.-P. FAROUX et J. RENAULT.

- *Thermodynamique*.
- *Mécanique 1 : Systèmes de points et notions de relativité*.
- *Électromagnétisme 1 : Électrostatique ; milieux conducteurs*.
- *Électromagnétisme 2 : Électrocinétique et éléments d'électronique*.
- *Mécanique 2 : Éléments de mécanique des solides et des fluides*.
- *Optique et physique ondulatoire : Optique géométrique et optique physique, phénomènes de propagation*.
- *Électromagnétisme 3 : Magnétostatique, induction, équations de Maxwell et compléments d'électronique*.
- *Électromagnétisme 4 : Milieux diélectriques et milieux aimantés*.

Exercices de mécanique, par J. RENAULT.

Exercices de thermodynamique, par M. BERTIN et J. RENAULT

Exercices d'électromagnétisme 1 : Électrostatique, électrocinétique et notions d'électronique, par J. RENAULT

Exercices d'électromagnétisme 2 : Magnétostatique et induction. Équations de Maxwell et ondes électromagnétiques, par J. RENAULT.

Exercices d'optique et de physique ondulatoire, par J. RENAULT.

Problèmes résolus d'électrostatique et dynamique des particules chargées, par H. LUMBROSO.

Problèmes résolus de mécanique du point et des systèmes de points, par H. LUMBROSO.

Problèmes résolus sur les circuits électriques, par H. LUMBROSO.

Problèmes résolus d'électronique, par H. LUMBROSO.

Problèmes résolus de mécanique des fluides, par H. LUMBROSO.

Problèmes résolus sur les ondes électromagnétiques, par H. LUMBROSO.

Cours de physique, par Y. DULAC. (Prépa Vété).

Exercices corrigés de physique, par Y. DULAC.

Les bases de l'électromagnétisme, par M. HULIN et J.-P. MAURY.

Équations de Maxwell. Ondes électromagnétiques dans le vide, par M. HULIN, N. HULIN et D. PERRIN.

CHIMIE

La chimie. Dictionnaire encyclopédique, par J. ANGENAULT.

Exercices et problèmes résolus de chimie, par M. LAFFITTE.

Cours de chimie organique, par P. ARNAUD.

Cours de chimie physique, par P. ARNAUD.

Exercices de chimie organique, par P. ARNAUD.

Cours de chimie minérale, par M. BERNARD.

Cours de chimie, T. 1, 1^{re} année, T. 2, 2^e année, par J. BOTTIN, J.-C. MALLET

Exercices et problèmes corrigés de chimie, par J. BOTTIN, et J.-C. MALLET (2

Formulaire de chimie générale, par F. DUPARC.

Le **COURS DE MATHÉMATIQUES** réunit toutes les notions de base de l'algèbre fondamentale et de l'algèbre linéaire indispensables tant aux concours d'entrée aux grandes écoles que pour entreprendre des études scientifiques à dominante mathématique.

De conception très élaborée, ce livre se veut avant tout un outil de travail. Plus de mille exercices qui collent au texte paragraphe par paragraphe, près de trois cents exemples développés, permettent une lecture active et une assimilation progressive.

Le texte est structuré pour rendre le repérage facile et rapide : les notions de base sont introduites progressivement. Les sujets enseignés en première années de classe préparatoire peuvent être étudiés séparément.

Les futurs élèves des grandes écoles, mais aussi les candidats à l'agrégation et les professeurs de lycées trouveront, en petits caractères, tous les approfondissements désirables.

Ce Cours de mathématiques se compose de 4 tomes :

1. Algèbre
2. Analyse
3. Compléments d'analyse
4. Algèbre bilinéaire et géométrie

Ce tome 1 propose 195 définitions, 305 théorèmes avec leur démonstration, 260 exemples et 1 093 exercices, du plus simple au plus élaboré.



ISBN 2-04-016450-2

